

2 Interpolation und Approximation

Ein Grundproblem der numerischen Praxis ist die Darstellung und Auswertung von Funktionen. Dabei ergeben sich folgende Aufgabenstellungen:

- (i) Eine Funktion $f(x)$ ist nur in einer diskreten Menge von Argumenten x_0, \dots, x_n bekannt und soll mit dieser Information rekonstruiert werden (z. B. zur graphischen Darstellung oder zur Auswertung an Zwischenstellen).
- (ii) Eine analytisch gegebene Funktion $f(x)$ soll auf der Rechenanlage so dargestellt werden, dass jederzeit Funktionswerte zu beliebigem Argument x leicht berechnet werden können (z. B. trigonometrische Funktionen).

In beiden Fällen hat man ein System mit unendlich vielen Freiheitsgraden, nämlich die funktionale Abhängigkeit $y = f(x)$, durch einen endlichen Datensatz zu simulieren. Hierzu bedient man sich gewisser Klassen P von einfach strukturierten Funktionen; z. B.:

$$\begin{array}{ll}
 \text{Polynome:} & p(x) = a_0 + a_1x + \dots + a_nx^n, \\
 \text{rationale Funktionen:} & r(x) = \frac{a_0 + a_1x + \dots + a_nx^n}{b_0 + b_1x + \dots + b_mx^m}, \\
 \text{trigonometrische Polynome:} & t(x) = \frac{1}{2}a_0 + \sum_{k=1}^n \{a_k \cos(kx) + b_k \sin(kx)\}, \\
 \text{Exponentialsummen:} & e(x) = \sum_{k=1}^n a_k \exp(b_kx).
 \end{array}$$

Definition 2.1: *Geschieht die Zuordnung eines Elementes $g \in P$ zur Funktion f durch Fixieren von Funktionswerten*

$$g(x_i) = y_i := f(x_i), \quad i = 0, \dots, n,$$

so spricht man von „Interpolation“. Ist $g \in P$ als in einem gewissen Sinne „beste“ Darstellung von f zu bestimmen, z. B.:

$$\max_{a \leq x \leq b} |f(x) - g(x)| \quad \text{minimal für } g \in P,$$

oder

$$\left(\int_a^b |f(x) - g(x)|^2 dx \right)^{1/2} \quad \text{minimal für } g \in P,$$

so spricht man allgemein von „Approximation“. Die jeweilige Wahl der Konstruktion von $g \in P$ hängt von der zu erfüllenden Aufgabe ab. Offenbar ist die Interpolation eine spezielle Art der Approximation mit

$$\max_{i=0, \dots, n} |f(x_i) - g(x_i)| \quad \text{minimal für } g \in P.$$

2.1 Polynominterpolation

Wir bezeichnen mit P_n den Vektorraum der Polynome vom Grad kleiner oder gleich n :

$$P_n := \{p(x) = a_0 + a_1x + \dots + a_nx^n \mid a_i \in \mathbb{R}, i = 0, \dots, n\}.$$

Definition 2.2: Die sog. „Lagrange¹-Interpolationsaufgabe“ besteht darin, zu $n + 1$ paarweise verschiedenen Stützstellen (auch „Knoten“ genannt) $x_0, \dots, x_n \in \mathbb{R}$ und gegebenen Knotenwerten $y_0, \dots, y_n \in \mathbb{R}$ ein Polynom $p \in P_n$ zu bestimmen mit der Eigenschaft

$$p(x_i) = y_i, \quad i = 0, \dots, n. \quad (2.1.1)$$

Satz 2.1 (Lagrange-Interpolation): Die Lagrange-Interpolationsaufgabe ist eindeutig lösbar.

Beweis: Wir zeigen zunächst die Eindeutigkeit. Sind $p_1, p_2 \in P_n$ zwei Lösungen, so gilt für $p := p_1 - p_2 \in P_n$: $p(x_i) = 0$, $i = 0, \dots, n$, d. h.: p hat $n + 1$ Nullstellen, und ist folglich identisch Null (folgt mit Hilfe des Satzes von Rolle²). Zur Existenz betrachten wir die Gleichungen $p(x_i) = y_i$, $i = 0, \dots, n$. Dies kann man interpretieren als ein lineares Gleichungssystem mit $n + 1$ Gleichungen für die $n + 1$ unbekanntenen Koeffizienten a_0, \dots, a_n von $p \in P_n$. Wegen der Eindeutigkeit von p muss dieses System dann notwendig eine Lösung haben. Q.E.D.

Zur Konstruktion des Interpolationspolynoms $p \in P_n$ verwendet man etwa die sog. „Lagrange-Basispolynome“

$$L_i^{(n)}(x) := \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j} \in P_n, \quad i = 0, \dots, n.$$

Dass der Satz von Polynomen $\{L_i^{(n)}, i = 0, \dots, n\}$ tatsächlich eine Basis von P_n ist, wird als Übungsaufgabe gestellt. Offenbar ist

$$L_i^{(n)}(x_k) = \begin{cases} 1, & \text{falls } i = k \\ 0, & \text{falls } i \neq k \end{cases} =: \delta_{ik} \quad (\text{Kronecker-Symbol}).$$

Definition 2.3: Das Polynom

$$p := \sum_{i=0}^n y_i L_i^{(n)} \in P_n \quad (2.1.2)$$

¹Joseph Louis de Lagrange (1736–1813): Französischer Mathematiker; 1766–87 Direktor der mathem. Klasse der Berliner Akademie, dann Professor in Paris; bahnbrechende Arbeiten zur Variationsrechnung, zur komplexen Funktionentheorie sowie zur theor. Mechanik und Himmelsmechanik.

²Michel Rolle (1652–1719): Französischer Mathematiker und Autodidakt; wirkte in Paris und leistete Beiträge zur Analysis, Algebra und Geometrie; der nach ihm benannte Satz wurde 1691 publiziert.

hat dann die gewünschten Eigenschaften $p(x_j) = y_j$. Es wird die „Lagrange-Darstellung“ des Interpolationspolynoms zu den Stützpunkten $(x_0, y_0), \dots, (x_n, y_n)$ genannt (abgekürzt „Lagrange-Interpolationspolynom“).

Die Lagrange-Darstellung des Interpolationspolynoms hat den Nachteil, dass sich die verwendeten Basisfunktionen von P_n bei Hinzunahme eines weiteren Stützpunktes (x_{n+1}, y_{n+1}) völlig ändern. Dies wird vermieden bei Verwendung der sog. „Newton³-Basispolynome“

$$N_0(x) := 1; \quad i = 1, \dots, n: \quad N_i(x) := \prod_{j=0}^{i-1} (x - x_j).$$

Aus diesen lässt sich das Interpolationspolynom systematisch aufbauen. Für den Ansatz

$$p(x) = \sum_{i=0}^n a_i N_i(x)$$

findet man durch sukzessive Auswertung in x_0, \dots, x_n das gestaffelte Gleichungssystem

$$\begin{aligned} y_0 &= p(x_0) = a_0 \\ y_1 &= p(x_1) = a_0 + a_1(x_1 - x_0) \\ &\vdots \\ y_n &= p(x_n) = a_0 + a_1(x_n - x_0) + \dots + a_n(x_n - x_0) \dots (x_n - x_{n-1}), \end{aligned}$$

woraus sich die Koeffizienten a_i rekursiv berechnen lassen. Die Hinzunahme eines weiteren Punktes (x_{n+1}, y_{n+1}) ist nun leicht durch Fortsetzung dieses Prozesses mit der Basisfunktion N_{n+1} zu bewerkstelligen. In der Praxis bestimmt man die a_i jedoch auf eine andere, numerisch stabilere Weise, die im Folgenden beschrieben wird.

Satz 2.2 (Newton-Darstellung): Das Lagrange-Interpolationspolynom zu den Punkten $(x_i, y_i), i = 0, \dots, n$, lässt sich bzgl. der Newton-Polynombasis schreiben in der Form

$$p(x) = \sum_{i=0}^n y[x_0, \dots, x_i] N_i(x). \quad (2.1.3)$$

Dabei bezeichnen $y[x_0, \dots, x_i]$ die zu den Punkten (x_i, y_i) gehörenden sog. „dividierten Differenzen“, welche rekursiv definiert sind durch

$$\begin{aligned} i = 0, \dots, n: & \quad y[x_i] := y_i \\ k = 1, \dots, n - i: & \quad y[x_i, \dots, x_{i+k}] := \frac{y[x_{i+1}, \dots, x_{i+k}] - y[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i}. \end{aligned}$$

³Isaac Newton (1643–1727): Englischer Physiker und Mathematiker; Professor an der Universität Cambridge; entwickelte u. a. die Grundlagen der klassischen Mechanik und der Differentialrechnung.

Beweis: Es bezeichne $p_{i,i+k} \in P_k$ das Polynom, das die Punkte $(x_i, y_i), \dots, (x_{i+k}, y_{i+k})$, interpoliert. Speziell ist also $p_{0,n} = p$ das gesuchte Interpolationspolynom. Wir zeigen

$$p_{i,i+k}(x) = y[x_i] + y[x_i, x_{i+1}](x - x_i) + \dots + y[x_i, \dots, x_{i+k}](x - x_i) \dots (x - x_{i+k-1}),$$

was offensichtlich die Aussage des Satzes als Spezialfall beinhaltet. Der Beweis wird durch Induktion bzgl. der Indexdifferenz $k = (i+k) - i$ geführt. Für $k = 0$ ist $p_{i,i} = y_i = y[x_i]$, $i = 0, \dots, n$. Sei die Behauptung richtig für $k - 1 \geq 0$. Konstruktionsgemäß gilt nun

$$p_{i,i+k}(x) = p_{i,i+k-1}(x) + a(x - x_i) \dots (x - x_{i+k-1}).$$

Zu zeigen ist also, dass $a = y[x_i, \dots, x_{i+k}]$. Offenbar ist a der Koeffizient von x^k in $p_{i,i+k}(x)$. Nach Induktionsannahme ist weiter

$$\begin{aligned} p_{i,i+k-1}(x) &= \dots + y[x_i, \dots, x_{i+k-1}]x^{k-1}, \\ p_{i+1,i+k}(x) &= \dots + y[x_{i+1}, \dots, x_{i+k}]x^{k-1}, \end{aligned}$$

wobei „ \dots “ für Polynomanteile vom Grad kleiner oder gleich $k - 2$ steht. Wie man leicht verifiziert, interpoliert das durch

$$\begin{aligned} q(x) &= \frac{(x - x_i)p_{i+1,i+k}(x) - (x - x_{i+k})p_{i,i+k-1}(x)}{x_{i+k} - x_i} \\ &= p_{i,i+k-1}(x) + (x - x_i) \frac{p_{i+1,i+k}(x) - p_{i,i+k-1}(x)}{x_{i+k} - x_i} \end{aligned}$$

gegebene Polynom $q \in P_k$ die $k + 1$ Stützpunkte (x_j, y_j) , $j = i, \dots, i + k$. Wegen der Eindeutigkeit des Interpolationspolynoms ist dann notwendig $q \equiv p_{i,i+k}$. Der führende Koeffizient in $p_{i,i+k}(x)$ ist demnach

$$a = \frac{y[x_{i+1}, \dots, x_{i+k}] - y[x_i, \dots, x_{i+k-1}]}{x_{i+k} - x_i} = y[x_i, \dots, x_{i+k}],$$

was den Beweis vervollständigt. Q.E.D.

Korollar 2.1: Die Aussage von Satz 2.2 impliziert eine wichtige Invarianzeigenschaft der dividierten Differenzen. Der führende Koeffizient $y[x_0, \dots, x_n]$ des Lagrange-Interpolationspolynoms ist gleichzeitig der Koeffizient des Monoms x^n in seiner Standarddarstellung. Da dieser unabhängig von der Reihenfolge in der Anordnung der Punkte x_0, \dots, x_n ist, gilt dasselbe folglich auch für die dividierte Differenz, d. h.: Es gilt

$$y[\tilde{x}_0, \dots, \tilde{x}_n] = y[x_0, \dots, x_n] \tag{2.1.4}$$

für jede beliebige Permutation $\tilde{x}_0, \dots, \tilde{x}_n$ dieser Punkte.

Die im Beweis von Satz 2.2 verwendete Beziehung zwischen den Polynomen $p_{i,i+k}$ kann direkt zur rekursiven Berechnung des Interpolationspolynoms $p = p_{0,n}$ verwendet werden.

Definition 2.4: Das durch die Rekursion $p_{i,i}(x) = y_i$, $i = 0, \dots, n$, und

$$p_{i,i+k}(x) = p_{i,i+k-1}(x) + (x - x_i) \frac{p_{i+1,i+k}(x) - p_{i,i+k-1}(x)}{x_{i+k} - x_i}, \quad (2.1.5)$$

für $k = 1, \dots, n$, $i = 0, \dots, n-k$, erzeugte Polynom $p_{0,n}$ ist die sog. „Neville⁴-Darstellung“ des Interpolationspolynoms zu den Stützpunkten $(x_0, y_0), \dots, (x_n, y_n)$.

Bei der praktischen Berechnung der Neville-Darstellung geht man nach folgendem Schema vor:

x_0	y_0	$p_{0,1}(x)$	$p_{0,2}(x)$	$p_{0,3}(x)$	\dots	$p_{0,n-1}(x)$	$p_{0,n}(x)$
x_1	y_1	$p_{1,2}(x)$	$p_{1,3}(x)$	$p_{1,4}(x)$	\dots	$p_{1,n}(x)$	
x_2	y_2	$p_{2,3}(x)$	$p_{2,4}(x)$	$p_{2,5}(x)$	\dots		
\vdots	\vdots	\vdots	\vdots	\vdots			
x_{n-1}	y_{n-1}	$p_{n-1,n}(x)$					
x_n	y_n						

Auch hier ist die Hinzunahme eines weiteren Stützpunktes (x_{n+1}, y_{n+1}) problemlos. Die Neville-Darstellung des Interpolationspolynoms bietet eine sehr effiziente und numerisch stabile Möglichkeit zur Berechnung einzelner Funktionswerte $p(\xi)$ ($\xi \neq x_i$) ohne vorherige Bestimmung der Koeffizienten in der Newton-Darstellung. Dazu setzt man im obigen Neville-Schema einfach $x = \xi$ und verwendet zur Berechnung von $p_{i,k} := p_{i,k}(\xi)$ (aus Stabilitätsgründen) die Rekursionsformeln $p_{i,i} = y_i$, $i = 0, \dots, n$, und

$$\begin{aligned} p_{i,i+k} &= p_{i,i+k-1} + (\xi - x_i) \frac{p_{i+1,i+k} - p_{i,i+k-1}}{x_{i+k} - x_i} \\ &= p_{i,i+k-1} + \frac{p_{i,i+k-1} - p_{i+1,i+k}}{(\xi - x_{i+k}) / (\xi - x_i) - 1} \end{aligned} \quad (2.1.6)$$

für $k = 1, \dots, n$, $i = 0, \dots, n-k$.

2.1.1 Auswertung von Polynomen

Ist ein Polynom $p \in P_n$ in der Form $p(x) = a_0 + a_1x + \dots + a_nx^n$ gegeben, so werden einzelne Werte $p(\xi)$ mit Hilfe des „Horner-Schemas“ berechnet (siehe Kapitel 1):

$$b_n = a_n; \quad k = n-1, \dots, 0: \quad b_k \equiv a_k + \xi b_{k+1}; \quad p(\xi) = b_0.$$

Zu $p_n := p \in P_n$ wird durch

$$p_{n-1}(x) := b_1 + b_2x + \dots + b_nx^{n-1}$$

⁴Eric Harold Neville (1889–1961): Englischer Mathematiker; Professor an der Universität in Reading, England (1919–1954); Beiträge zur numerischen Mathematik, u. a. zur praktischen Polynominterpolation.

ein Polynom $p_{n-1} \in P_{n-1}$ definiert. Wegen $a_k = b_k - \xi b_{k+1}$ gilt offenbar

$$p(x) = (x - \xi)p_{n-1}(x) + r_0, \quad r_0 = p(\xi) = b_0,$$

d. h.: Das Horner-Schema leistet unter anderem die Abspaltung des Linearfaktors $x - \xi$ vom Polynom $p(x)$ (euklidischer Algorithmus). Weiter ist dann

$$\frac{p(x) - p(\xi)}{x - \xi} = p_{n-1}(x), \quad x \neq \xi,$$

d. h.: Für $x \rightarrow \xi$ folgt die Beziehung $p'(\xi) = p_{n-1}(\xi)$. Zur Berechnung von $p'(\xi)$ wird das Horner-Schema auf das Polynom p_{n-1} angewendet. Dies liefert Koeffizienten c_k , $k = 2, \dots, n$, sowie ein Polynom $p_{n-2} \in P_{n-2}$ mit der Eigenschaft

$$p_{n-1}(x) = (x - \xi)p_{n-2}(x) + r_1, \quad r_1 = p_{n-1}(\xi) = c_1.$$

Durch fortgesetzte Abspaltung des Linearfaktors $x - \xi$ erhält man so eine (endliche) Folge von Polynomen $p_n, p_{n-1}, p_{n-2}, \dots, p_0$ mit der Eigenschaft

$$p_{n-j}(x) = (x - \xi)p_{n-j-1}(x) + r_j, \quad j = 0, \dots, n-1; \quad p_0 = r_n,$$

und damit die Darstellung

$$p_n(x) = r_0 + r_1(x - \xi) + \dots + r_n(x - \xi)^n. \quad (2.1.7)$$

Vergleicht man dies mit der Taylor-Entwicklung von p an der Stelle ξ , so findet man

$$r_j = \frac{1}{j!} p^{(j)}(\xi), \quad j = 0, \dots, n. \quad (2.1.8)$$

Die Koeffizienten des Polynoms p_{n-j} seien mit $a_k^{(j)}$ bezeichnet:

$$p_{n-j}(x) = a_j^{(j)} + a_{j+1}^{(j)}x + \dots + a_n^{(j)}x^{n-j}, \quad j = 0, \dots, n.$$

Sie werden berechnet durch die rekursive Vorschrift

$$\begin{aligned} j = 0, \dots, n : \quad a_n^{(j+1)} &= a_n^{(j)}, \\ k = n-1, \dots, j : \quad a_k^{(j+1)} &= a_k^{(j)} + \xi a_{k+1}^{(j)}, \end{aligned}$$

und es gilt $p^{(j)}(\xi) = j! a_j^{(j+1)}$, $j = 0, \dots, n$.

Die Koeffizienten $a_k^{(j)}$ bilden das sog. „vollständige Horner-Schema“ zur Auswertung des Polynoms p im Punkt ξ .

Das Horner-Schema kann leicht zur Auswertung eines Polynoms

$$p(x) = a_0 + a_1(x - x_0) + \dots + a_n(x - x_0) \dots (x - x_{n-1})$$

in allgemeiner Newton-Darstellung modifiziert werden:

$$b_n = a_n; \quad k = n - 1, \dots, 0: \quad b_k \equiv a_k + (\xi - x_k)b_{k+1}; \quad p(\xi) = b_0.$$

Ist man an den Ableitung $p^{(j)}(\xi)$ eines in Newton-Form gegebenen Polynoms interessiert, so ist dies weitgehend äquivalent mit dem Problem, zu einer gegebenen Darstellung

$$p(x) = a_0 + a_1(x - x_0) + \dots + a_n(x - x_0)\dots(x - x_{n-1})$$

die Koeffizienten b_k , $k = 0, \dots, n$, in der Darstellung

$$p(x) = b_0 + b_1(x - y_0) + \dots + b_n(x - y_0)\dots(x - y_{n-1})$$

bzgl. anderer (paarweise verschiedener) Punkte y_k , $k = 0, \dots, n - 1$, zu bestimmen. Dies wird durch das folgende „verallgemeinerte Horner-Schema“ geleistet:

$$\begin{aligned} k = 0, \dots, n: & \quad a_k^{(0)}; \\ j = 0, \dots, n - 1: & \quad a_n^{(j+1)} = a_n^{(j)}; \\ k = n - 1, \dots, j: & \quad a_k^{(j+1)} = a_k^{(j)} + (y_j - x_{k-j})a_{k+1}^{(j+1)}, \quad b_j = a_j^{(j+1)}. \end{aligned}$$

2.1.2 Interpolation von Funktionen

Wir betrachten nun den Fall, dass die Knotenwerte y_i durch eine Funktion f auf einem die Stützpunkte x_i enthaltenden Intervall $[a, b]$ gegeben sind:

$$y_i = f(x_i), \quad x_i \in [a, b], \quad i = 0, \dots, n.$$

Zunächst stellt sich die Frage, wie gut das zugehörige Interpolationspolynom $p \in P_n$ die Funktion f auf $[a, b]$ approximiert. Im Folgenden bezeichne (x_0, \dots, x_n) das kleinste Intervall, das alle in den Klammern eingeschlossenen Punkte enthält. Ferner bezeichnet $C[a, b]$ den Vektorraum der über $[a, b]$ *stetigen* Funktionen und analog $C^k[a, b]$ den Vektorraum der über $[a, b]$ *k-mal stetig differenzierbaren* Funktionen.

Satz 2.3 (Interpolationsfehler 1): Sei $f \in C^{n+1}[a, b]$. Dann gibt es zu jedem $x \in [a, b]$ ein $\xi_x \in (x_0, \dots, x_n, x)$, so dass gilt:

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{j=0}^n (x - x_j). \quad (2.1.9)$$

Beweis: Für $x \in \{x_0, \dots, x_n\}$ ist alles klar. Sei also $x \in [a, b] \setminus \{x_0, \dots, x_n\}$. Wir setzen

$$l(t) := \prod_{j=0}^n (t - x_j), \quad c(x) := \frac{f(x) - p(x)}{l(x)}.$$

Die Funktion $F(t) = f(t) - p(t) - c(x)l(t)$ besitzt dann mindestens die $n+2$ Nullstellen x_0, \dots, x_n, x in $[a, b]$. Durch wiederholte Anwendung des Satzes von Rolle erschließt man, dass dann die Ableitung $F^{(n+1)}$ eine Nullstelle $\xi_x \in (x_0, \dots, x_n, x)$ hat. Mit

$$0 = F^{(n+1)}(\xi_x) = f^{(n+1)}(\xi_x) - p^{(n+1)}(\xi_x) - c(x)l^{(n+1)}(\xi_x) = f^{(n+1)}(\xi_x) - c(x)(n+1)!$$

folgt die Behauptung.

Q.E.D.

Satz 2.4 (Interpolationsfehler 2): Sei $f \in C^{n+1}[a, b]$. Dann gestattet der Fehler bei der Polynominterpolation für $x \in [a, b] \setminus \{x_0, \dots, x_n\}$ die Darstellung

$$f(x) - p(x) = f[x_0, \dots, x_n, x] \prod_{j=0}^n (x - x_j), \quad (2.1.10)$$

mit der Notation $f[x_i, \dots, x_{i+k}] := y[x_i, \dots, x_{i+k}]$, und es ist

$$f[x_0, \dots, x_n, x] = \int_0^1 \int_0^{t_1} \dots \int_0^{t_{n-1}} \int_0^{t_n} f^{(n+1)}(x_0 + t_1(x_1 - x_0) + \dots + t_n(x_n - x_{n-1}) + t(x - x_n)) dt dt_n \dots dt_2 dt_1.$$

Beweis: Der Beweis wird durch Induktion nach der Anzahl der Stützstellen (in der Reihung x_0, x_1, x_2, \dots) geführt. Für $n=0$ gilt trivialerweise

$$f(x) - p_0(x) = f(x) - f(x_0) = \begin{cases} f[x_0, x](x - x_0), \\ (x - x_0) \int_0^1 f'(x_0 + t(x - x_0)) dt. \end{cases}$$

Sei die Behauptung nun richtig für $n-1 \geq 0$. Dann ist

$$\begin{aligned} f(x) - p_n(x) &= f(x) - \sum_{i=0}^n f[x_0, \dots, x_i] \prod_{j=0}^{i-1} (x - x_j) \\ &= f(x) - p_{n-1}(x) - f[x_0, \dots, x_n] \prod_{j=0}^{n-1} (x - x_j) \\ &= f[x_0, \dots, x_{n-1}, x] \prod_{j=0}^{n-1} (x - x_j) - f[x_0, \dots, x_n] \prod_{j=0}^{n-1} (x - x_j) \\ &= \frac{f[x_0, \dots, x_{n-1}, x] - f[x_0, \dots, x_n]}{x - x_n} \prod_{j=0}^n (x - x_j), \end{aligned}$$

und somit, wegen $f[x_0, \dots, x_{n-1}, x] = f[x, x_0, \dots, x_{n-1}]$, unter Ausnutzung der Definition der dividierten Differenzen:

$$f(x) - p_n(x) = f[x_0, \dots, x_n, x] \prod_{j=0}^n (x - x_j).$$

Ferner ist nach Induktionsvoraussetzung

$$\begin{aligned}
 & f[x_0, \dots, x_{n-1}, x] - f[x_0, \dots, x_n] \\
 &= \int_0^1 \int_0^{t_1} \dots \int_0^{t_{n-1}} \{ f^{(n)}(x_0 + t_1(x_1 - x_0) + \dots + t_n(x - x_{n-1})) \\
 &\quad - f^{(n)}(x_0 + t_1(x_1 - x_0) + \dots + t_n(x_n - x_{n-1})) \} dt_n \dots dt_1 \\
 &= \int_0^1 \int_0^{t_1} \dots \int_0^{t_n} \underbrace{\frac{d}{dt} f^{(n)}(x_0 + \dots + t_n(x_n - x_{n-1}) + t(x - x_n))}_{f^{(n+1)}(\dots)(x - x_n)} dt dt_n \dots dt_1
 \end{aligned}$$

und folglich, nach Definition der dividierten Differenzen,

$$f[x_0, \dots, x_n, x] = \int_0^1 \int_0^{t_1} \dots \int_0^{t_{n-1}} \int_0^{t_n} f^{(n+1)}(\dots) dt dt_n \dots dt_1.$$

Dies vervollständigt den Beweis. Q.E.D.

Die obige Integraldarstellung der dividierten Differenzen $f[x_0, \dots, x_n]$ gestattet ihre stetige Fortsetzung für den Fall, daß einige der Stützstellen zusammenfallen:

$$f[x_0, \dots, x_r, x_r, \dots, x_n] := \lim_{\varepsilon \rightarrow 0} f[x_0, \dots, x_r, x_r + \varepsilon, \dots, x_n].$$

Im Extremfall $x_0 = \dots = x_n$ wird

$$f[x_0, \dots, x_n] = \int_0^1 \int_0^{t_1} \dots \int_0^{t_{n-1}} f^{(n)}(x_0) dt_n \dots dt_2 dt_1 = \frac{1}{n!} f^{(n)}(x_0),$$

und das Newton-Interpolationspolynom geht über in das Taylor-Polynom n -ten Grades von f in x_0 :

$$p_n(x) = \sum_{i=0}^n f[x_0, \dots, x_i] \prod_{j=0}^{i-1} (x - x_j) = \sum_{i=0}^n \frac{1}{i!} f^{(i)}(x_0) (x - x_0)^i.$$

Ferner sehen wir, dass sich die dividierten Differenzen einer hinreichend oft differenzierbaren Funktion durch Zwischenwerte der entsprechenden Ableitungen ausdrücken lassen:

$$f[x_0, \dots, x_n, x] = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x), \quad (2.1.11)$$

mit einer Zwischenstelle $\xi_x \in \overline{(x_0, \dots, x_n, x)}$.

Wir wollen nun allgemein den Fehler bei der Lagrange-Interpolation diskutieren:

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{j=0}^n (x - x_j). \quad (2.1.12)$$

Für großes n wird $1/n!$ sehr klein, und das Produkt wird klein, wenn die Stützstellen immer dichter zusammenrücken. Sind die Ableitungen von f beschränkt auf $[a, b]$, so gilt also sicher

$$\max_{a \leq x \leq b} |f(x) - p(x)| \rightarrow 0 \quad (n \rightarrow \infty).$$

Meist haben die Ableitungen der zu interpolierenden Funktionen jedoch ein zu starkes Wachstum für $n \rightarrow \infty$, z. B.:

$$f(x) = (1 + x^2)^{-1}, \quad |f^{(n)}(x)| \approx 2^n n! O(|x|^{-2-n}),$$

so dass gleichmäßige Konvergenz des Approximationsprozesses nicht mehr zu erwarten ist.

Beispiel 2.1: Für die Funktion $f(x) = |x|$, $x \in [-1, 1]$, ergibt die Lagrange-Interpolation in den Stützstellen $x_i = -1 + ih$, $i = 0, \dots, 2m$, mit $h = 1/m$ und $x \notin \{x_i, i = 0, \dots, 2m\}$ das Verhalten

$$p_m(x) \not\rightarrow f(x) \quad (m \rightarrow \infty).$$

Dieser Effekt ist nicht auf nicht differenzierbare Funktionen beschränkt, wie das obige Beispiel $f(x) = (1 + x^2)^{-1}$, $x \in [-5, 5]$ zeigt (s. Übungsaufgabe).

Bemerkung 2.1: Der Approximationssatz von Weierstrass⁵ besagt, dass jede Funktion $f \in C[a, b]$ beliebig gut gleichmäßig auf $[a, b]$ durch Polynome approximiert werden kann. Die Vermutung, dass dies mit Lagrange-Interpolationspolynomen geschehen kann, ist jedoch i. Allg. falsch.

Ein weiterer Defekt der Lagrange-Interpolation besteht in ihrer großen Fehlerempfindlichkeit. Fehlerhafte Daten $y_i + \Delta y_i$ wirken sich auf die Gestalt des Polynoms nicht nur lokal bei der Stützstelle x_i aus, sondern verändern den Verlauf auch relativ dramatisch über dem ganzen Intervall.

Beispiel 2.2: $x_i = -1 + ih$, $i = 0, \dots, 2m$, $h = 1/m$; $y_i = 0$ für $i \neq m$, $y_m = \varepsilon$;

$$p(x) = \varepsilon \prod_{\substack{j=0 \\ j \neq m}}^{2m} \frac{x - x_j}{x_j} \quad (\text{Lagrange-Polynom zum Aufpunkt } x_m = 0)$$

Dies liegt daran, dass auch für „gutartige“ Funktionen der Term $|\prod_{j=0}^n (x - x_j)|$ für $x \notin \overline{(x_0, \dots, x_n)}$ sehr schnell anwächst. Die Verwendung der Polynominterpolation zur weitreichenden „Extrapolation“ ist also nur bedingt zu empfehlen. Hierfür erweist sich die Interpolation mit *rationalen* Funktionen als wesentlich geeigneter.

⁵Karl Theodor Weierstrass (1815–1897): Deutscher Mathematiker; ab 1856 Professor in Berlin; begründete die moderne, strenge Analysis.

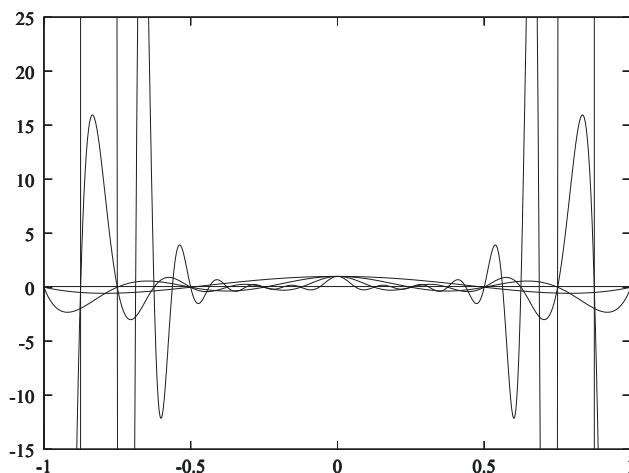


Abbildung 2.1: Störungsverlauf bei wachsendem Polynomgrad

2.1.3 Hermite-Interpolation

Die Aufgabenstellung der Lagrange-Interpolation lässt sich verallgemeinern zur sog. „Hermite⁶-Interpolation“:

Definition 2.5: Die Hermite-Interpolationsaufgabe lautet wie folgt:

Gegeben: $x_i, i = 0, \dots, m$ (paarweise verschieden)

$y_i^{(k)}, i = 0, \dots, m, k = 0, \dots, \mu_i$ ($\mu_i \geq 0$).

Gesucht: $p \in P_n, n = m + \sum_{i=0}^m \mu_i: p^{(k)}(x_i) = y_i^{(k)}$.

Die Punkte x_i werden als $(\mu_i + 1)$ -fache Stützstellen bezeichnet.

Analog zu Satz 2.1 beweist man:

Satz 2.5 (Hermite-Interpolation): Die Hermite-Interpolationsaufgabe besitzt eine eindeutige Lösung.

Sind die Knotenwerte $y_i^{(k)} = f^{(k)}(x_i)$ durch eine Funktion f gegeben, so gilt:

⁶Charles Hermite (1822–1901): Französischer Mathematiker; Professor an der École Polytechnique und der Sorbonne in Paris; Beiträge zur Zahlentheorie und zur Theorie elliptischer Funktionen.

Satz 2.6 (Interpolationsfehler): Ist $f \in C^{n+1}[a, b]$, so gibt es zu jedem $x \in [a, b]$ ein $\xi_x \in (x_0, \dots, x_m, x)$, so dass für die Lösung $p \in P_n$ der Hermite-Interpolationsaufgabe gilt:

$$\begin{aligned} f(x) - p(x) &= f[x_0, \dots, x_0, \dots, x_m, \dots, x_m, x] \prod_{i=0}^m (x - x_i)^{\mu_i+1} \\ &= \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{i=0}^m (x - x_i)^{\mu_i+1}. \end{aligned} \quad (2.1.13)$$

Beweis: Analog zu Satz 2.3 bzw. Satz 2.4.

Q.E.D.

Aus Stetigkeitsgründen besitzt das Hermite-Interpolationspolynom einer Funktion $f \in C^{n+1}[a, b]$ die Darstellung

$$\begin{aligned} p(x) &= \sum_{i=0}^m \sum_{r=1}^{\mu_i+1} f[\underbrace{x_0, \dots, x_0}_{(\mu_0+1)\text{-mal}}, \dots, \underbrace{x_{i-1}, \dots, x_{i-1}}_{(\mu_{i-1}+1)\text{-mal}}, \underbrace{x_i, \dots, x_i}_{r\text{-mal}}] \times \\ &\quad \times \prod_{j=0}^{i-1} (x - x_j)^{\mu_j+1} (x - x_i)^{r-1}. \end{aligned}$$

Die dividierten Differenzen $f[\dots]$ sind durch ein zur Lagrange-Interpolation analoges Rekursionsschema bestimmt, wobei immer dann, wenn Differenzen wegen des Zusammenfallens von Stützpunkten nicht nach der Definitionsformel gebildet werden können, sinngemäß die Stützwerte *höherer Ordnung* einzusetzen sind; z. B.:

$$\begin{aligned} y[x_i, x_i] &= y_i^{(1)}, & y[x_i, x_i, x_{i+1}] &= \frac{y[x_i, x_{i+1}] - y_i^{(1)}}{x_{i+1} - x_i}, \\ y[x_i, x_i, x_i] &= \frac{1}{2} y_i^{(2)}, & y[x_i, x_i, x_i, x_{i+1}] &= \frac{y[x_i, x_i, x_{i+1}] - \frac{1}{2} y_i^{(2)}}{x_{i+1} - x_i}. \end{aligned}$$

Eine weitere Verallgemeinerung der Lagrange- und Hermite-Interpolation ist die sog. „Hermite/Birkhoff⁷-Interpolation“, bei der Funktions- bzw. Ableitungswerte in verschiedenen Punkten beliebig gemischt vorgegeben werden, z. B.:

$$p \in P_3 : \quad p(0) = 1, \quad p'(1) = 2, \quad p(2) = 0, \quad p'(3) = 3.$$

Die Frage nach der Lösbarkeit der allgemeinen Hermite/Birkhoff-Interpolationsaufgabe ist noch nicht vollständig geklärt. Es können im Gegensatz zur Lagrange- und Hermite-Aufgabe alle Fälle von „eindeutig lösbar“ über „unendlich mehrdeutig lösbar“ bis „unlösbar“ auftreten (Übungsaufgabe).

⁷George David Birkhoff (1884–1944): US-Amerikanischer Mathematiker; Professor an der Harvard University, Boston, USA; Beiträge zu sehr verschiedenen Gebieten der Mathematik: dynamische Systeme (Ergoden-Satz), Differentialgleichungen, Gravitationstheorie und Mathematik der Ästhetik.

2.2 Extrapolation zum Limes

Eine wichtige Anwendung der Polynominterpolation ist die sog. „Richardson⁸-Extrapolation (zum Limes)“. Ein numerischer Prozess liefere für jeden Wert eines positiven Parameters $h \in \mathbb{R}_+$ ($h \rightarrow 0$) einen Wert $a(h)$. Gesucht ist die nicht direkt berechenbare Größe

$$a(0) = \lim_{h \rightarrow 0} a(h). \quad (2.2.14)$$

Zur Annäherung von $a(0)$ berechnet man $a(h_i)$ für gewisse Werte h_i , $i = 0, \dots, n$, und nimmt den Wert $p_n(0)$ des zugehörigen Interpolationspolynoms zu $(h_i, a(h_i))$ als Schätzung für $a(0)$.

Beispiel 2.3: Wir betrachten die numerische Realisierung der Regel von l’Hopital⁹. Zur Berechnung von

$$a(0) := \lim_{x \rightarrow +0} \frac{\cos(x) - 1}{\sin(x)} \quad (= 0)$$

setzen wir

$$a(x) := \frac{(\cos(x) - 1)}{\sin(x)}$$

und interpolieren $a(x)$ an einigen Stützstellen h_i nahe bei 0:

$$\begin{aligned} h_0 &= \frac{1}{8}, & a(h_0) &= -6.258151 \cdot 10^{-2}, \\ h_1 &= \frac{1}{16}, & a(h_1) &= -3.126018 \cdot 10^{-2}, \\ h_2 &= \frac{1}{32}, & a(h_2) &= -1.562627 \cdot 10^{-2}. \end{aligned}$$

Das interpolierende Polynom ist in Lagrange-Darstellung:

$$p_2(x) = a(h_0) \frac{(x - \frac{1}{16})(x - \frac{1}{32})}{(\frac{1}{8} - \frac{1}{16})(\frac{1}{8} - \frac{1}{32})} + a(h_1) \frac{(x - \frac{1}{8})(x - \frac{1}{32})}{(\frac{1}{16} - \frac{1}{8})(\frac{1}{16} - \frac{1}{32})} + a(h_2) \frac{(x - \frac{1}{8})(x - \frac{1}{16})}{(\frac{1}{32} - \frac{1}{8})(\frac{1}{32} - \frac{1}{16})},$$

und wir erhalten

$$a(0) \sim p_2(0) = -1.02 \cdot 10^{-5}.$$

Numerisch günstiger wäre die Berechnung von $p_2(0)$ mit Hilfe des Neville-Algorithmus nach dem Schema

$$p_{i,i+k}(0) = p_{i,i+k-1}(0) + \frac{p_{i,i+k-1}(0) - p_{i+1,i+k}(0)}{x_{i+k}/x_i - 1}, \quad k = 1, 2.$$

⁸Lewis Fry Richardson (1881–1953): Englischer Mathematiker und Physiker; wirkte an verschiedenen Institutionen in England und Schottland; typischer „angewandter Mathematiker“; leistete Pionierbeiträge zur Modellierung und Numerik in der Wettervorhersage.

⁹Guillaume F. A. Marquis de l’Hopital (1661–1704): Französischer Mathematiker; Schüler von Johann Bernoulli; veröffentlichte 1696 das erste Lehrbuch der Differentialrechnung, welches auch die nach ihm benannte Regel enthält.

i	x_i	$p_{i,i}(0) = a(h_i)$	$p_{i,i+1}(0)$	$p_{i,i+2}(0)$
0	$x_0 = \frac{1}{8}$	$-6.258151 \cdot 10^{-2}$	$6.115 \cdot 10^{-5}$	$-1.02 \cdot 10^{-5}$
1	$x_1 = \frac{1}{16}$	$-3.126018 \cdot 10^{-2}$	$7.64 \cdot 10^{-6}$	
2	$x_2 = \frac{1}{32}$	$-1.562627 \cdot 10^{-2}$		

Beispiel 2.4: Wir betrachten die numerische Differentiation. Für C^1 -Funktionen f ist

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = f'(x),$$

und für $f \in C^2$ folgt durch Taylor-Entwicklung

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \frac{h}{2} f''(\zeta_x), \quad \zeta_x \in \overline{(x, x+h)}.$$

Im Falle $f \in C^3$ erhält man eine bessere Näherung zu $f'(x)$ durch den zentralen Differenzenquotienten

$$\frac{f(x+h) - f(x-h)}{2h} = f'(x) + \frac{h^2}{6} f'''(\zeta_x), \quad \zeta_x \in \overline{(x+h, x-h)}.$$

Ist f analytisch, so gilt sogar

$$a(h) := \frac{f(x+h) - f(x-h)}{2h} = f'(x) + \sum_{i=1}^{\infty} \frac{f^{(2i+1)}(x)}{(2i)!} h^{2i},$$

d. h.: $a(h)$ ist eine „gerade“ Funktion in h . Zur Extrapolation von $a(h)$ verwendet man daher zweckmäßigerweise auch gerade Polynome, d. h. Polynome in h^2 . Als Beispiel betrachten wir $f(x) = \sin(x)$ mit $f'(0) = \cos(0) = 1$ und

$$a(h) \equiv \frac{\sin(h) - \sin(-h)}{2h} = \frac{\sin(h)}{h}.$$

Auswertung von $a(h)$ in

$$h_0 = \frac{1}{8}, \quad h_1 = \frac{1}{16}, \quad h_2 = \frac{1}{32} \quad (\text{„heimliche“ Knoten: } -\frac{1}{8}, -\frac{1}{16}, -\frac{1}{32})$$

$$a(h_0) = 0.9973979, \quad a(h_1) = 0.9993491, \quad a(h_2) = 0.9998372,$$

ergibt dann

$$p_2(h) = a(h_0) \frac{(h^2 - \frac{1}{16^2})(h^2 - \frac{1}{32^2})}{(\frac{1}{8^2} - \frac{1}{16^2})(\frac{1}{8^2} - \frac{1}{32^2})} + \dots, \quad p_2(0) = 0.999999926.$$

Der folgende Satz liefert die theoretische Grundlage der Richardson-Extrapolation.

Satz 2.7 (Extrapolationsfehler): Für die Funktion $a(h)$, $h \in \mathbb{R}_+$, sei bekannt, dass eine asymptotische Entwicklung der Form

$$a(h) = a_0 + \sum_{j=1}^n a_j h^{jq} + a_{n+1}(h) h^{(n+1)q} \quad (2.2.15)$$

gilt, mit einem $q > 0$, und gewissen Koeffizienten a_j und $a_{n+1}(h) = a_{n+1} + o(1)$ für $h \rightarrow 0$. Sei $(h_k)_{k=0,1,2,\dots}$ eine monoton fallende Folge positiver Zahlen mit der Eigenschaft

$$0 < \frac{h_{k+1}}{h_k} \leq \rho < 1. \quad (2.2.16)$$

Für das Interpolationspolynom $p_n^{(k)} \in P_n$ (in h^q) durch $(h_k^q, a(h_k)), \dots, (h_{k+n}^q, a(h_{k+n}))$ gilt dann:

$$a(0) - p_n^{(k)}(0) = O(h_k^{(n+1)q}) \quad (k \rightarrow \infty). \quad (2.2.17)$$

Beweis: Wir setzen zur Abkürzung $z = h^q$ und $z_k = h_k^q$. Das Interpolationspolynom zu den Stützpunkten $(z_{k+i}; a(h_{k+i}))$, $i = 0, \dots, n$, ist in Lagrange-Darstellung:

$$p_n(z) = \sum_{i=0}^n a(h_{k+i}) L_{k+i}^{(n)}(z), \quad L_{k+i}^{(n)}(z) = \prod_{\substack{l=0 \\ l \neq i}}^n \frac{z - z_{k+l}}{z_{k+i} - z_{k+l}}.$$

Aus der Fehlerdarstellung

$$f(x) - p_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\zeta_x) \prod_{i=0}^n (x - x_{k+i}), \quad \zeta_x \in [0, h_0],$$

für $f \equiv 1$ sowie $f(x) = x^r$ liest man ab, dass (Übungsaufgabe)

$$\sum_{i=0}^n z_{k+i}^r L_{k+i}^{(n)}(0) = \begin{cases} 1 & , \text{ für } r = 0, \\ 0 & , \text{ für } r = 1, \dots, n, \\ (-1)^n \prod_{i=0}^n z_{k+i} & , \text{ für } r = n+1. \end{cases}$$

Damit erschließen wir:

$$\begin{aligned} p_n(0) &= \sum_{i=0}^n \left\{ a_0 + \sum_{j=1}^n a_j z_{k+i}^j + a_{n+1}(h_{k+i}) z_{k+i}^{n+1} \right\} L_{k+i}^{(n)}(0) \\ &= a_0 \sum_{i=0}^n L_{k+i}^{(n)}(0) + \sum_{j=1}^n a_j \left\{ \sum_{i=0}^n z_{k+i}^j L_{k+i}^{(n)}(0) \right\} + \\ &\quad a_{n+1} \sum_{i=0}^n z_{k+i}^{n+1} L_{k+i}^{(n)}(0) + \sum_{i=0}^n o(1) z_{k+i}^{n+1} L_{k+i}^{(n)}(0), \end{aligned}$$

und somit

$$p_n(0) = a_0 + a_{n+1}(-1)^n \prod_{i=0}^n h_{k+i}^q + o(h_k^{(n+1)q}).$$

Hier wurde (2.2.16) benutzt, um die von h_k unabhängige Abschätzung

$$\left| L_{k+i}^{(n)}(0) \right| = \prod_{l=0, l \neq i}^n \left| \frac{z_{k+l}}{z_{k+i} - z_{k+l}} \right| = \prod_{l=0, l \neq i}^n \left| \frac{1}{z_{k+i}/z_{k+l} - 1} \right| \leq \gamma(n, \rho) \quad (2.2.18)$$

zu garantieren. Wegen

$$\prod_{i=0}^n h_{k+i}^q = O(h_k^{(n+1)q})$$

ergibt sich (2.2.17) gleichmäßig für alle k .

Q.E.D.

Üblicherweise wird ein Extrapolationsprozess anhand des folgenden „Extrapolationstableau“ zur Berechnung der Werte $p_{i,i+k} = p_{i,i+k}(0)$ durchgeführt. Es sei daran erinnert, dass $p_{i,i+k}(h)$ das Polynom (in h^q) ist, welches die Punkte $(h_i^q, a(h_i)), \dots, (h_{i+k}^q, a(h_{i+k}))$ interpoliert. Die Funktionswerte $p_{i,i+k}$ erhält man nach dem Neville-Algorithmus:

$$p_{i,i+k} = p_{i,i+k-1} + \frac{p_{i,i+k-1} - p_{i+1,i+k}}{x_{i+k}/x_i - 1}.$$

Der Konvention folgend setzen wir $a_{ik} \equiv p_{i-k,i}$:

	$a_{i0} = a(h_i)$				
h_0	a_{00}	Extrapolationstableau			
h_1	a_{10}	\rightarrow	a_{11}		
h_2	a_{20}	\rightarrow	a_{21}	\rightarrow	a_{22}
\vdots	\vdots	\vdots	\vdots	\ddots	
h_i	a_{i0}	a_{i1}	a_{i2}	\cdots	a_{ii}

Die Tableaueinträge werden sukzessive nach der folgenden Rekursionsformel berechnet:

$$\begin{aligned} i = 0, 1, 2, \dots: & \quad a_{i0} = a(h_i), \\ i = 1, 2, 3, \dots, k = 1, 2, 3, \dots, i: & \quad a_{ik} = a_{i,k-1} + \frac{a_{i,k-1} - a_{i-1,k-1}}{(h_{i-k}/h_i)^q - 1}. \end{aligned} \quad (2.2.19)$$

Nach Satz 2.7 gilt dann für festes k :

$$a(0) - a_{ik} = O(h_{i-k}^{(k+1)q}) \quad (i \rightarrow \infty), \quad (2.2.20)$$

vorausgesetzt die Schrittweitenfolge genügt der Bedingung (2.2.16). Gebräuchliche Folgen sind gegeben durch $h_i \equiv h_0/n_i$ mit

$$i) \quad n_i = 2^i, \quad ii) \quad n_i = 2, 4, 6, 8, 12, 16, \dots, \quad iii) \quad n_i = 1, 2, 3, \dots \text{ (unzulässig).}$$

2.2.1 Fehlerkontrolle

Zur praktischen Durchführung der Extrapolation (zum Limes $h=0$) gehört ein Kriterium, wann der Extrapolationsprozess abubrechen ist. Sei eine zu erzielende Fehlertoleranz TOL vorgegeben. Der Fehlerdarstellung (siehe den Beweis von Satz 2.7)

$$a_{ik} = a(0) + a_{k+1}(-1)^k \prod_{j=0}^k h_{i-k+j}^q + o(h_{i-k}^{(k+1)q})$$

entnehmen wir, dass für festes k und genügend großes i der Fehler $a_{ik} - a(0)$ monoton gegen Null konvergiert (falls $a_{k+1} \neq 0$). Mit den neu definierten Größen

$$b_{ik} \equiv 2a_{i+1,k} - a_{ik}$$

gilt dann im Falle $q \geq 1$:

$$\begin{aligned} b_{ik} - a(0) &= 2\{a_{i+1,k} - a(0)\} - \{a_{ik} - a(0)\} \\ &= 2a_{k+1}(-1)^k \prod_{j=0}^k h_{i+1-k+j}^q + o(h_{i+1-k}^{(k+1)q}) - a_{k+1}(-1)^k \prod_{j=0}^k h_{i-k+j}^q + o(h_{i-k}^{(k+1)q}) \\ &= \prod_{j=0}^k h_{i-k+j}^q \left\{ -a_{k+1}(-1)^k + 2 \left(\frac{h_{i+1}}{h_{i-k}} \right)^q a_{k+1}(-1)^k \right\} + o(h_{i-k}^{(k+1)q}). \end{aligned}$$

Wegen $h_{i+1}^q/h_{i-k}^q \ll 1$ gilt also in erster Näherung

$$b_{ik} - a(0) \doteq -(-1)^k a_{k+1} \prod_{j=0}^k h_{i-k+j}^q,$$

und folglich

$$a_{ik} - a(0) \doteq -(b_{ik} - a(0)), \quad (2.2.21)$$

für festes k und genügend großes i . Wegen der monotonen Konvergenz $a_{ik} - a(0) \rightarrow 0$ ($i \rightarrow \infty$) gilt also asymptotisch entweder $a_{ik} \leq a(0) \leq b_{ik}$ oder $a_{ik} \geq a(0) \geq b_{ik}$, und beide Seiten konvergieren monoton gegen $a(0)$ für $i \rightarrow \infty$. Dieses Verhalten der Folgen $(a_{ik})_{i \in \mathbb{N}}$ und $(b_{ik})_{i \in \mathbb{N}}$ (für festes k) kann zur Konstruktion eines Abbruchkriteriums herangezogen werden:

$$|a_{ik} - b_{ik}| < \text{TOL} \quad \Rightarrow \quad \text{STOP}. \quad (2.2.22)$$

Bemerkung 2.2: Für die Praxis lohnt es sich, den Extrapolationsprozess vollständig durchzuführen, d. h. die Diagonalelemente a_{ii} des Tableaus zu berechnen. In der Tat kann man zeigen, dass die „Diagonalfolge“ $(a_{ii})_{i \in \mathbb{N}}$ schneller gegen $a(0)$ konvergiert als jede der Spaltenfolgen $(a_{ik})_{i \in \mathbb{N}}$, $k \geq 0$; es gilt $a_{ii} - a(0) = o(h_{i-k}^{(k+1)q})$.

2.3 Spline-Interpolation

Die Lagrange-Interpolationspolynome eignen sich nicht besonders gut zur Approximation von (nicht glatten) Funktionen, da sie bei Vermehrung der Stützstellenzahl dazu neigen, zwischen den Stützstellen immer größere Werte anzunehmen. Dies ist eine Folge ihrer „Steifheit“ bedingt durch die Forderung von C^∞ -Übergängen in den Knoten. Zur Reduzierung dieser Steifheit setzt man die interpolierende Funktion φ nur als „stückweise polynomial“ bzgl. der Zerlegung $a = x_0 < x_1 < \dots < x_n = b$ an. In den Knoten x_i werden dann geeignete Differenzierbarkeitseigenschaften (z. B. $\varphi \in C^2[a, b]$) gefordert.

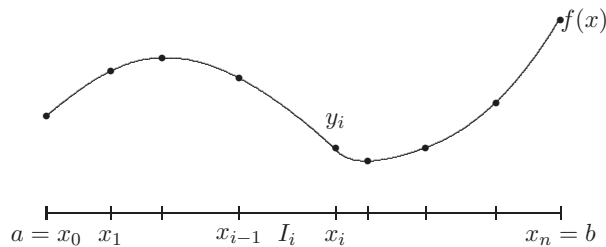


Abbildung 2.2: *Spline-Approximation*

Die Länge des Teilintervalls $I_i = [x_{i-1}, x_i]$ ist $h_i = x_i - x_{i-1}$, und die Feinheit der ganzen Intervallunterteilung wird durch die Größe $h := \max_{i=1, \dots, n} h_i$ beschrieben. Auf einer solchen Intervallzerlegung werden Vektorräume von stückweise polynomialen Funktionen betrachtet

$$S_h^{(k,r)}[a, b] = \{p \in C^r[a, b], p|_{I_i} \in P_k(I_i)\}$$

für $k, r \in \{0, 1, 2, \dots\}$. Zu einem Satz von Stützwerten in Punkten aus dem Intervall $[a, b]$, die etwa wieder von einer zu interpolierenden Funktion f genommen werden, wird dann eine „Interpolierende“ $p \in S_h^{(k,r)}[a, b]$ mit Hilfe von geeigneten Interpolationsbedingungen bestimmt. Wir betrachten im Folgenden einige einfache Beispiele.

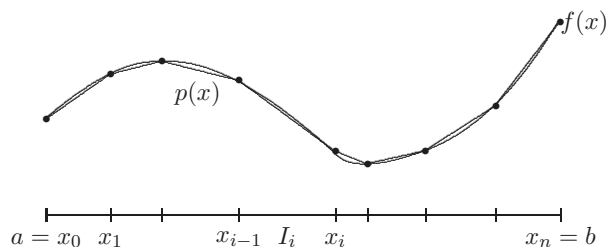


Abbildung 2.3: *Stückweise lineare Interpolation*

Beispiel 2.5: Die stückweise lineare Lagrange-Interpolation (Fall $k=1, r=0$) approximiert eine gegebene Funktion f auf $[a, b]$ durch einen Polygonzug in den Stützstellen $x_i, i=0, \dots, n$:

$$p \in S_h^{(1,0)}[a, b] = \{p \in C[a, b], p|_{I_i} \text{ linear}\}, \quad p(x_i) = f(x_i), \quad i = 0, \dots, n.$$

Anwendung der Fehlerabschätzung für die Lagrange-Interpolation separat auf jedem der Teilintervalle I_i ergibt die globale Fehlerabschätzung

$$\max_{x \in [a, b]} |f(x) - p(x)| \leq \frac{1}{2} h^2 \max_{x \in [a, b]} |f''(x)|.$$

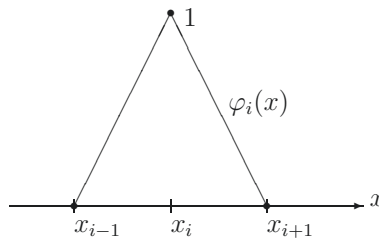


Abbildung 2.4: „Lineare“ Knotenbasisfunktion

Für die Konstruktion der Interpolierenden verwendet man die sog. „Knotenbasis“ von $S_h^{(1,0)}[a, b]$ bestehend aus den „Dachfunktionen“ $\varphi_i \in S_h^{(1,0)}[a, b], i = 0, \dots, n$, die durch die Bedingung

$$\varphi_i(x_j) = \delta_{ij}$$

eindeutig bestimmt sind. Dass dies tatsächlich eine Basis ergibt, ist durch elementare Argumente einzusehen und sei als Übungsaufgabe gestellt. Die Interpolierende p von f lässt sich dann in der Form

$$p(x) = \sum_{i=0}^n f(x_i) \varphi_i(x).$$

darstellen. Diese Konstruktion ist analog zur Lagrange-Darstellung des Lagrange-Interpolationspolynoms. Da die Dachfunktionen φ_i nur in den direkt an den jeweiligen Aufpunkt x_i angrenzenden Teilintervallen von Null verschieden sind, nennt man diese Basis „lokal“. Höhere globale Glattheit als Stetigkeit ist sinnvoll mit stückweise linearen Interpolierenden offensichtlich nicht erzielbar. Dazu benötigt man Polynomansätze höherem Grades.

Beispiel 2.6: Wir betrachten noch die stückweise Interpolation mit kubischen Polynomen ($k=3, r=0$ oder $r=1$): $S_h^{(3,0)}[a, b]$ und $S_h^{(3,1)}[a, b]$. Zur Erzielung globaler Stetigkeit

($r=0$) verwendet man als Interpolationsbedingungen

$$p(x_i) = f(x_i), \quad p(x_{ij}) = f(x_{ij}),$$

mit jeweils zwei zusätzlichen Interpolationspunkten $x_{ij} \in I_i$, $i=1, \dots, n$, $j=1, 2$. Durch stückweise kubische Lagrange-Interpolation ist dadurch eindeutig eine global stetige Interpolierende $p \in S_h^{(3,0)}[a, b]$ festgelegt. Analog erhält man durch stückweise kubische Hermite-Interpolation aus den Interpolationsbedingungen

$$p(x_i) = f(x_i), \quad p'(x_i) = f'(x_i), \quad i = 0, \dots, n,$$

eine global stetig differenzierbare Interpolierende $p \in S_h^{(3,1)}[a, b]$. In beiden Fällen folgt aus den jeweiligen Fehlerabschätzungen die globale Fehlerabschätzung

$$\max_{x \in [a, b]} |f(x) - p(x)| \leq \frac{1}{4!} h^4 \max_{x \in [a, b]} |f^{(4)}(x)|.$$

Analog zum Fall $k = 1$ lassen sich auch hier wieder „lokale“ Knotenbasen der Ansatzräume $S_h^{(3,0)}$ und $S_h^{(3,1)}$ angeben. Derartige rein lokale Interpolationsprozesse haben Anwendungen etwa bei der numerischen Lösung von gewöhnlichen und partiellen Differentialgleichungen.

Die Forderung nach höherer globaler Glattheit, etwa Interpolation in $S_h^{(3,2)}$, führt auf die sog. „Spline-Interpolation“, welche von großer praktischer Bedeutung z. B. bei der glatten Darstellung von Flächen in der Computergrafik ist. Der Begriff „Spline“ stammt aus dem Englischen und bedeutet soviel wie „biegsames Kurvenlineal“ („Biegestab“):

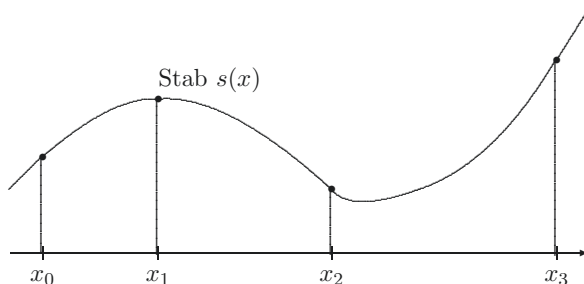


Abbildung 2.5: (Kubische) Spline-Funktion

Nach einem Grundgesetz der Mechanik (Prinzip vom Minimum der potentiellen Gesamtenergie) stellt sich der Biegestab so ein, dass die Gesamtkrümmung minimiert wird, d. h. in linearisierter Näherung:

$$\int_{x_0}^{x_n} s''(x)^2 dx = \min! \quad (2.3.23)$$

bzgl. aller möglichen interpolierenden (hinreichend glatten) Funktionen. Außerhalb des Interpolationsintervalls $[x_0, x_n]$ kann $s(x)$ als linear angenommen werden, d. h.

$$s''(x_0) = s''(x_n) = 0. \quad (2.3.24)$$

Diese sog. „natürlichen“ Randbedingungen stellen sich also automatisch ein, wenn der Stab nicht willkürlich zu einem anderen Verhalten gezwungen wird („erzwungene“ Randbedingungen: z. B. $s'(x_0) = s'(x_n) = 0$). Wir betrachten hier der Einfachheit halber nur die am häufigsten verwendeten „kubischen“ Spline-Funktionen; sie ergeben sich auf natürliche Weise aus dem obigen Biegestabmodell aufgrund der Forderung (2.3.23).

Definition 2.6: Eine Funktion $s_n : [a, b] \rightarrow \mathbb{R}$ heißt „kubischer Spline“ bzgl. einer Zerlegung $a = x_0 < x_1 < \dots < x_n = b$, wenn gilt:

- (i) $s_n \in C^2[a, b]$;
- (ii) $s_n|_{[x_{i-1}, x_i]} \in P_3, \quad i = 1, \dots, n$.

Gilt zusätzlich

$$(iii) \quad s_n''(a) = s_n''(b) = 0,$$

so wird der kubische Spline „natürlich“ genannt.

Wir fragen nun nach der Existenz des interpolierenden kubischen Splines zu vorgegebenen Knotenwerten

$$s_n(x_i) = y_i, \quad i = 0, \dots, n.$$

Satz 2.8 (Spline-Interpolation): Der interpolierende kubische Spline existiert und ist eindeutig bestimmt durch zusätzliche Vorgabe von $s_n''(a)$ und $s_n''(b)$.

Beweis: Jeder kubische Spline (bzgl. der Zerlegung $a = x_0 < \dots < x_n = b$) hat die Form $s(x)|_{[x_{i-1}, x_i]} = p_i(x)$, $i = 1, \dots, n$, mit Polynomen $p_i \in P_3$. Die jeweiligen 4 Koeffizienten dieser n Polynome ergeben $4n$ freie Parameter. Zu ihrer Bestimmung stehen folgende lineare Beziehungen zur Verfügung:

$s(x_i) = y_i, \quad i = 0, \dots, n$:	$2n$	Gleichungen
$s' \in C[a, b]$:	$n - 1$	"
$s'' \in C[a, b]$:	$n - 1$	"
Zusatzbedingungen	:	2	"
Σ	:	$4n$	"

Zum Nachweis der Existenz einer Lösung dieses (quadratischen) Gleichungssystems genügt es wieder zu verifizieren, dass das zugehörige homogene System nur die triviale Lösung hat. Dazu führen wir die folgende Funktionenmenge ein:

$$N \equiv \{w \in C^2[a, b] \mid w(x_i) = 0, i = 0, \dots, n\}.$$

Seien nun $s_n^{(1)}, s_n^{(2)}$ zwei interpolierende Splines. Für die Differenz $s \equiv s_n^{(1)} - s_n^{(2)} \in N$ gilt dann mit beliebigem $w \in N$:

$$\begin{aligned} \int_a^b s''(x)w''(x) dx &= \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} s''(x)w''(x) dx \\ &= \sum_{i=0}^{n-1} \left\{ s''w' \Big|_{x_i}^{x_{i+1}} - s^{(3)}w \Big|_{x_i}^{x_{i+1}} + \int_{x_i}^{x_{i+1}} s^{(4)}(x)w(x) dx \right\} \\ &= \sum_{i=0}^{n-1} s''w' \Big|_{x_i}^{x_{i+1}} = s''(b)w'(b) - s''(a)w'(a) = 0. \end{aligned}$$

Speziell für $w \equiv s \in N$ ist also

$$\int_a^b |s''(x)|^2 dx = 0,$$

d. h.: s ist linear. Wegen $s(a) = s(b) = 0$ folgt $s \equiv 0$.

Q.E.D.

Die obige Orthogonalitätsbeziehung hat die interessante Konsequenz, dass sich der interpolierende Spline durch eine besonders geringe Oszillation auszeichnet.

Satz 2.9: Für den interpolierenden, natürlichen, kubischen Spline s_n gilt

$$\int_a^b |s_n''(x)|^2 dx \leq \int_a^b |f''(x)|^2 dx \quad (2.3.25)$$

bzgl. aller anderen Funktionen $f \in C^2[a, b]$ mit $f(x_i) = y_i, i = 0, \dots, n$.

Beweis: Sei N wieder definiert wie im Beweis von Satz 2.8. Jede interpolierende Funktion $f \in C^2[a, b]$ kann in der Form $f = s_n + w$ mit einem $w \in N$ geschrieben werden. Wir haben (siehe oben)

$$\int_a^b s_n''(x)w''(x) dx = 0 \quad \forall w \in N.$$

Wegen der Identität

$$\begin{aligned} \int_a^b |f''(x)|^2 dx &= \int_a^b |s_n''(x) + w''(x)|^2 dx \\ &= \int_a^b |s_n''(x)|^2 dx + 2 \underbrace{\int_a^b s_n''(x)w''(x) dx}_{=0} + \underbrace{\int_a^b |w''(x)|^2 dx}_{\geq 0} \end{aligned}$$

folgt die Behauptung.

Q.E.D.

Die Aussage von Satz 2.9 lässt sich dahingehend umkehren, dass jede interpolierende Funktion $s \in C^2[a, b]$, $s(x_i) = y_i$, mit der Eigenschaft (2.3.25) notwendig ein natürlicher kubischer Spline ist.

Zur expliziten Berechnung des interpolierenden Splines s_n schreiben wir seine Bestandteile $s_n|_{[x_{i-1}, x_i]} = p_i \in P_3$ in der Form

$$p_i(x) = a_0^{(i)} + a_1^{(i)}(x - x_i) + a_2^{(i)}(x - x_i)^2 + a_3^{(i)}(x - x_i)^3, \quad i = 1, \dots, n,$$

und bestimmen die $4n$ Koeffizienten $a_0^{(i)}, \dots, a_3^{(i)}$. Dies wird im Folgenden für den interpolierenden „natürlichen“ Spline durchgeführt:

– Die Interpolationsbedingung $p_i(x_i) = y_i$, $p_i(x_{i-1}) = y_{i-1}$ impliziert:

$$(1) \quad a_0^{(i)} = y_i, \quad i = 1, \dots, n,$$

und mit $h_i = x_i - x_{i-1}$:

$$(2) \quad y_{i-1} - y_i = -a_1^{(i)}h_i + a_2^{(i)}h_i^2 - a_3^{(i)}h_i^3, \quad i = 1, \dots, n.$$

– Die Randbedingung $p_1''(x_0) = p_n''(x_n) = 0$ impliziert:

$$(3) \quad a_2^{(1)} - 3a_3^{(1)}h_1 = 0, \quad a_2^{(n)} = 0.$$

– Die Stetigkeit der 1. Ableitung $p_i'(x_i) = p_{i+1}'(x_i)$ impliziert:

$$(4) \quad a_1^{(i)} = a_1^{(i+1)} - 2a_2^{(i+1)}h_{i+1} + 3a_3^{(i+1)}h_{i+1}^2, \quad i = 1, \dots, n-1.$$

– Die Stetigkeit der 2. Ableitung $p_i''(x_i) = p_{i+1}''(x_i)$ impliziert:

$$(5) \quad a_2^{(i)} = a_2^{(i+1)} - 3a_3^{(i+1)}h_{i+1}, \quad i = 1, \dots, n-1.$$

Damit haben wir $4n$ Gleichungen (1)-(5) für die $a_k^{(i)}$ gefunden. Zunächst werden nun die $a_1^{(i)}$ und $a_3^{(i)}$ durch die $a_2^{(i)}$ ausgedrückt. Zur Vereinfachung wird $a_2^{(0)} := 0$ und $a_2^{(n+1)} := 0$ gesetzt.

Die Gleichungen (3) und (5) ergeben ($i \rightarrow i-1$):

$$(6) \quad a_3^{(i)} = \frac{a_2^{(i)} - a_2^{(i-1)}}{3h_i}, \quad i = 1, \dots, n.$$

(2) und (6) ergeben:

$$(7) \quad \begin{aligned} a_1^{(i)} &= \frac{y_i - y_{i-1}}{h_i} + a_2^{(i)} h_i - a_3^{(i)} h_i^2 \\ &= \frac{y_i - y_{i-1}}{h_i} + h_i \left\{ a_2^{(i)} - \frac{a_2^{(i)} - a_2^{(i-1)}}{3} \right\} \\ &= \frac{y_i - y_{i-1}}{h_i} + \frac{h_i}{3} \{2a_2^{(i)} + a_2^{(i-1)}\}, \quad i = 1, \dots, n. \end{aligned}$$

(4), (7) und (6) ergeben:

$$\begin{aligned} &\underbrace{\frac{y_i - y_{i-1}}{h_i} + \frac{h_i}{3} \{2a_2^{(i)} + a_2^{(i-1)}\}}_{= a_1^{(i)}} = \\ &= \underbrace{\frac{y_{i+1} - y_i}{h_{i+1}} + \frac{h_{i+1}}{3} \{2a_2^{(i+1)} + a_2^{(i)}\}}_{= a_1^{(i+1)}} - 2a_2^{(i+1)} h_{i+1} + 3a_3^{(i+1)} h_{i+1}^2 \\ &= \frac{y_{i+1} - y_i}{h_{i+1}} + \frac{h_{i+1}}{3} \{2a_2^{(i+1)} + a_2^{(i)}\} - h_{i+1} \{a_2^{(i+1)} + a_2^{(i)}\}, \quad i = 1, \dots, n-1. \end{aligned}$$

Dies wird für $i = 1, \dots, n-1$ umgeschrieben zu

$$h_i a_2^{(i-1)} + 2(h_i + h_{i+1}) a_2^{(i)} + h_{i+1} a_2^{(i+1)} = 3 \left\{ \frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i} \right\}.$$

Damit haben wir für den $(n-1)$ -Vektor $(a_2^{(1)}, \dots, a_2^{(n-1)})^T$ ein $(n-1) \times (n-1)$ -Gleichungssystem aufgestellt; die Koeffizientenmatrix

$$A = \begin{bmatrix} 2(h_1 + h_2) & h_2 & & & 0 \\ h_2 & 2(h_2 + h_3) & \ddots & & \\ & h_3 & \dots & & \\ & & & \ddots & 2(h_{n-2} + h_{n-1}) & h_{n-1} \\ 0 & & & & h_{n-1} & 2(h_{n-1} + h_n) \end{bmatrix}$$

ist „symmetrisch“, d. h.: $a_{ij} = a_{ji}$, und „strikt diagonaldominant“, d. h.:

$$\sum_{\substack{j=1 \\ j \neq i}}^{n-1} |a_{ij}| < |a_{ii}|. \quad (2.3.26)$$

Hieraus folgt nach einem bekannten Satz der Linearen Algebra (s. auch Kapitel 4) ebenfalls die Regularität von A . Durch Lösung dieses Gleichungssystems erhält man zunächst die Koeffizienten $a_2^{(1)}, \dots, a_2^{(n-1)}$; $a_2^{(n)} = 0$ aufgrund der Randbedingungen. Einsetzen der Werte in (7) und (6) liefert dann die anderen Koeffizienten $a_1^{(1)}, \dots, a_1^{(n)}$ sowie $a_3^{(1)}, \dots, a_3^{(n)}$. Zur Lösung des tridiagonalen (symmetrischen) Gleichungssystems kann eine spezielle Variante der Gauß-Elimination verwendet werden (s. Kapitel 4).

Interpolierende Spline-Funktionen besitzen wesentlich bessere Approximationseigenschaften für

$$h := \max_{i=0, \dots, n-1} |x_{i+1} - x_i| \rightarrow 0$$

als die Lagrange-Interpolationspolynome. Allgemein konvergiert

$$\max_{a \leq x \leq b} |f(x) - s_n(x)| \rightarrow 0 \quad (h \rightarrow 0) \quad (2.3.27)$$

sogar für absolutstetiges f mit $\int_a^b |f'(x)|^2 dx < \infty$.

Satz 2.10 (Approximationsfehler): Sei $f \in C^4[a, b]$. Erfüllt der interpolierende kubische Spline $s_n''(a) = f''(a)$ und $s_n''(b) = f''(b)$, so gilt

$$\max_{a \leq x \leq b} |f(x) - s_n(x)| \leq \frac{1}{2} h^4 \max_{a \leq x \leq b} |f^{(4)}(x)|. \quad (2.3.28)$$

Beweis: Siehe Schaback/Werner [5].

Q.E.D.

Neben den guten Approximationseigenschaften weisen Spline-Funktionen auch eine wesentlich bessere Stabilität gegenüber kleinen Störungen in den Interpolationsdaten y_i auf als Lagrange-Polynome; lokale Störungen klingen nach rechts und links schnell ab.

2.4 Trigonometrische Interpolation

In vielen Anwendungsbereichen treten „periodische“ Funktionen, d. h. Funktionen mit der Eigenschaft $f(x + \omega) = f(x)$, $x \in \mathbb{R}$, auf, mit der sog. „Periode“ $\omega > 0$. Hier bietet sich die Interpolation mit „trigonometrischen Summen“ an:

$$t_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^m \left\{ a_k \cos\left(\frac{kx}{\omega}\right) + b_k \sin\left(\frac{kx}{\omega}\right) \right\}, \quad n := 2m,$$

welche ebenfalls ω -periodisch sind. O.B.d.A. können wir im Folgenden $\omega = 2\pi$ annehmen. Das Interpolationsintervall ist dann $[0, 2\pi]$, und die Stützstellen werden äquidistant gewählt zu

$$x_k = k \frac{2\pi}{n+1}, \quad k = 0, \dots, n.$$

Beim Arbeiten mit trigonometrischen Summen erweist sich die „komplexe“ Schreibweise als vorteilhaft.

Satz 2.11 (Trigonometrische Interpolation): Zu gegebenen Zahlen $y_0, \dots, y_n \in \mathbb{C}$ gibt es genau eine Funktion der Gestalt ($i = \sqrt{-1}$)

$$t_n^*(x) = \sum_{k=0}^n c_k e^{ikx}, \quad (2.4.29)$$

welche den Interpolationsbedingungen $t_n^*(x_j) = y_j$ ($j = 0, \dots, n$) genügt. Die Koeffizienten sind bestimmt durch

$$c_k = \frac{1}{n+1} \sum_{j=0}^n y_j e^{-ijx_k}, \quad k = 0, \dots, n. \quad (2.4.30)$$

Beweis: Mit den Abkürzungen

$$w = e^{ix}, \quad w_k = e^{ix_k} = e^{ik2\pi/(n+1)}, \quad k = 0, \pm 1, \dots, \pm n,$$

wird

$$t_n^*(x) = p_n(w) = \sum_{k=0}^n c_k w^k, \quad p_n(\cdot) \in P_n.$$

Offenbar gilt $w_k^{n+1} = e^{ik2\pi} = 1$, d. h.: Die w_k sind sog. $(n+1)$ -te „Einheitswurzeln“. Ferner ist

$$w_k^j = e^{j(ik2\pi)/(n+1)} = e^{k(ij2\pi)/(n+1)} = w_k^j.$$

Die trigonometrische Interpolationsbedingung

$$t_n^*(x_k) = y_k, \quad k = 0, \dots, n,$$

ist damit äquivalent zur polynomialen Interpolationsbedingung (im Komplexen)

$$p_n(w_k) = y_k, \quad k = 0, \dots, n,$$

Diese wiederum ist nach Satz 2.1 durch ein eindeutig bestimmtes Polynom $p_n \in P_n$ erfüllbar. Zur Berechnung der zugehörigen Koeffizienten c_k schreiben wir

$$\sum_{j=0}^n y_j w_k^{-j} = \sum_{j=0}^n p_n(w_j) w_k^{-j} = \sum_{j=0}^n \left(\sum_{l=0}^n c_l w_j^l \right) w_k^{-j} = \sum_{l=0}^n c_l \left(\sum_{j=0}^n w_j^{l-k} \right).$$

Die w_k sind Wurzeln von $w^{n+1} - 1 = (w-1)(w^n + w^{n-1} + \dots + 1) = 0$. Wegen $w_k \neq 1$ für $k = \pm 1, \dots, \pm n$, muss also $\sum_{j=0}^n w_k^j = 0$ sein. Dies ergibt

$$\sum_{j=0}^n w_j^{l-k} = \sum_{j=0}^n w_{l-k}^j = \begin{cases} n+1, & l=k, \\ 0, & l \neq k. \end{cases}$$

Also ist

$$\sum_{j=0}^n y_j w_k^{-j} = c_k(n+1),$$

bzw.

$$c_k = \frac{1}{n+1} \sum_{j=0}^n y_j e^{-ijx_k}.$$

Dies vervollständigt den Beweis.

Q.E.D.

Für Satz 2.11 ist es unwesentlich, ob n gerade oder ungerade ist. Im Folgenden müssen diese beiden Fälle aber unterschieden werden. Wir wollen zeigen, wie mit Hilfe der Aussagen von Satz 2.11 die gestellte trigonometrische Interpolationsaufgabe gelöst werden kann. Dies führt zur sog. „diskreten Fourier¹⁰-Analysis“.

Satz 2.12 (Diskrete Fourier-Analyse): Für $n \in \mathbb{N}_0$ gibt es zu gegebenen reellen Zahlen y_0, \dots, y_n genau ein trigonometrisches Polynom der Form

$$t_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^m \{a_k \cos(kx) + b_k \sin(kx)\} + \frac{\theta}{2} a_{m+1} \cos((m+1)x),$$

mit $t_n(x_j) = y_j$, $j = 0, \dots, n$, wobei

$$\begin{aligned} \theta = 0, \quad m = \frac{1}{2}n, \quad & \text{falls } n \text{ gerade,} \\ \theta = 1, \quad m = \frac{1}{2}(n-1), \quad & \text{falls } n \text{ ungerade.} \end{aligned}$$

Die Koeffizienten sind bestimmt durch

$$a_k = \frac{2}{n+1} \sum_{j=0}^n y_j \cos(jx_k), \quad b_k = \frac{2}{n+1} \sum_{j=0}^n y_j \sin(jx_k).$$

Beweis: (i) Sei t_n^* das Interpolationspolynom nach Satz 2.11:

$$t_n^*(x) = \sum_{k=0}^n c_k e^{ikx}, \quad c_k = \frac{1}{n+1} \sum_{j=0}^n y_j e^{-ijx_k}, \quad k = 0, \dots, n.$$

Wegen der 2π -Periodizität von e^{-ix} gilt

$$e^{-ijx_{n+1-k}} = e^{-ij(n+1-k)2\pi/(n+1)} = e^{-ij2\pi+ijx_k} = e^{ijx_k},$$

¹⁰Jean-Baptiste Baron de Fourier (1768–1830): Französischer Mathematiker und Physiker; Mitglied der Pariser Akademie; lehrte an der École Polytechnique; begleitete Napoleon auf seinem Feldzug nach Ägypten; zählt zu den bedeutendsten Mathematikern des 19. Jahrhunderts; fand bei seinen Arbeiten zur Theorie der Wärmeleitung die Darstellbarkeit periodischer Funktionen durch trigonometrische Reihen.

und folglich für $k = 1, \dots, m$:

$$c_{n+1-k} = \frac{1}{n+1} \sum_{j=0}^n y_j e^{-ijx_{n-1-k}} = \frac{1}{n+1} \sum_{j=0}^n y_j e^{ijx_k} =: c_{-k}.$$

Mit dieser Bezeichnung setzen wir

$$a_k := c_k + c_{-k}, \quad b_k := i(c_k - c_{-k}), \quad k = 1, \dots, m,$$

und $a_{m+1} := 2c_{m+1}$ im Falle $n = 2m + 1$ ungerade, und definieren das trigonometrische Polynom

$$t_n(x) := \frac{1}{2}a_0 + \sum_{k=1}^m \{a_k \cos(kx) + b_k \sin(kx)\} + \frac{\theta}{2}a_{m+1} \cos((m+1)x).$$

(ii) Wir wollen zeigen, dass $t_n^*(x_j) = t_n(x_j) = y_j$, $j = 0, \dots, n$, ist. Zunächst gilt

$$\begin{aligned} t_n(x_j) &= c_0 + \sum_{k=1}^m \{(c_k + c_{-k}) \cos(kx_j) + i(c_k - c_{-k}) \sin(kx_j)\} + \\ &\quad + \theta c_{m+1} \cos((m+1)x_j) \\ &= c_0 + \sum_{k=1}^m c_k \{\cos(kx_j) + i \sin(kx_j)\} + \sum_{k=1}^m c_{-k} \{\cos(kx_j) - \\ &\quad - i \sin(kx_j)\} + \theta c_{m+1} \{\cos((m+1)x_j) + i \sin((m+1)x_j)\}. \end{aligned}$$

Da $\sin((m+1)x_j) = 0$ (wegen $(m+1)x_j = j\pi$ für $n = 2m + 1$ ungerade), folgt bei Beachtung von $e^{iz} = \cos(z) + i \sin(z)$

$$t_n(x_j) = c_0 + \sum_{k=1}^m c_k e^{ikx_j} + \sum_{k=1}^m c_{-k} e^{-ikx_j} + \theta c_{m+1} e^{i(m+1)x_j}.$$

Mit Hilfe von $c_{-k} = c_{n+1-k}$ und $e^{-ikx_j} = e^{ikj2\pi/(n+1)} = e^{i(n+1-k)j2\pi/(n+1)} = e^{i(n+1-k)x_j}$ ergibt sich wie gewünscht

$$t_n(x_j) = \sum_{k=0}^n c_k e^{ikx_j} = y_j.$$

Das trigonometrische Polynom $t_n(\cdot)$ erfüllt also die Interpolationsbedingungen.

(iii) Als nächstes betrachten wir die Koeffizienten a_k und b_k . Unter Verwendung der Beziehungen

$$\sin(z) = \frac{1}{2i}(e^{iz} - e^{-iz}), \quad \cos(z) = \frac{1}{2}(e^{iz} + e^{-iz}),$$

gilt:

$$a_k = c_k + c_{-k} = \frac{1}{n+1} \sum_{j=0}^n y_j \{e^{-ijx_k} + e^{ijx_k}\} = \frac{2}{n+1} \sum_{j=0}^n y_j \cos(jx_k),$$

$$b_k = i(c_k - c_{-k}) = \frac{1}{n+1} \sum_{j=0}^n y_j i \{e^{-ijx_k} - e^{ijx_k}\} = \frac{2}{n+1} \sum_{j=0}^n y_j \sin(jx_k).$$

Dies zeigt die gewünschte Darstellung der Koeffizienten und auch, dass sie reell sind.

(iv) Es bleibt, die Eindeutigkeit des interpolierenden trigonometrischen Polynoms zu zeigen. Die $n+1$ Bedingungen $t_n(x_j) = y_j$, $j = 0, \dots, n$, lassen sich als lineares Gleichungssystem für die $n+1$ unbekanntenen Koeffizienten a_k, b_k auffassen. Da dieses System nach dem eben Gezeigten für alle rechten Seiten y_0, \dots, y_n lösbar ist, sind die Lösungen auch eindeutig. Q.E.D.

Bei der trigonometrischen Interpolation von (stetigen) Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ sind in Satz 2.12 die Werte y_j durch $f(x_j)$ zu ersetzen. Ist f allerdings nicht 2π -periodisch, wird es zunächst einmal zu einer 2π -periodischen Funktion \tilde{f} gemacht (s. Abb. 2.6):

$$\tilde{f}(x) := \begin{cases} f(x), & x \in (0, 2\pi), \\ \frac{1}{2} \{f(0) + f(2\pi)\}, & x = 0, \\ 2\pi\text{-periodisch auf } \mathbb{R} \text{ fortgesetzt.} \end{cases}$$

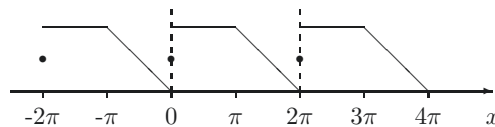


Abbildung 2.6: *Periodische Fortsetzung*

Besitzt f im Intervall $[0, 2\pi]$ eine Unstetigkeitsstelle ζ , so definiert man noch

$$\tilde{f}(\zeta) := \frac{1}{2} \{f(\zeta_+) + f(\zeta_-)\}.$$

Dies ist dadurch motiviert, dass bekanntlich die Fourier-Reihe einer stückweise stetigen Funktion in den Unstetigkeitsstellen gerade gegen diesen Mittelwert konvergiert.

In Anwendungen tritt häufig der Fall auf, dass eine Funktion f nur auf dem Intervall $[0, \pi]$ gegeben ist. Ihre 2π -periodische Fortsetzung auf ganz \mathbb{R} würde dann i. Allg. bei $x = \pi$ eine Unstetigkeit in f oder in f' besitzen, welche die Approximationsgüte des trigonometrischen Interpolationspolynoms auf $[0, 2\pi]$ reduziert. Durch geeignete Wahl der Fortsetzung kann dieser Effekt häufig gemildert werden. Ist f auf $[0, \pi]$ gegeben mit

$f(0) = f(\pi) = 0$, so wird eine 2π -periodische Fortsetzung von f erklärt durch

$$\tilde{f}(x) := \begin{cases} f(x), & x \in [0, \pi], \\ -f(2\pi - x), & x \in [\pi, 2\pi], \\ 2\pi\text{-periodisch auf } \mathbb{R} \text{ fortgesetzt.} \end{cases}$$

Offenbar ist dann \tilde{f} in $x = \pi$ stetig differenzierbar (falls f in $x = \pi$ einseitig differenzierbar war):

$$\lim_{h \rightarrow +0} \frac{\tilde{f}(\pi + h) - \tilde{f}(\pi)}{h} = \lim_{h \rightarrow +0} \frac{-f(\pi - h) + f(\pi)}{h} = f'(\pi)$$

$$\lim_{h \rightarrow +0} \frac{\tilde{f}(\pi) - \tilde{f}(\pi - h)}{h} = \lim_{h \rightarrow +0} \frac{f(\pi) + f(\pi - h)}{h} = f'(\pi).$$

Beispiel 2.7: a) Beispiel einer „ungeraden“ periodischen Fortsetzung:

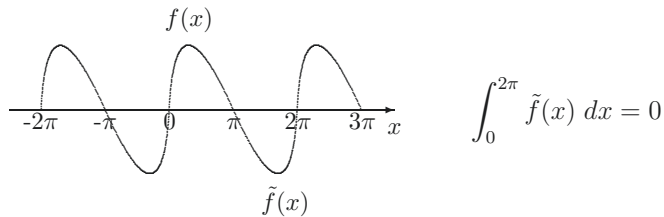


Abbildung 2.7: Ungerade periodische Fortsetzung

b) Beispiel einer „geraden“ periodischen Fortsetzung:

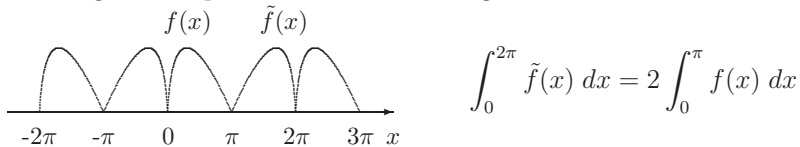


Abbildung 2.8: Gerade periodische Fortsetzung

Je nachdem ob $f : [0, \pi] \rightarrow \mathbb{R}$ als „ungerade“ oder als „gerade“ Funktion 2π -periodisch fortgesetzt wird, ergeben sich bei der trigonometrischen Interpolation von f die folgenden Sonderfälle: (Beachte $f(0) = f(\pi) = 0$.)

a) Gerade Fortsetzung:

$$x_k = k \frac{2\pi}{n+1}, \quad k = 0, \dots, n = 2m + 1$$

$$y_k = \begin{cases} f(x_k), & k = 0, \dots, m, \\ f(2\pi - x_k), & k = m + 1, \dots, n \end{cases} \quad \Rightarrow \quad y_k = y_{n+1-k}, \quad k = 1, \dots, n.$$

Für $k = 1, \dots, m$ folgt:

$$\begin{aligned}
 c_k &= \frac{1}{n+1} \sum_{j=0}^n y_j e^{-ijx_k} = \frac{1}{n+1} \sum_{j=0}^n y_{n+1-j} e^{-ijx_k} \\
 &= \frac{1}{n+1} \sum_{j=0}^n y_j e^{-i(n+1-j)x_k} = c_{n+1-k} = c_{-k}, \\
 b_k &= i(c_k - c_{-k}) = 0, \\
 a_k &= 2c_k = \frac{2}{n+1} \sum_{j=0}^m y_j e^{-ijx_k} + \frac{2}{n+1} \sum_{j=m+1}^n y_j e^{-ijx_k} \\
 &= \frac{2}{n+1} \sum_{j=0}^m y_j e^{-ijx_k} + \frac{2}{n+1} \sum_{j=1}^m \underbrace{y_{n+1-j}}_{=y_j} e^{-i(n+1-j)x_k} \\
 &= \frac{2}{n+1} \sum_{j=0}^m y_j e^{-ijx_k} + \frac{2}{n+1} \sum_{j=1}^m y_j e^{ijx_k} \quad (n = 2m + 1) \\
 &= \frac{4}{n+1} \sum_{j=1}^m y_j \underbrace{\frac{1}{2}(e^{ijx_k} + e^{-ijx_k})}_{=\cos(jx_k)} \quad (y_0 = y_{m+1} = 0) \\
 &= \frac{2}{m+1} \sum_{j=1}^m y_j \cos(jx_k).
 \end{aligned}$$

Die gerade fortgesetzte Funktion lässt sich also durch eine „Cosinus-Summe“ interpolieren:

$$s(x) = \frac{1}{2}a_0 + \sum_{k=1}^{m+1} a_k \cos(kx), \quad a_k = \frac{2}{m+1} \sum_{j=1}^m \tilde{f}(x_j) \cos(jx_k).$$

b) Ungerade Fortsetzung:

$$\begin{aligned}
 x_k &= k \frac{2\pi}{n+1}, \quad k = 0, \dots, n = 2m + 1, \\
 y_k &= \begin{cases} f(x_k), & k = 0, \dots, m \\ -f(2\pi - x_k), & k = m + 1, \dots, n \end{cases} \Rightarrow y_k = -y_{n+1-k}, \quad k = 1, \dots, n
 \end{aligned}$$

Dies ergibt analog wie im Fall (a):

$$c_k = -c_{-k} \Rightarrow a_k = 0, \quad b_k = 2ic_k = \dots$$

Die ungerade fortgesetzte Funktion lässt sich also durch eine „Sinus-Summe“ interpolieren:

$$s(x) = \sum_{k=1}^m b_k \sin(kx), \quad b_k = \frac{2}{m+1} \sum_{j=1}^m \tilde{f}(x_j) \sin(jx_k).$$

2.4.1 Diskrete Fourier-Transformation

Wir diskutieren nun noch die effiziente Berechnung des trigonometrischen Interpolationspolynoms.

Definition 2.7: Die in Satz 2.12 behandelte Aufgabenstellung wird „diskrete Fourier-Analyse“ genannt: Den $n + 1$ Werten $y_j = f(x_j)$, ($j = 0, \dots, n$) einer Funktion $f : [0, 2\pi] \rightarrow \mathbb{R}$ werden (eindeutig) die $n + 1$ Koeffizienten a_k, b_k des trigonometrischen Interpolationspolynoms zugeordnet

$$t_n(x) = \frac{1}{2}a_0 + \sum_{k=1}^m \{a_k \cos(kx) + b_k \sin(kx)\} + \frac{\theta}{2} a_{m+1} \cos((m+1)x).$$

Die Abbildung $\{y_j\} \rightarrow \{a_k, b_k\}$ heißt „diskrete Fourier-Transformation“; sie ist offenbar umkehrbar.

Interpretiert man $\cos(kx)$, $\sin(kx)$ als Grundschwingungsformen eines 2π -periodischen Prozesses $y = f(x)$, so bedeutet eine (diskrete) Fourier-Analyse die Bestimmung der jeweiligen Anteile a_k, b_k dieser Grundschwingungen am Prozeß. Zur Berechnung dieser Koeffizienten a_k, b_k bzw. zur Bestimmung des Polynoms $t_n(\cdot)$ nach den Formeln

$$a_k = c_k + c_{-k}, \quad b_k = i(c_k - c_{-k})$$

sind $n + 1$ Summen der Form

$$c_k := \frac{1}{n+1} \sum_{j=0}^n y_j e^{ijk2\pi/(n+1)} = \sum_{j=0}^n \bar{y}_j w^{jk}, \quad k = 0, \dots, n,$$

zu berechnen mit den Abkürzungen (Man beachte, dass $w^{n+1} = 1$ ist.)

$$\bar{y}_j := \frac{1}{n+1} y_j, \quad w := e^{-i2\pi/(n+1)}.$$

Das Horner-Schema erfordert dazu jeweils n (komplexe) Operationen (1 komplexe Multiplikation und 1 komplexe Addition), d. h.: Insgesamt sind $n^2 + n$ Operationen erforderlich. Für große n und bei mehrfacher Ausführung der Fourier-Analyse bedeutet dies einen beträchtlichen numerischen Aufwand. Im Jahre 1965 gaben Cooley und Tukey¹¹ einen Algorithmus an, der diese Aufgabe wesentlich effizienter löst, die sog. „Schnelle Fourier-Transformation“ („Fast Fourier Transform“ oder auch kurz „FFT“). Wir erläutern diese für den Spezialfall $n + 1 = 2^p$, $p \in \mathbb{N}$:

¹¹John Wilder Tukey (1915–2000): US-Amerikanischer Mathematiker; arbeitete seit 1945 an der Princeton University, seit 1956 als Direktor der Statistics Research Group; bekannt vor allem durch die gemeinsam mit J. W. Cooley (IBM Cooperation) entwickelte FFT: An algorithm for the machine calculation of complex Fourier series, Math. Comput. 19, 297-301 (1965); wichtige Beiträge auch zur praktischen Statistik.

Idee der FFT: Durch Aufspaltung der Summe

$$\begin{aligned} c_k &= \bar{y}_0 w^{0k} + \bar{y}_1 w^{1k} + \bar{y}_2 w^{2k} + \bar{y}_3 w^{3k} + \dots + \bar{y}_{2^{p-2}} w^{(2^{p-2})k} + \bar{y}_{2^{p-1}-1} w^{(2^{p-1}-1)k} \\ &= \left\{ \underbrace{\bar{y}_0 (w^2)^{0k}} + \underbrace{\bar{y}_1 (w^2)^{0k} w^k} \right\} + \left\{ \underbrace{\bar{y}_2 (w^2)^{1k}} + \underbrace{\bar{y}_3 (w^2)^{1k} w^k} \right\} + \dots \\ &\quad \dots + \left\{ \underbrace{\bar{y}_{2^{p-2}} (w^2)^{(2^{p-1}-1)k}} + \underbrace{\bar{y}_{2^{p-1}-1} (w^2)^{(2^{p-1}-1)k} w^k} \right\} \end{aligned}$$

bzgl. gerader („—“) und ungerader („ \dots “) Indizes erhält man

$$c_k = \sum_{j=0}^{2^{p-1}-1} \bar{y}_{2j} (w^2)^{jk} + \sum_{j=0}^{2^{p-1}-1} \bar{y}_{2j+1} (w^2)^{jk} w^k.$$

Sei $k_1 \in \{0, 1, \dots, 2^{p-1}-1\}$ der ganzzahlige Rest bei Division von k durch 2^{p-1} :

$$k \equiv k_1 \pmod{2^{p-1}}.$$

Wegen $w^{2^p} = 1$ gilt mit einem geeigneten $\lambda \in \mathbb{N}$

$$(w^2)^{jk} = (w^2)^{\lambda 2^p j} (w^2)^{jk_1} = (w^{2^p})^{\lambda j} (w^2)^{jk_1} = (w^2)^{jk_1}.$$

Folglich ist

$$c_k = \sum_{j=0}^{2^{p-1}-1} \bar{y}_{2j} (w^2)^{jk_1} + w^k \sum_{j=0}^{2^{p-1}-1} \bar{y}_{2j+1} (w^2)^{jk_1}$$

bzw.

$$c_k = \tilde{c}_{k_1} + w^k \bar{c}_{k_1}, \quad k = 0, \dots, 2^p - 1, k \equiv k_1 \pmod{2^{p-1}},$$

mit den Teilsummen

$$\tilde{c}_{k_1} := \sum_{j=0}^{2^{p-1}-1} \bar{y}_{2j} (w^2)^{jk_1}, \quad \bar{c}_{k_1} := \sum_{j=0}^{2^{p-1}-1} \bar{y}_{2j+1} (w^2)^{jk_1}, \quad k_1 = 0, \dots, 2^{p-1}-1.$$

Zur Berechnung der $n+1 = 2^p$ Größen c_k genügt es also, die $2 \cdot 2^{p-1}$ Größen \tilde{c}_{k_1} , \bar{c}_{k_1} zu bestimmen, d. h.: Die Fourier-Analyse mit $n+1 = 2^p$ Termen wird ersetzt durch zwei Fourier-Analysen mit jeweils 2^{p-1} Termen. Anwendung derselben Aufspaltung auf \tilde{c}_{k_1} und \bar{c}_{k_1} ergibt vier Fourier-Analysen mit jeweils 2^{p-2} Termen u.s.w.. Am Schluss verbleiben als Startpunkt des Algorithmus 2^p Fourier-Analysen mit jeweils einem Term. Diese „trivialen“ Fourier-Analysen ordnen den einzelnen Stützpunkten x_k , ($k = 1, \dots, n$) gerade die Werte y_k zu, welche die jeweiligen Koeffizienten der trigonometrischen Approximation 0-ter Ordnung durch konstante Funktionen darstellen.

Satz 2.13 (Schnelle Fourier-Transformation): *Im Fall $n+1 = 2^p$ löst die „Schnelle Fourier-Transformation“ das Problem der Berechnung der $n+1$ Fourier-Koeffizienten $\{c_0, \dots, c_n\}$, mit $2(n+1) \log_2(n+1)$ (komplexen) arithmetischen Operationen.*

Beweis: Sei r_p die Anzahl der (komplexen) Operationen zur Berechnung aller Koeffizienten $\{c_0, \dots, c_n\}$ im Falle $n+1 = 2^p$. Sind die Größen $\tilde{c}_{k_1}, \bar{c}_{k_1}$ ($k_1 = 0, \dots, 2^{p-1} - 1$) bekannt, so erfordert die Berechnung der Potenzen w^k ($k = 1, \dots, n$) und der Koeffizienten $c_k = \tilde{c}_{k_1} + w^k \bar{c}_{k_1}$ ($k = 0, \dots, n$) offenbar höchstens $(n-1) + (n+1) = 2n \leq 2 \cdot 2^p$ Operationen. (Eine große Ersparnis wäre durch vorausgehende Berechnung und Speicherung der w^k zu erzielen!) Also wird

$$r_p \leq 2 \cdot r_{p-1} + 2 \cdot 2^p, \quad p = 1, 2, 3, \dots$$

Wir wollen zeigen, dass $r_p \leq 2p \cdot 2^p$ gilt. Ausgehend von $r_0 = 0$ folgt durch Induktion:

$$r_p \leq 2 \cdot r_{p-1} + 2 \cdot 2^p = 2 \cdot (2(p-1) \cdot 2^{p-1}) \leq 2p \cdot 2^p = 2 \log_2(n+1) (n+1),$$

was den Beweis vervollständigt.

Q.E.D.

Beispiel 2.8: $n = 2^7 - 1 = 127$

$$n^2 + n = 16.256, \quad 2(n+1) \underbrace{\log_2(n+1)}_{\sim 7} = 1.792.$$

Implementierung: Bei der konkreten Implementierung der FFT geht man genau entgegengesetzt zu ihrer obigen Herleitung vor. Zunächst werden die 2^p ein-elementigen Fourier-Analysen durchgeführt, welche die gegebenen Stützwerte y_j , ($j = 1, \dots, 2^p - 1$) verwenden. Danach werden nur noch die rekursiven Formeln

$$c_k = \tilde{c}_{k_1} + w^k \bar{c}_{k_1}, \quad k = 0, \dots, 2^p - 1, \quad k \equiv k_1 \pmod{2^{p-1}},$$

abgearbeitet.

2.5 Gauß-Approximation

Wir fassen im Folgenden die Menge $C[a, b]$ der über einem Intervall $[a, b]$ stetigen reell- bzw. komplex-wertigen Funktionen als einen (unendlich dimensionalen) Vektorraum über dem Zahlkörper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ auf.

Bemerkung 2.3: Die Aussagen dieses Abschnitts gelten sinngemäß auch für den Vektorraum $R[a, b]$ der über dem Intervall $[a, b]$ Riemann-integrierbaren Funktionen oder sogar für den Vektorraum $L^2(a, b)$ der über (a, b) im Lebesgue-Sinne quadrat-integrierbaren Funktionen.

Gegeben sei eine Funktion $f \in C[a, b]$ sowie ein endlich dimensionaler Teilraum $S \subset C[a, b]$, dessen Elemente zur Approximation von f dienen sollen, z. B.: $S = P_n$, Raum

der Polynome vom Grad $\leq n$. Im Gegensatz zur Interpolation verwendet die „Gauß-Approximation“ das sog. „quadratische Mittel“

$$\|f\| \equiv \left(\int_a^b |f(x)|^2 dx \right)^{1/2}$$

als Maß für die Güte einer Approximation. d. h.: Gesucht ist ein $g \in S$, so dass

$$\|f - g\| = \min_{\varphi \in S} \|f - \varphi\|. \quad (2.5.31)$$

Ein $g \in S$ mit dieser Eigenschaft heißt dann „Bestapproximation“ von f (in S bzgl. $\|\cdot\|$). Durch $\|\cdot\|$ ist auf $C[a, b]$ eine „Norm“ gegeben, d. h. eine Funktion $\|\cdot\| : C[a, b] \rightarrow \mathbb{R}_+$ mit analogen Eigenschaften wie die wohlbekannte euklidische Vektornorm auf dem \mathbb{K}^n . Es sei an die folgenden Eigenschaften einer Norm erinnert:

1. Definitheit: $\|f\| \in \mathbb{R}_+$, $\|f\| = 0 \Rightarrow f = 0$.
2. Sublinearität: $\|f + g\| \leq \|f\| + \|g\|$ (Dreiecksungleichung),
3. Homogenität: $\|\alpha f\| = |\alpha| \|f\|$, $\alpha \in \mathbb{K}$.

Das Analogon zum euklidischen Skalarprodukt ist in diesem Fall das „ L^2 -Skalarprodukt“

$$(f, g) \equiv \int_a^b f(t) \overline{g(t)} dt, \quad (f, f) = \|f\|^2.$$

Hierfür liegen wieder die für ein „Skalarprodukt“ charakteristischen Eigenschaften vor:

1. Definitheit: $(f, f) \in \mathbb{R}_+$, $(f, f) = 0 \Rightarrow f = 0$.
2. Linearität: $(\alpha f + g, h) = \alpha(f, g) + (h, g)$, $\alpha \in \mathbb{K}$,
3. Symmetrie: $(f, g) = \overline{(g, f)}$.

Nicht jede Norm gehört zu einem Skalarprodukt; z. B.: die sog. „ L^p -Normen“ $\|\cdot\|_p$ für $p \in [1, \infty) \setminus \{2\}$, sowie die sog. „Maximumnorm“ $\|\cdot\|_\infty$:

$$\|f\|_p := \left(\int_a^b \|f(x)\|^p dx \right)^{1/p}, \quad \|f\|_\infty := \max_{a \leq x \leq b} |f(x)|.$$

Wir notieren noch die wichtige „Höldersche¹² Ungleichung“ für Skalarprodukte (auch „Schwarzsche¹³ Ungleichung“ im Fall allgemeiner Skalarprodukte)

$$|(f, g)| \leq \|f\| \|g\|.$$

¹²Ludwig Otto Hölder (1859–1937): Deutscher Mathematiker; Professor in Tübingen; Beiträge zunächst zur Theorie der Fourier-Reihen und später vor allem zur Gruppentheorie; fand 1884 die nach ihm benannte Ungleichung.

¹³Hermann Schwarz (1843–1921): Deutscher Mathematiker; wirkte in Halle, Göttingen und Berlin; leistete grundlegende Arbeiten zur Funktionentheorie, Differentialgeometrie und Variationsrechnung.

Versehen mit dem L^2 -Skalarprodukt wird $C[a, b]$ zu einem sog. „unitären Raum“. Für die Gauß¹⁴-Approximation in unitären Räumen haben wir die folgende fundamentale Aussage:

Satz 2.14 (Allgemeine Gauß-Approximation): *Seien H ein unitärer Raum und $S \subset H$ ein endlich dimensionaler Teilraum. Dann existiert zu jedem $f \in H$ eine eindeutig bestimmte „beste Approximation“ $g \in S$:*

$$\|f - g\| = \min_{\varphi \in S} \|f - \varphi\|. \quad (2.5.32)$$

Beweis: (i) Wir wollen die Eigenschaft der besten Approximation zunächst durch eine etwas handlichere Bedingung charakterisieren. Sei $g \in S$ eine beste Approximation. Dann besitzt für beliebiges, fest gewähltes $\varphi \in S$ die quadratische Funktion

$$F_\varphi(t) := \|f - g - t\varphi\|^2, \quad t \in \mathbb{R},$$

bei $t = 0$ ein Minimum. Folglich ist

$$\frac{d}{dt} F(t)|_{t=0} = \frac{d}{dt} \|f - g - t\varphi\|^2|_{t=0} = 0.$$

Ausgeschrieben bedeutet dies $(f - g - t\varphi, \varphi)|_{t=0} = 0$ und folglich

$$(f - g, \varphi) = 0 \quad \forall \varphi \in S. \quad (2.5.33)$$

Diese Beziehung kann man so interpretieren, dass der Fehler $f - g$ auf dem approximierenden Teilraum S „orthogonal“ ist bzgl. des L^2 -Skalarprodukts (\cdot, \cdot) .

(ii) Genüge nun umgekehrt $g \in S$ der Bedingung (2.5.33). Dann gilt mit einem beliebigen $\varphi \in S$:

$$\|f - g\|^2 = (f - g, f - g) = (f - g, f - \varphi) + (f - g, \varphi - g) \leq \|f - g\| \|f - \varphi\|$$

und folglich

$$\|f - g\| \leq \inf_{\varphi \in S} \|f - \varphi\|,$$

d. h.: g ist auch beste Approximation.

(iii) Eindeutigkeit der besten Approximation: Seien $g_1, g_2 \in S$ zwei Bestapproximationen. Dann gilt notwendig

$$(f - g_1, \varphi) = 0 = (f - g_2, \varphi) \quad \forall \varphi \in S,$$

und folglich

$$(g_1 - g_2, \varphi) = 0 \quad \forall \varphi \in S.$$

¹⁴Carl Friedrich Gauß (1777–1855): Bedeutender deutscher Mathematiker, Astronom und Physiker; wirkte in Göttingen; fundamentale Beiträge zur Arithmetik, Algebra und Geometrie, Begründer der modernen Zahlentheorie, Bestimmung von Planetenbahnen durch „Gauß-Ausgleich“, Arbeiten zum Erdmagnetismus und Konstruktion eines elektromagnetischen Telegraphen.

Wählen wir $\varphi := g_1 - g_2 \in S$, ergibt sich $\|g_1 - g_2\|^2 = 0$ und somit $g_1 = g_2$.

(iv) Existenz der besten Approximation: Der endlich dimensionale Teilraum $S \subset H$ besitzt eine Basis $\{\psi_1, \dots, \psi_n\}$, $n := \dim H$, bzgl. derer sich die gesuchte besten Approximation $g \in S$ darstellen lässt in der following Form:

$$g = \sum_{k=1}^n \alpha_k \psi_k.$$

Einsetzen dieses Ansatzes in die notwendige Orthogonalitätsbedingung (2.5.33) ergibt

$$\left(f - \sum_{i=1}^n \alpha_i \psi_i, \varphi\right) = (f, \varphi) - \sum_{k=1}^n \alpha_k (\psi_k, \varphi) = 0 \quad \forall \varphi \in S.$$

Dies ist bei sukzessiver Wahl von $\varphi := \psi_i$ für $i = 1, \dots, n$, äquivalent zu dem linearen $n \times n$ -Gleichungssystem

$$\sum_{k=1}^n (\psi_k, \psi_i) \alpha_k = (f, \psi_i), \quad i = 1, \dots, n. \quad (2.5.34)$$

Mit der Notation

$$\alpha := (\alpha_k)_{k=1}^n, \quad b := ((f, \psi_i))_{i=1}^n, \quad A := ((\psi_k, \psi_i))_{i,k=1}^n,$$

lässt sich dies in der kompakten Form $A\alpha = b$ schreiben. Die Matrix A ist als „Gramsche Matrix“ der Basis $\{\psi_1, \dots, \psi_n\}$ regulär. Dies ersieht man aus der Beziehung

$$\bar{\alpha}^T A \alpha = \sum_{i,k=1}^n \bar{\alpha}_i \alpha_k (\psi_k, \psi_i) = (g, g),$$

welche die Injektivität von A impliziert. Ferner ist A offenbar symmetrisch“ (im Fall $\mathbb{K} = \mathbb{R}$) bzw. „hermitesch“ (im Fall $\mathbb{K} = \mathbb{C}$) und folglich „positiv definit“. Das Gleichungssystem $A\alpha = b$ ist also für jede rechte Seite b , d. h. für jedes $f \in H$ eindeutig lösbar. Folglich bestimmt die Orthogonalitätsbedingung (2.5.33) eindeutig ein Element $g \in S$, welches dann notwendig eine Bestapproximation von f ist. Q.E.D.

Zur Konstruktion der besten Approximation $g \in S$ zu einer Funktion $f \in H$ kann zunächst das Gleichungssystem (2.5.34) dienen:

$$\sum_{k=1}^n (\psi_k, \psi_i) \alpha_k = (f, \psi_i), \quad i = 1, \dots, n. \quad (2.5.35)$$

Es besitzt eine besonders einfache Lösung, wenn die Basis $\{\psi_1, \dots, \psi_n\}$ ein „Orthonormalsystem“ („ONS“) ist, d. h.: $(\psi_k, \psi_i) = \delta_{ki}$. Dann gilt offenbar

$$\alpha_k = (f, \psi_k), \quad k = 1, \dots, n,$$

d. h.: Die beste Approximation ist explizit gegeben durch

$$g = \sum_{k=1}^n (f, \psi_k) \psi_k. \quad (2.5.36)$$

Hilfssatz 2.1 (Gram-Schmidt-Algorithmus): Zu jeder Basis $\{\psi_1, \dots, \psi_n\}$ von S lässt sich ein Orthonormalsystem mit dem „Gram¹⁵-Schmidt¹⁶-Algorithmus“ konstruieren:

$$\begin{aligned} \tilde{\varphi}_1 &:= \psi_1, & \varphi_1 &:= \|\tilde{\varphi}_1\|^{-1} \tilde{\varphi}_1, \\ k = 2, \dots, n : & \quad \tilde{\varphi}_k &:= \psi_k - \sum_{i=1}^{k-1} (\psi_k, \varphi_i) \varphi_i, & \quad \varphi_k &:= \|\tilde{\varphi}_k\|^{-1} \tilde{\varphi}_k. \end{aligned}$$

Das Ergebnis ist ein Orthonormalsystem $\{\varphi_1, \dots, \varphi_n\}$ in S .

Beweis: Dies zeigt man durch Induktion nach $n = \dim S$. Im Fall $\psi_1 \neq 0$ ist φ_1 wohldefiniert. Sei nun $\{\varphi_1, \dots, \varphi_{n-1}\}$ wohldefiniert und ONS. Im Fall

$$\tilde{\varphi}_n = \psi_n - \sum_{k=1}^{n-1} (\psi_n, \varphi_k) \varphi_k = 0$$

wäre $\{\psi_1, \dots, \psi_n\}$ linear abhängig, im Widerspruch zur Annahme. Also ist φ_n wohldefiniert. Weiter ist für $k = 1, \dots, n-1$:

$$(\varphi_n, \varphi_k) = (\psi_n, \psi_k) - \sum_{i=1}^{n-1} (\psi_n, \varphi_i) \underbrace{(\varphi_i, \varphi_k)}_{= \delta_{ik}} = 0,$$

und $\|\varphi_n\| = 1$ nach Konstruktion.

Q.E.D.

Im Folgenden wollen wir die obigen allgemeinen Aussagen zur Gauß-Approximation mit Polynomen anwenden. O.B.d.A. legen wir das Intervall $[a, b] = [-1, 1]$ zugrunde (gegebenenfalls Variablentransformation). Nach Satz 2.14 existiert zu jedem $f \in C[-1, 1]$ die eindeutig bestimmte beste Approximation $g \in S = P_n$. Zu ihrer Berechnung sei zunächst die Basis $\{1, x, \dots, x^n\}$ von P_n herangezogen. Die Koeffizienten in der Darstellung $g = \sum_{k=0}^n \alpha_k x^k$ ergeben sich dann als Lösung eines linearen Gleichungssystems mit der Koeffizientenmatrix $A = (a_{ik})_{i,k=0}^n$, wobei

¹⁵Jørgen Pedersen Gram (1850–1916): Dänischer Mathematiker, Mitarbeiter und später Eigentümer einer Versicherungsgesellschaft, Beiträge zur Algebra (Invariantentheorie), Wahrscheinlichkeitstheorie, Numerik und Forstwissenschaft; das u.ä. nach ihm benannte Orthogonalisierungsverfahren geht aber wohl auf Laplace zurück und wurde bereits von Cauchy 1836 verwendet.

¹⁶Erhard Schmidt (1876–1959): Deutscher Mathematiker, Professor in Berlin, Gründer des dortigen Instituts für Angewandte Mathematik 1920, nach dem Krieg Direktor des Mathematischen Instituts der Akademie der Wissenschaften der DDR; Beiträge zur Theorie der Integralgleichungen und der Hilbert-Räume sowie später zur Topologie.

$$A = 2 \begin{bmatrix} 1 & 0 & 1/3 & 0 & 1/5 & \cdots \\ 0 & 1/3 & 0 & 1/5 & & \\ 1/3 & 0 & 1/5 & 0 & & \\ 0 & 1/5 & & \ddots & & \\ 1/5 & & & & \ddots & \\ \vdots & & & & & \ddots \end{bmatrix}.$$

$$a_{ik} = \int_{-1}^1 x^k x^i dx = \begin{cases} 0 & , \text{ falls } i+k \text{ ungerade} \\ \frac{2}{i+k+1} & , \text{ falls } i+k \text{ gerade} \end{cases}$$

Diese Matrix (sog. Hilbert¹⁷-Matrix) ist zwar regulär, doch ist ihre Invertierung so extrem schlecht konditioniert, dass für große n die Berechnung von g auf diesem Wege unmöglich ist. Statt dessen wird die Basis $\{1, \dots, x^n\}$ bzgl. des (reellen) Skalarprodukts ($\mathbb{K} = \mathbb{R}$)

$$(f, g) := \int_{-1}^1 f(x)g(x) dx$$

orthonormalisiert. Das Ergebnis fassen wir in folgendem Satz zusammen.

Satz 2.15 (Legendre-Polynome): *Durch Orthogonalisierung der natürlichen Monombasis $\{1, x, \dots, x^n\}$, $n \in \mathbb{N}$, mit dem Gram-Schmidt-Algorithmus ergeben sich Polynome $p_k \in P_k$ (nicht normalisiert), welche sich in der Form*

$$p_k(x) = \frac{k!}{(2k)!} \frac{d^k}{dx^k} (x^2 - 1)^k, \quad k = 0, 1, \dots, n, \quad (2.5.37)$$

darstellen lassen. Für sie gilt die zweistufige Rekursionsformel

$$\begin{aligned} p_0(x) &\equiv 1, & p_1(x) &= x, \\ p_{k+1}(x) &= xp_k(x) - \frac{k^2}{4k^2-1} p_{k-1}(x), & k &= 1, 2, \dots, n-1. \end{aligned} \quad (2.5.38)$$

sowie

$$\|p_k\| = \frac{k!^2}{(2k)!} \sqrt{\frac{2^{2k+1}}{2k+1}}, \quad p_k(1) = \frac{2^k k!^2}{(2k)!}. \quad (2.5.39)$$

¹⁷David Hilbert (1862–1943): Bedeutender deutscher Mathematiker; wirkte in Königsberg und Göttingen; begründete u. a. den axiomatischen Aufbau der Mathematik; zum Wesen der Axiomatik (in der Geometrie) sagte er „Man muss jederzeit anstelle von Punkten, Geraden, Ebenen - Tische, Stühle, Bierseidel sagen können“.

Durch Normierung bei $x = 1$ erhält man die sog. „Legendre¹⁸-Polynome“

$$L_k(x) := \frac{(2k)!}{2^k k!^2} p_k(x), \quad L_k(1) = 1, \quad k = 0, 1, 2, \dots, n. \quad (2.5.40)$$

Beweis: Wir führen den Beweis in mehreren Schritten, wobei einige Rechnungen als Übungsaufgabe gestellt sind.

(i) Wir zeigen zunächst, dass die durch (2.5.37) definierten Polynome $p_k \in P_k$ bzgl. des Skalarprodukts (\cdot, \cdot) orthogonal sind. Dies ergibt sich durch partielle Integration über dem Intervall $[-1, 1]$ (Übungsaufgabe). Analog erschließen wir die Beziehungen (2.5.39).

(ii) Der führende Term von $p_k(x)$ ist gemäß der Definition x^k . Folglich sind die p_k gerade die durch den Gram-Schmidt-Algorithmus aus der Monombasis erzeugten orth. Polynome.

(iii) Das Polynom $(x^2 - 1)^k$ ist eine gerade Funktion. Seine k -ten Ableitungen sind dann ungerade oder gerade je nachdem, ob k ungerade oder gerade ist. Folglich ist

$$p_k(x) = (-1)^k p_k(-x).$$

(iv) Wegen $p_{k+1}(x) = x^{k+1} + \dots$ ist $p_{k+1}(x) - xp_k(x) = \gamma_k x^k + \gamma_{k-1} x^{k-1} + \dots + \gamma_0$ mit gewissen Koeffizienten $\gamma_k, \dots, \gamma_0$. Ist nun $k+1$ gerade, so ist das Polynom $p_{k+1}(x) - xp_k(x)$ gerade aber x^k ungerade, so dass notwendig $\gamma_k = 0$ sein muss. Es gibt daher eine Darstellung

$$p_{k+1}(x) - xp_k(x) = \sum_{i=0}^{k-1} \gamma_i p_i(x)$$

mit den Polynomen p_0, \dots, p_{k-1} , die ja als Orthogonalsystem eine Basis von P_{k-1} bilden. Wegen der Orthogonalität der p_k folgt dann für $j = 0, \dots, k-2$:

$$0 = (p_{k+1} - xp_k, L_j) = \sum_{i=0}^{k-1} \gamma_i (p_i, p_j) = \gamma_j \|p_j\|^2,$$

bzw. $\gamma_0 = \dots = \gamma_{k-2} = 0$. Es besteht also eine zweistufige Rekursion der Form

$$p_{k+1}(x) = xp_k(x) + \gamma_{k-1} p_{k-1}(x).$$

Zur Bestimmung des Koeffizienten γ_{k-1} verwenden wir $p_k(1) = \frac{k!^2}{(2k)!} 2^k$. Es ergibt sich

$$\begin{aligned} \gamma_{k-1} &= \frac{p_{k+1}(1) - p_k(1)}{p_{k-1}(1)} = \frac{\frac{(k+1)!^2}{(2k+2)!} 2^{k+1} - \frac{k!^2}{(2k)!} 2^k}{\frac{(k-1)!^2}{(2k-2)!} 2^{k-1}} \\ &= \frac{4k^2(k+1)^2 - 2k^2(2k+2)(2k+1)}{(2k+2)(2k+1)2k(2k-1)} = \frac{k(k+1) - k(2k+1)}{(2k+1)(2k-1)} = -\frac{k^2}{4k^2-1}, \end{aligned}$$

wie behauptet.

Q.E.D.

¹⁸Adrien-Marie Legendre (1752–1833): Französischer Mathematiker; Mitglied der Pariser Akademie der Wissensch.; Beiträge zur Himmelsmechanik, Zahlentheorie und Geometrie.

Die Rekursionsformel (2.5.38) für die Polynome p_n ist ein Spezialfall eines allgemeinen Resultats für „orthogonale Polynome“; siehe den folgenden Satz 2.17.

Die Gauß-Approximation mit orthogonalen Polynomen hat den Vorteil der formal einfachen Berechenbarkeit der besten Approximation

$$g(x) = \sum_{k=0}^n \left(\int_{-1}^1 |p_k(x)|^2 dx \right)^{-1} \left(\int_{-1}^1 f(\xi) p_k(\xi) d\xi \right) p_k(x).$$

Die Berechnung der Koeffizienten erfordert i. Allg. numerische Quadratur.

Die Maximalabweichung der Gauß-Approximierenden

$$\|f - g\|_\infty = \max_{-1 \leq x \leq 1} |f(x) - g(x)|$$

wird jedoch i. Allg. groß; insbesondere in der Nähe der Intervallenden treten große Fehler auf. Zur Unterdrückung dieses Defektes verwendet man das gewichtete Skalarprodukt

$$(f, g)_\omega = \int_{-1}^1 f(x)g(x)\omega(x) dx, \quad \omega(x) \equiv \frac{1}{\sqrt{1-x^2}},$$

wodurch eine stärkere Bindung in der zugehörigen Fehlernorm

$$\|f - g\|_\omega = \left(\int_{-1}^1 |f(x) - g(x)|^2 \frac{dx}{\sqrt{1-x^2}} \right)^{1/2}$$

an den Intervallenden impliziert wird. Zur Anwendung von Satz 2.14 wird die Basis $\{1, x, \dots, x^n\}$ von P_n nun bzgl. dieses neuen Skalarproduktes orthogonalisiert.

Satz 2.16 (Tschebyscheff-Polynome): *Durch Orthogonalisierung der natürlichen Monombasis $\{1, x, \dots, x^n\}$, $n \in \mathbb{N}$, bzgl. des gewichteten Skalarprodukts $(\cdot, \cdot)_\omega$ mit dem Gram-Schmidt-Algorithmus ergeben sich Polynome $p_k \in P_k$ (nicht normalisiert), welche sich in der Form*

$$p_0(x) \equiv 1, \quad p_k(x) = 2^{k-1} \cos[k \arccos(x)], \quad k = 1, 2, \dots, n, \quad (2.5.41)$$

darstellen lassen. Für sie gilt die zweistufige Rekursionsformel

$$\begin{aligned} p_0(x) &\equiv 1, \quad p_1(x) = x, \\ p_{k+1}(x) &= 4xp_k(x) - 4p_{k-1}(x), \quad k = 1, 2, \dots, n, \end{aligned} \quad (2.5.42)$$

sowie

$$\|p_k\|_\omega = \begin{cases} \sqrt{\pi}, & k = 0 \\ \sqrt{\pi/2}, & k \neq 0, \end{cases} \quad p_k(1) = 2^{k-1}. \quad (2.5.43)$$

Durch Normierung bei $x = 1$ erhält man die sog. „Tschebyscheff¹⁹-Polynome“

$$T_k(x) = \cos[k \arccos(x)], \quad T_k(1) = 1, \quad k = 0, 1, 2, \dots \quad (2.5.44)$$

Beweis: Zunächst gilt für die Funktionen $g_k := \cos[k \arccos(x)]$ im Fall $k = 0, 1$: $g_0 \equiv 1$, $g_1(x) = x$. Weiter folgt aus der Identität

$$\cos((n+1)x) + \cos((n-1)x) = 2 \cos(x) \cos(nx)$$

die rekursive Beziehung

$$\begin{aligned} g_{k+1}(x) &= \cos[(k+1) \arccos(x)] \\ &= 2 \cos[\arccos(x)] \cos[k \arccos(x)] - \cos[(k-1) \arccos(x)] \\ &= 2xg_k - g_{k-1}. \end{aligned}$$

Weiter erhält man mit Hilfe der Variablentransformation $x = \cos(\theta)$ mit

$$dx = -\sin \theta d\theta = -\sqrt{1 - \cos^2(\theta)} d\theta = -\sqrt{1 - x^2} d\theta$$

die Beziehung

$$\int_{-1}^1 g_k(x)g_j(x)\omega(x) dx = -\int_{\pi}^0 \cos(k\theta) \cos(j\theta) d\theta = \begin{cases} \pi, & k = j = 0 \\ \pi/2, & k = j \neq 0 \\ 0, & k \neq j \end{cases}.$$

Hieraus entnehmen wir, dass die g_k tatsächlich Polynome k -ten Grades über $[-1, 1]$ sind, paarweise orthogonal bzgl. $(\cdot, \cdot)_{\omega}$ sind und den führenden Koeffizienten $g_k(x) = 2^{k-1}x^k + \dots$ haben. Die skalierten Polynome $p_0 \equiv 1$ und $p_k := 2^{k-1}g_k$ haben dann den führenden Koeffizienten $p_k(x) = x^k + \dots$ und genügen der zweistufigen Rekursionsformel

$$p_{k+1}(x) = 4xp_k - 4p_{k-1}.$$

Sie sind also gerade die durch das Gram-Schmidtsche Orthogonalisierungsverfahren erzeugten Polynome. Q.E.D.

Die beste Approximation einer Funktion $f \in C[-1, 1]$ in P_n bzgl. des gewichteten Skalarproduktes $(\cdot, \cdot)_{\omega}$ hat also die Gestalt

$$g = \sum_{k=0}^n \alpha_k T_k(x)$$

¹⁹Pafnuty Lvovich Tschebyscheff (russ.: Chebyshev) (1821–1894): Russischer Mathematiker; Professor in St. Petersburg; Beiträge zur Zahlentheorie, Wahrscheinlichkeitstheorie und vor allem zur Approximationstheorie; entwickelte eine allgemeine Theorie orthogonaler Polynome.

mit den Koeffizienten

$$\alpha_0 = \frac{1}{\pi} \int_{-1}^1 f(x) \omega(x) dx, \quad \alpha_k = \frac{2}{\pi} \int_{-1}^1 f(x) T_k(x) \omega(x) dx, \quad k = 1, \dots, n.$$

Das Bestehen von zweistufigen Rekursionsformeln für die Gauß-Legendre sowie für die Tschebyscheff-Polynome legt nahe, dass dies vielleicht generell für orthogonale Polynomen bzgl. Skalarprodukten der betrachteten Form $(\cdot, \cdot)_\omega$ gilt. Dazu gilt der folgende Satz.

Satz 2.17 (Orthogonale Polynome): *Das allgemeine Skalarprodukt (\cdot, \cdot) auf $C[-1, 1]$ habe die auf dem Vektorraum P der Polynome die Symmetrieeigenschaft*

$$(p, xq) = (xp, q) \quad \forall p, q \in P. \quad (2.5.45)$$

Dann genügen die durch das Gram-Schmidt-Verfahren aus der Basis $\{1, x, x^2, \dots\}$ gewonnenen orthogonalen Polynome p_k , $k = 0, 1, 2, \dots$ (nicht normalisiert), den folgenden rekursiven Beziehungen beginnend mit $p_0(x) \equiv 1$, $p_1(x) = x - \beta_0$:

$$p_{k+1}(x) = (x - \beta_k)p_k - \gamma_k p_{k-1}, \quad k = 1, 2, 3, \dots, \quad (2.5.46)$$

mit den Koeffizienten

$$\beta_k = \frac{(xp_k, p_k)}{\|p_k\|^2}, \quad k = 0, 1, 2, \dots, \quad \gamma_k = \frac{\|p_k\|^2}{\|p_{k-1}\|^2}, \quad k = 1, 2, 3, \dots$$

Beweis: Das Gram-Schmidt-Verfahren erzeugt die Polynome p_k nach der Vorschrift

$$p_0 \equiv 1, \quad k = 1, 2, \dots: \quad p_k = x^k - \sum_{i=0}^{k-1} \frac{(x^k, p_i)}{\|p_i\|^2} p_i.$$

Also ist $p_0 \equiv 1$ und $p_1 = x - \beta_0$. Wir setzen $q_{k+1} := (x - \beta_k)p_k - \gamma_k p_{k-1}$. Dann gilt offenbar wegen $p_k \perp P_{k-1}$ und der Symmetrieeigenschaft von (\cdot, \cdot) :

$$\begin{aligned} (q_{k+1}, p_k) &= (xp_k, p_k) - \beta_k \|p_k\|^2 - \gamma_k \underbrace{(p_{k-1}, p_k)}_{=0} = 0, \\ (q_{k+1}, p_{k-1}) &= (xp_k, p_{k-1}) - \beta_k \underbrace{(p_k, p_{k-1})}_{=0} - \gamma_k \underbrace{\|p_{k-1}\|^2}_{= \|p_k\|^2} = (p_k, \underbrace{xp_{k-1} - p_k}_{\in P_{k-1}}) = 0, \end{aligned}$$

sowie für $j = 0, \dots, k-2$:

$$(q_{k+1}, p_j) = (p_k, \underbrace{xp_j}_{\in P_{k-1}}) - \beta_k \underbrace{(p_k, p_j)}_{=0} - \gamma_k \underbrace{(p_{k-1}, p_j)}_{=0} = 0.$$

Also ist q_{k+1} orthogonal zu $P_k = \text{Span}\{p_0, \dots, p_k\}$ und hat die Gestalt

$$q_{k+1}(x) = x^{k+1} + r(x), \quad r \in P_k.$$

Entwickelt man r nach $\{p_0, \dots, p_k\}$,

$$r(x) = \sum_{i=0}^k (r, p_i) \|p_i\|^{-2} p_i(x),$$

so ergibt sich mit

$$\begin{aligned} q_{k+1}(x) &= x^{k+1} + \sum_{i=0}^k \frac{(r, p_i)}{\|p_i\|^2} p_i(x) \\ &= x^{k+1} + \sum_{i=0}^k \underbrace{\{(q_{k+1}, p_i) - (x^{k+1}, p_i)\}}_{=0} \|p_i\|^{-2} p_i(x) = p_{k+1}(x) \end{aligned}$$

schließlich die Behauptung.

Q.E.D.

2.6 Tschebyscheff-Approximation

Im Folgenden betrachten wir nur reell-wertige Funktionen. Die Gauß-Approximation hat gewisse Probleme mit der gleichmäßigen Approximation auf dem ganzen zugrunde liegenden Intervall; sie ist aber verhältnismäßig einfach zu realisieren. Die sog. „Tschebyscheff-Approximation“ verwendet direkt die „Maximumnorm“

$$\|f\|_\infty = \max_{a \leq x \leq b} |f(x)|$$

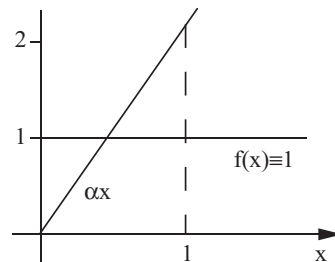
zur Bestimmung der besten Approximation $g \in S \subset C[a, b]$.

$$\|f - g\|_\infty = \min_{\varphi \in S} \|f - \varphi\|_\infty. \quad (2.6.47)$$

Die Norm $\|\cdot\|_\infty$ auf $C[a, b]$ wird nicht durch ein Skalarprodukt erzeugt; die Existenz der besten Approximation kann also nicht aus Satz 2.12 erschlossen werden. Tatsächlich ist i. Allg. die beste Approximation nicht einmal eindeutig bestimmt. Ihre Existenz ist jedoch allgemein in normierten Räumen gesichert.

Beispiel 2.9: Dieses Beispiel belegt die mögliche Mehrdeutigkeit der „besten“ Tschebyscheff-Approximation.

$$\begin{aligned} [a, b] &= [0, 1], \quad f(x) \equiv 1 \\ S &= \{g \mid g(x) = \alpha x, \alpha \in \mathbb{R}\}, \quad \dim S = 1 \\ \|f - g\|_\infty &\geq 1 \quad \forall g \in S \\ \|f - g\|_\infty &= 1 \quad \forall g = \alpha x, \quad 0 \leq \alpha \leq 2 \end{aligned}$$



Satz 2.18 (Allgemeine Tschebyscheff-Approximation): Sei E ein normierter Vektorraum mit Norm $\|\cdot\|$ und $S \subset E$ ein endlich dimensionaler Teilraum. Dann gibt es zu jedem $f \in E$ eine beste Approximation $g \in S$:

$$\|f - g\| = \min_{\varphi \in S} \|f - \varphi\|. \quad (2.6.48)$$

Beweis: Ein $g_0 \in S$ mit $\|g_0\| > 2\|f\|$ kann keine beste Approximation sein, da

$$\|f - g_0\| \geq \|g_0\| - \|f\| > \|f\| = \|f - 0\| \geq \inf_{\varphi \in S} \|f - \varphi\|.$$

Die optimale Approximation ist also in der beschränkten Teilmenge

$$S_0 := \{\varphi \in S : \|\varphi\| \leq 2\|f\|\} \subset S$$

zu suchen. Sie ist abgeschlossen und, da S endlich dimensional ist, kompakt (Satz von Bolzano/Weierstrass). Die auf S stetige Funktion $F(\varphi) := \|f - \varphi\|$ nimmt dann auf S_0 ein Minimum g an, d. h.:

$$\|f - g\| = \min_{\varphi \in S_0} \|f - \varphi\| = \min_{\varphi \in S} \|f - \varphi\|.$$

Q.E.D.

Die Eindeutigkeit der Tschebyscheff-Approximation in $C[a, b]$ wird durch die sog. „Haarsche²⁰ Bedingung“ (H) an den Ansatzraum $S \subset C[a, b]$ mit $\dim S = n$ garantiert:

Definition 2.8: (H) Man sagt, dass der (endlich dimensionale) Teilraum $S \subset C[a, b]$ der „Haarschen Bedingung“ genügt, wenn die Lagrange-Interpolationsaufgabe $g(x_i) = y_i$, $i = 1, \dots, n$ mit beliebigen Stützstellen $a \leq x_1 < x_2 < \dots < x_n \leq b$ und Werten $y_1, \dots, y_n \in \mathbb{R}$ stets durch ein $g \in S$ lösbar ist.

Für einen Teilraum $S \subset C[a, b]$ ist die Haarsche Bedingung äquivalent zur eindeutigen Lösbarkeit der Lagrange-Interpolationsaufgabe. Dies sieht man wie folgt:

Sei $\{g_1, \dots, g_n\}$ eine Basis von S . Die Existenz eines interpolierenden $g = \sum a_i g_i \in S$ ist äquivalent zur Lösbarkeit des linearen Gleichungssystems

$$\sum_{i=1}^n a_i g_i(x_j) = y_j, \quad j = 1, \dots, n, \quad (2.6.49)$$

für den Koeffizientenvektor $(a_1, \dots, a_n)^T$. Die Haarsche Bedingung ist äquivalent zur Regularität der Matrix $(g_i(x_j))_{i,j=1,\dots,n}$, d. h. zur eindeutigen Lösbarkeit der Interpolationsaufgabe.

²⁰Alfréd Haar (1885–1933): Ungarischer Mathematiker; Professor in Kolozsvár (Cluj), Budapest und Szeged; viele wichtige Beiträge zur Approximationstheorie („Haarsche Bedingung“) und Analysis auf Gruppen („Haar measure“).

Beispiel 2.10: Wir geben Beispiele von Systemen S , für welche die Haarsche Bedingung (H) erfüllt ist.

1. Die Polynomräume P_n erfüllen Bedingung (H) auf jedem Intervall $[a, b]$.
2. Der Raum $S = \text{Span}\{x, \dots, x^n\}$ erfüllt Bedingung (H) nur, wenn $0 \notin [a, b]$ ist:

$$x_1 = 0 \Rightarrow x_1^k = 0, \quad k = 1, \dots, n \quad \Rightarrow \quad \det(x_i^j)_{i,j=1, \dots, n} = 0.$$

Satz 2.19 (Tschebyscheffscher Alternantensatz): Für den Teilraum $S \subset C[a, b]$ mit $\dim S = n$ sei die Haarsche Bedingung (H) erfüllt. Dann ist die Tschebyscheff-Approximation $g \in S$, einer Funktion $f \in C[a, b]$ durch folgende Eigenschaft charakterisiert:

(A) Es existieren $m \geq n + 1$ Stellen $a \leq x_0 < \dots < x_m \leq b$ (sog. „Alternante“, so dass für die Fehlerfunktion $e(x) = f(x) - g(x)$ gilt:

$$|e(x_i)| = \|e\|_\infty, \quad e(x_i) = -e(x_{i+1}); \quad i = 1, \dots, n. \quad (2.6.50)$$

Insbesondere ist die beste Approximation eindeutig bestimmt.

Beweis: Für den nicht trivialen Beweis der Alternantenaussage verweisen wir auf die Literatur; z. B.: Schaback/Werner [5] Q.E.D.

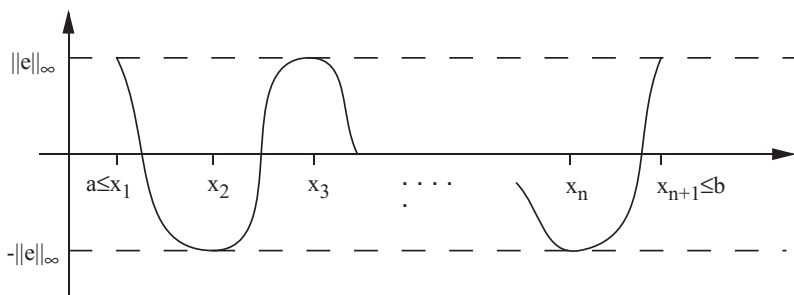


Abbildung 2.9: Schema der Alternantenregel

Korollar 2.2: Der Alternantensatzes impliziert, dass die beste Approximation für den Spezialfall $S = P_{n-1}$ eindeutig bestimmt ist.

Beweis: Seien g_1, g_2 zwei beste Approximationen mit $e_1 = f - g_1$, $e_2 = f - g_2$. Für $\lambda \in (0, 1)$ ist dann $\|\lambda e_2\|_\infty < \|e_1\|_\infty$, so dass der Graph von $\lambda e_2(x)$ den von $e_1(x)$ mindestens n -mal schneidet (siehe Abbildung 2.9). Jede der Funktionen $\varphi_\lambda(x) = e_1(x) - \lambda e_2(x)$ hat also mindestens n Nullstellen. Durch Grenzübergang $\lambda \rightarrow 1$ folgt, dass $\varphi_1(x) = e_1(x) - e_2(x) = g_2(x) - g_1(x)$ mindestens n (ihrer Vielfachheit entsprechend oft gezählte) Nullstellen besitzt. Wegen $g_2 - g_1 \in P_{n-1}$ ergibt sich zwangsläufig $g_1 \equiv g_2$. Q.E.D.

Bemerkung 2.4: Der Alternantensatz ist die Grundlage des sog. „Remez²¹-Algorithmus“ zur Konstruktion der Tschebyscheff-Approximation. Wäre eine Alternante $\{x_1, \dots, x_{n+1}\}$ bekannt, so könnte man bei gegebener Basis $\{\varphi_1, \dots, \varphi_n\}$ von S die beste Approximation

$$g = \sum_{i=1}^n \alpha_i \varphi_i$$

sowie die Größe $\alpha_{n+1} := \sigma \|f - g\|_\infty$, $\sigma \in \{-1, 1\}$, aus dem linearen Gleichungssystem

$$\sum_{i=1}^n \alpha_i \varphi_i(x_k) + (-1)^k \alpha_{n+1} = f(x_k), \quad k = 1, \dots, n+1,$$

berechnen. Der Remez-Algorithmus besteht aus der systematischen iterativen Suche nach der Alternante $\{x_1, \dots, x_{n+1}\}$. In jedem Schritt wird mit der Näherung $\{x_i^{(t)}, \dots, x_{n+1}^{(t)}\}$ das Gleichungssystem für $\alpha_1^{(t)}, \dots, \alpha_n^{(t)}$ und $\alpha_{n+1}^{(t)}$ gelöst und für

$$g(t) = \sum_{i=1}^n \alpha_i^{(t)} \varphi_i$$

das Optimalitätskriterium

$$\|f - g^{(t)}\|_\infty = |\alpha_{n+1}^{(t)}|$$

abgefragt. I. Allg. konvergiert $\{x_1^{(t)}, \dots, x_{n+1}^{(t)}\}$ gegen $\{x_1, \dots, x_{n+1}\}$, allerdings nicht in endlich vielen Schritten.

Beispiel 2.11: $f(x) = \cos(x)$, $[a, b] = [0, \pi/2]$, $S = P_1$

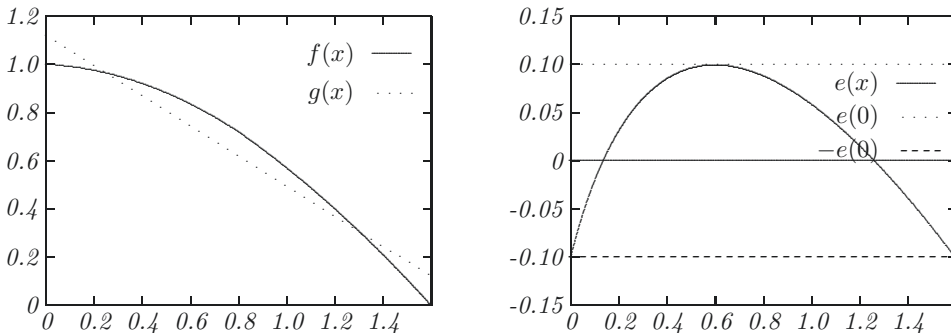


Abbildung 2.10: Anwendung der Alternantenregel

²¹Evgeny Yakovlevich Remez (1896–1975): Russischer Mathematiker; Professor an der Universität Kiew (1935); Beiträge zur konstruktiven Approximationstheorie („Remez-Algorithmus“) und zur numerischen Lösung von Differentialgleichungen.

2.6.1 „Optimale“ Lagrange-Interpolation

Zur Anwendung der Tschebyscheff-Approximation stellen wir die Frage nach der „optimalen“ Wahl der Stützstellen bei der Lagrange-Interpolation. Für das Lagrange-Interpolationspolynom $p \in P_n$ einer Funktion $f \in C^{n+1}[a, b]$ zu den Stützstellen $a \leq x_0 < x_1 < \dots < x_n \leq b$ gilt nach Satz 2.3 die Fehlerdarstellung

$$f(x) - p_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\zeta_x) L(x), \quad x \in [x_0, \dots, x_n],$$

wobei $L(x) = \prod_{j=0}^n (x - x_j)$. Die Stützstellen x_0, \dots, x_n solle so bestimmt werden, dass

$$\max_{a \leq x \leq b} |L(x)| = \|L\|_\infty \quad (2.6.51)$$

minimal wird. Damit hätte man eine „optimale“ Darstellbarkeit von Funktionen aus $C^{n+1}[a, b]$ durch Lagrange-Interpolationspolynome in P_n . Nun ist $L(x) = x^{n+1} - p$ mit einem $p \in P_n$, d. h.: Die Aufgabe „optimale“ Stützstellen zu bestimmen, ist äquivalent zur Konstruktion der Tschebyscheff-Approximation zu $f(x) = x^{n+1}$ bzgl. $S = P_n$. Nach dem Alternantensatz hat die Fehlerfunktion $e = x^{n+1} - p$ mindestens $n + 1$ Nullstellen im Intervall $[a, b]$.

Satz 2.20 (Optimale Stützstellen): Auf dem Intervall $[a, b] = [-1, 1]$ ist die Tschebyscheff-Approximation $g \in P_n$ zu $f(x) = x^{n+1}$ gegeben durch

$$g(x) = x^{n+1} - 2^{-n} T_{n+1}(x) \quad (2.6.52)$$

mit dem $(n + 1)$ -ten Tschebyscheff-Polynom

$$T_{n+1}(x) = \cos[(n+1) \arccos(x)].$$

Die Nullstellen

$$x_k = \cos\left(\frac{\pi}{2} \frac{2k+1}{n+1}\right), \quad k = 0, \dots, n \quad (2.6.53)$$

von T_{n+1} sind gerade die „optimalen“ Stützstellen der Lagrange-Interpolation auf $[-1, 1]$.

Beweis: Das Polynom $T_{n+1} \in P_{n+1}$ hat die $n+1$ Nullstellen $x_k = \cos\left(\frac{\pi}{2} \frac{2k+1}{n+1}\right)$, $k = 0, \dots, n$, und es gilt (Übungsaufgabe)

$$\max_{-1 \leq x \leq 1} \prod_{k=0}^n |x - x_k| = 2^{-n}.$$

Der rekursiven Beziehung

$$T_0 \equiv 1, \quad T_1(x) = x, \quad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x),$$

entnimmt man, dass $T_{n+1}(x) = 2^n x^{n+1} + q(x)$ mit einem $q \in P_n$. Also ist

$$2^{-n} T_{n+1}(x) = \prod_{k=0}^n (x - x_k) = L(x).$$

Weiter nimmt $T_{n+1}(x) = \cos((n+1) \arccos(x))$ im Intervall $[-1, 1]$ offenbar genau $(n+2)$ -mal einen Extremwert an, abwechselnd ± 1 (siehe Abb. 2.11). Diese $n+2$ Extremalstellen bilden dann eine Alternante für die Approximation $g(x) = x^{n+1} - 2^{-n} T_{n+1}(x) \in P_n$ zu x^{n+1} . Nach dem Alternantensatz ist g also die eindeutig beste Approximation zu x^{n+1} . Mit den Nullstellen x_k von $T_{n+1}(x)$ gilt folglich

$$\begin{aligned} \max_{-1 \leq x \leq 1} \left| \prod_{k=0}^n (x - x_k) \right| &= 2^{-n} \max_{-1 \leq x \leq 1} |T_{n+1}(x)| \\ &= \max_{-1 \leq x \leq 1} |x^{n+1} - [x^{n+1} - 2^{-n} T_{n+1}(x)]| \\ &= \min_{p \in P_n} \max_{-1 \leq x \leq 1} |x^{n+1} - p(x)| \\ &= \min_{-1 \leq \zeta_0 < \dots < \zeta_n \leq 1} \max_{-1 \leq x \leq 1} \left| \prod_{k=0}^n (x - \zeta_k) \right| \end{aligned}$$

und damit die behauptete Optimalitätseigenschaft.

Q.E.D.

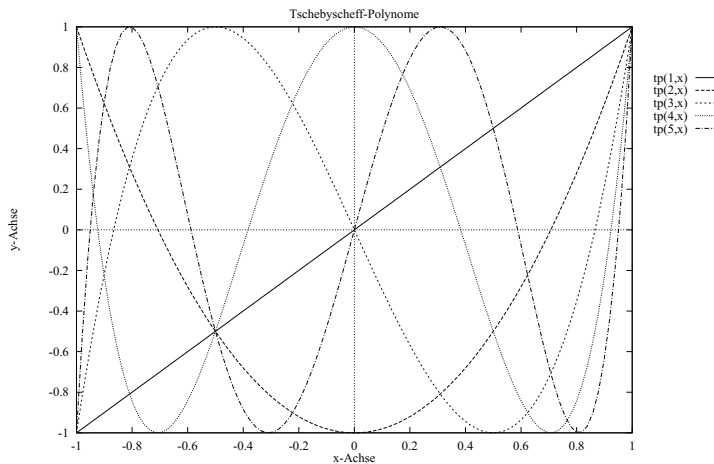


Abbildung 2.11: Anwendung der Alternantenregel

Die optimalen Stützstellen der Lagrange-Interpolation auf einem beliebigen Intervall $[a, b]$ gewinnt man aus dem Resultat von Satz 2.20 mit Hilfe der Transformation φ :

$[a, b] \rightarrow [-1, 1]$:

$$x = \varphi(y) := \frac{2}{b-a}y - \frac{a+b}{b-a},$$

zu

$$y_k = \frac{b-a}{2}x_k + \frac{a+b}{2}, \quad k = 0, \dots, n.$$

Für den Interpolationsfehler einer Funktion $f \in C^{n+1}[a, b]$ gilt dann

$$\|f - p_n\|_\infty \leq \frac{1}{(n+1)!} \frac{(b-a)^{n+1}}{2^{2n+1}} \|f^{(n+1)}\|_\infty. \quad (2.6.54)$$

2.7 Übungsaufgaben

Übung 2.1: Gegeben seien $n+1$ paarweise verschiedene Punkte $x_i \in \mathbb{R}^1, i = 0, 1, \dots, n$, und die zugehörigen $n+1$ sog. Lagrange-Polynome

$$L_i^{(n)}(x) = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j}, \quad i = 0, \dots, n.$$

Man zeige, dass die Polynome $\{L_i^{(n)}, i = 0, \dots, n\}$, eine Basis des Polynomraums P_n (Vektorraum aller Polynome vom Grad kleiner oder gleich n) bilden, und dass die folgenden Beziehungen gelten:

$$\begin{aligned} i) \quad & \sum_{i=0}^n L_i^{(n)}(x) = 1, \quad x \in \mathbb{R}^1, \\ ii) \quad & \sum_{i=0}^n x_i^k L_i^{(n)}(0) = 0, \quad k = 1, \dots, n, \\ iii) \quad & \sum_{i=0}^n x_i^{n+1} L_i^{(n)}(0) = (-1)^n \prod_{i=0}^n x_i. \end{aligned}$$

(Hinweis: Man verwende die Eindeutigkeit des Lagrange-Interpolationspolynoms und die Darstellung des Fehlers bei der Lagrange-Interpolation.)

Übung 2.2: Für verschiedene Orte wurde an einem bestimmten Tag die folgenden Tageslängen gemessen:

Ort	Tageslänge	Lage
A	17h 28m	55,7°
B	18h 00m	57,7°
C	18h 31m	59,3°
D	19h 56m	62,6°
E	22h 34m	65,6°

Man bestimme die Tageslänge am Ort F bei $61,7^\circ$ durch Auswertung des zugehörigen Interpolationspolynoms mit Hilfe des Neville-Algorithmus. (Es genügt 4-stellige Dezimalrechnung.)

Übung 2.3: Es soll eine 10-stellige Wertetabelle von

$$f(x) = \int_0^x \sin(t)^2 dt, \quad x \in [0, \pi],$$

erstellt werden (in Festkommadarstellung), so dass kubische Lagrange-Interpolation einen Fehler kleiner als $5 \cdot 10^{-9}$ für jeden Wert von x im Intervall $[0, \pi]$ ergibt. Reichen dazu die Werte zu 250 äquidistant verteilten Stützstellen aus? (Hinweis: Der Auswertungsfehler setzt sich zusammen aus dem absoluten Interpolationsfehler und dem absoluten Rundungsfehler in den zur Interpolation verwendeten Stützwerten.)

Übung 2.4: Gegeben sei die Funktion $f(x) = e^{\lambda x}$, $\lambda \in \mathbb{R}$, auf einem Intervall $[a, b]$. Man zeige, dass in diesem Fall der Fehler $f - p_n$ der Lagrange-Interpolation von f über beliebig verteilten $n + 1$ (paarweise verschiedenen) Stützstellen aus $[a, b]$ für $n \rightarrow \infty$ gleichmäßig gegen Null konvergiert:

$$\max_{x \in [a, b]} |f(x) - p_n(x)| \rightarrow 0 \quad (n \rightarrow \infty).$$

Was unterscheidet diese Funktion von dem in der Vorlesung angegebenen Beispiel $f(x) = (1 + x^2)^{-1}$, für welches die Lagrange-Interpolation für $n \rightarrow \infty$ nicht konvergiert (s. Aufgabe 3.5)?

Übung 2.5: Wir betrachten die Hermite-Interpolationsaufgabe, zu (paarweise verschiedenen) Stützstellen x_i ($i = 0, \dots, m$) und zu gegebenen Werten $y_i^{(0)}, y_i^{(1)}$ ($i = 0, \dots, m$) ein Polynom $p \in P_n$, $n = 2m + 1$, so zu bestimmen, dass

$$p(x_i) = y_i^{(0)}, \quad p'(x_i) = y_i^{(1)} \quad (i = 0, \dots, m).$$

Man zeige:

- Die Hermite-Interpolationsaufgabe besitzt eine eindeutig bestimmte Lösung.
- Im Falle der Interpolation einer $(n+1)$ -mal differenzierbaren Funktion f , d. h. $p(x_i) = f(x_i)$, $p'(x_i) = f'(x_i)$ ($i = 0, \dots, m$) gilt die Fehlerdarstellung

$$f(x) - p(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!} \prod_{i=0}^m (x - x_i)^2$$

mit x -abhängigen Zwischenstellen $\xi_x \in [a, b]$. (Hinweis: Man modifiziere die entsprechende Argumentation der Vorlesung oder des Skriptums für die Lagrange-Interpolation.)

Übung 2.6 (Praktische Aufgabe): Man berechne die Lagrange-Interpolationspolynome der Funktionen

$$f(x) = \frac{1}{1 + 25x^2}, \quad g(x) = \sqrt{|x|}, \quad -1 \leq x \leq 1,$$

in Neville-Darstellung, jeweils zu den Stützstellen $x_i = -1 + ih$, $i = 0, \dots, n$, mit $h = 2/n$, für $n = 5, 10, 15, 20$. Man stelle die Polynome grafisch dar und vergleiche die Polynomgraphen mit den richtigen Funktionsverläufen.

Übung 2.7: Auf einer Zerlegung $0 = x_0 < x_1 < \dots < x_{N-1} < x_N = 1$ des Intervalls $I = [0, 1]$ werde eine Funktion stückweise quintisch interpoliert, d. h.: Auf jedem der Teilintervalle $I_k = [x_{k-1}, x_k]$, $k = 1, \dots, N$, wird f durch ein Polynom $p_5^{(k)} \in P_5$ interpoliert, wobei jeweils einer der folgenden Sätze an Bedingungen verwendet wird:

- (i) $p_5^{(k)}(\xi) = f(\xi)$, $\xi \in \{x_{k-1} + jh_k/5, j = 0, \dots, 5\}$, $h_k = x_k - x_{k-1}$,
- (ii) $p_5^{(k)}(\xi) = f(\xi)$, $(p_5^{(k)})'(\xi) = f'(\xi)$, $(p_5^{(k)})''(\xi) = f''(\xi)$, $\xi \in \{x_{k-1}, x_k\}$.

Im ersten Fall (Lagrange-Interpolation) ist die resultierende zusammengesetzte Funktion auf dem Intervall I stetig, und im zweiten Fall (Hermite-Interpolation) zweimal stetig differenzierbar. Man zeige, dass für $f \in C^6[0, 1]$ in beiden Fällen die Abschätzung gilt:

$$\max_{x \in I} |f(x) - p(x)| \leq \frac{h^6}{720} \max_{x \in I} |f^{(6)}(x)|, \quad h := \max_{k=1, \dots, N} h_k.$$

(Hinweis: Man wende die Fehlerdarstellung der Vorlesung für die Lagrange- bzw. die Hermite-Interpolation auf jedem der Teilintervalle I_k an.)

Übung 2.8: Für die Funktion $f(x) = \cosh(x)$ ist die Wertetabelle gegeben

x	$f(x)$
0.52	1,1382741
0.56	1,1609408
0.60	1,1854652
0.64	1,2118867
0.68	1,2402474.

Man bestimme durch Extrapolation eines geeigneten Differenzenquotienten möglichst gute Näherungen zum Ableitungswert $f'(0.6) = 0,63665358\dots$.

Übung 2.9: Welche von den Indexfolgen

- (i) $n_i = 2i - 1$, $i \in \mathbb{N}$,
- (ii) $n_i = 3^i$, $i \in \mathbb{N}$,
- (iii) $n_i = i^2$, $i \in \mathbb{N}$,

für Schrittweiten $h_i = h/n_i$ ist zulässig für die Extrapolation zum Limes?

Übung 2.10: Es bezeichne S_0 den Vektorraum der kubischen, natürlichen Spline-Funktionen zu den Stützstellen $x_0 = 0, x_1 = 1, x_2 = 2$.

a) Sind die folgenden Funktionen in S_0 ?

$$\begin{aligned} (i) \quad & f(x) = x^3 - x^2, \\ (ii) \quad & f(x) = x^2(x - 6) - (x - 2)^3, \\ (iii) \quad & f(x) = \max\{0, x - 1\}^3 - \frac{1}{2}x^3. \end{aligned}$$

b) Man bestimme den interpolierenden Spline $s_2 \in S_0$ für $f(x) = x^3$. Wie lautet das Ergebnis, wenn die natürlichen Randbedingungen durch $s_2''(x_0) = f''(x_0)$, und $s_2''(x_2) = f''(x_2)$ ersetzt werden.

Übung 2.11 (Praktische Aufgabe): Zur Bestimmung des halben Umfangs

$$\pi = 3,1415\,92653\,58979\,32384\,62643\dots$$

des Einheitskreises verwendeten schon die „alten Griechen“ den Umfang einbeschriebener regulärer Polygone. Mit Hilfe des 96-seitigen Polygons fand z. B. Archimedes den Wert $\pi \approx 3,142$. Die allgemeine Formel für den Umfang T_n des n -seitigen einbeschriebenen, regulären Polygons ist

$$(*) \quad T_n = 2n \sin(\pi/n) \quad \rightarrow \quad 2\pi \quad (n \rightarrow \infty).$$

Man kann die T_n für $n = 6 \cdot 2^i$ mit Hilfe der Rekursionsformel (nachprüfen!)

$$T_6 = 6, \quad T_{2n} = 2\sqrt{2n^2 - n\sqrt{4n^2 - T_n^2}},$$

ohne Auswertung des Sinus berechnen.

i) Man bestimme mit Hilfe der Richardson-Extrapolation aus den Stützwerten

$$\{T_n, n = 6 \cdot 2^i, i = 0, \dots, k\},$$

für $k = 1, \dots, 30$, Approximationen für π und plote den resultierenden Fehler in Abhängigkeit von k . (Hinweis: Man setze $x_n = 1/n$ und $T(x_n) := T_n$ und extrapoliere die Funktion $T(x) := 2/x \sin(\pi x)$ nach $x = 0$.)

ii) Man wiederhole die Rechnung mit den Stützwerten $\{T_i, i = 3, 4, \dots, k\}$ für $k = 4, \dots, 33$, wobei die benötigten Werte T_i direkt aus der Definitionsformel (*) bestimmt werden sollen. Wie sind die beobachteten Phänomene zu erklären?

Übung 2.12: Man zeige, dass die Funktionen

$$\varphi_0(x) = 1/\sqrt{2x}, \quad \varphi_k(x) = \frac{1}{\sqrt{\pi}} \cos(kx), \quad \psi_k(x) = \frac{1}{\sqrt{\pi}} \sin(kx), \quad k = 1, \dots, n,$$

ein Orthonormalsystem des Teilraums $T_n \subset C[-\pi, \pi]$ der trigonometrischen Polynome vom Grad kleiner oder gleich n bzgl. des L^2 -Skalarprodukts über dem Intervall $[-\pi, \pi]$ bilden und bestimme die beste Approximation der Funktion $f(x) = x$ in T_n .

Übung 2.13: Man zeige, dass die durch

$$\varphi_k(x) = \frac{k!}{(2k)!} \frac{d^k}{dx^k} (x^2 - 1)^k, \quad k = 0, 1, \dots, n,$$

definierten Polynome orthogonal bzgl. des L^2 -Skalarprodukts über $[-1, 1]$ sind, und dass

$$\|\varphi_k\| = \frac{k!^2}{(2k)!} \sqrt{\frac{2^{2k+1}}{2k+1}}, \quad \varphi_k(1) = \frac{k!^2}{(2k)!} 2^k.$$

Durch Normierung erhält man hieraus die sog. Gauß-Legendre-Polynome

$$L_k(x) := \frac{(2k)!}{k!^2 2^k} \varphi_k(x), \quad L(1) = 1.$$

(Hinweis: Man verwende partielle Integration, leite für die Integrale $I_k := \int_{-1}^1 (1-x^2)^k dx$ eine einstufige Rekursionsformel her und differenziere die Funktion $(x^2 - 1)^k$.)

Übung 2.14: Man zeige die folgenden Eigenschaften einer Norm $\|\cdot\|$ auf einem Vektorraum E (nicht notwendig endlich dimensional):

a) $\|x - y\| \geq \left| \|x\| - \|y\| \right|, \quad x, y \in E.$

b) Die Funktion $N(\cdot) = \|\cdot\| : E \rightarrow \mathbb{R}$ ist stetig.

c) Ist E endlich dimensional, so sind alle Normen auf E äquivalent, d. h.: Zu je zwei Normen $\|\cdot\|_1, \|\cdot\|_2$ auf E gibt es stets Konstanten $M \geq m > 0$, so dass

$$m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1, \quad x \in E.$$

(Hinweis: Für eine Basis $\{e^1, \dots, e^n\}$ von E betrachte man die Funktion $F(\alpha_1, \dots, \alpha_n) := \left\| \sum_{i=1}^n \alpha_i e^i \right\|$ auf der Einheitskugel des \mathbb{R}^n .)

Übung 2.15: Man bestimme die Gauß-Approximationen der Funktion $f(x) = \sqrt{x}$ bzgl. der L^2 -Norm über dem Intervall $[0, 1]$ in den Polynomräumen P_0 , P_1 und P_2 . Man stelle die Ergebnisse graphisch dar.

Übung 2.16: (Praktische Aufgabe) Man berechne rekursiv die Gauß-Legendre- und die Tschebyscheff-Polynome L_k und T_k auf dem Intervall $[-1, 1]$ aus den Beziehungen

$$\begin{aligned} \text{a) } \quad L_k(x) &:= \frac{(2k)!}{2^k k!^2} \varphi_k(x), \quad k = 0, 1, 2, \dots, \\ \varphi_0(x) &= 1, \quad \varphi_1(x) = x, \quad \varphi_{k+1}(x) = x\varphi_k(x) - \frac{k^2}{4k^2 - 1} \varphi_{k-1}, \end{aligned}$$

$$\text{b) } \quad T_0(x) = 1, \quad T_1(x) = x, \quad T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x),$$

für $k = 0, 1, 2, \dots, 10$ und stelle sie grafisch dar.