

Approximationstheorie

Vorlesungsskriptum WS 2020/21

R. Verfürth

Fakultät für Mathematik, Ruhr-Universität Bochum

Inhaltsverzeichnis

Kapitel I. Einführung	5
I.1. Beispiele für Approximationsaufgaben	5
I.2. Allgemeine Existenzaussagen	7
I.3. Allgemeine Eindeutigkeitsaussagen	11
Kapitel II. Lineare Čebyšev-Approximation	19
II.1. Kolmogoroff-Kriterium	19
II.2. Charakterisierungssatz	23
II.3. Haarscher Eindeutigkeitssatz	28
II.4. Strenge Eindeutigkeit	33
II.5. Satz von Mairhuber-Curtis	36
Kapitel III. Čebyšev-Approximation durch Polynome	39
III.1. Čebyšev-Polynome	39
III.2. Approximation von $(x - a)^{-1}$	42
III.3. Vergleichssatz von Bernstein	46
Kapitel IV. Asymptotische Aussagen	51
IV.1. Satz von Weierstraß und Lethargietheorem	51
IV.2. Glattheitsmaße	56
IV.3. Jackson-Sätze	58
IV.4. Ungleichungen von Bernstein und Markov	66
IV.5. Umkehrsätze von Bernstein und Zygmund	71
IV.6. Resultate negativen Charakters	76
Kapitel V. Numerische Behandlung der Čebyšev-Approximation	87
V.1. Satz von Stiefel	87
V.2. Der Remez-Algorithmus	92
V.3. Diskretisierungstheorie	95
Kapitel VI. Nichtlineare Approximation	99
VI.1. Einführung	99
VI.2. Approximation im Hilbert-Raum	101
VI.3. Lineare Approximation durch stückweise konstante Funktionen	108
VI.4. Nichtlineare Approximation durch stückweise konstante Funktionen	114
VI.5. Adaptive Approximation	119
VI.6. Haar-Wavelets	122

VI.7. Haar-Wavelets im \mathbb{R}^d und glatte Wavelets	126
VI.8. Anwendungsbeispiele	131
Index	133
Literaturverzeichnis	137

KAPITEL I

Einführung

I.1. Beispiele für Approximationsaufgaben

BEISPIEL I.1.1 (Funktionsanpassung). Auf einem Computer müssen transzendente Funktionen, wie z. B. die Exponentialfunktion, durch Algorithmen approximiert werden, die nur die vier Grundrechenarten benötigen. Wegen der Additionstheoreme kann man sich bei der Exponentialfunktion auf Argumente x im Intervall $[0, 1]$ beschränken. Der naheliegendste Ansatz ist das Taylor-Polynom

$$e^x \approx \sum_{k=0}^n \frac{1}{k!} x^k.$$

Der Fehler ist

$$\max_{0 \leq x \leq 1} \left| e^x - \sum_{k=0}^n \frac{1}{k!} x^k \right| = \frac{e}{(n+1)!}.$$

Stattdessen kann man auch die beste Approximation durch Polynome suchen:

$$\max_{0 \leq x \leq 1} \left| e^x - \sum_{k=0}^n a_k x^k \right| \rightarrow \min_{a_k}.$$

Dies ist ein *Approximationsproblem*. Man kann zeigen, dass für den Fehler der besten Approximation gilt

$$\min_{a_0, \dots, a_n} \max_{0 \leq x \leq 1} \left| e^x - \sum_{k=0}^n a_k x^k \right| = \frac{e}{2^{2n+1} (n+1)!}.$$

Man gewinnt also einen Faktor 2^{2n+1} . Einen noch größeren Genauigkeitsgewinn kann man erzielen, wenn man die Division einbezieht und Approximationen der Form $\frac{p(x)}{q(x)}$ mit Polynomen p und q betrachtet.

BEISPIEL I.1.2 (Angleichung von Messdaten). Zu Parametern t_1, \dots, t_n sind Daten x_1, \dots, x_n gemessen worden. Man vermutet einen linearen Zusammenhang zwischen Parametern und Daten. Dies führt auf das Approximationsproblem

$$\sum_{i=1}^n (x_i - a - bt_i)^2 \rightarrow \min_{a,b}$$

das Problem der *kleinsten Fehlerquadrate*. Wie man sich leicht überlegt, sind die optimalen Werte für a und b Lösungen eines zweidimensionalen linearen Gleichungssystems.

BEISPIEL I.1.3 (Lösen einer parabolischen Differentialgleichung). In einem homogenen Stab, dessen Länge der Einfachheit halber auf π normiert sei, ist die Temperaturverteilung in Abhängigkeit von Ort und Zeit zu bestimmen. Die Stabenden werden stets auf Temperatur 0 gehalten; die Anfangstemperatur des Stabes ist bekannt. Dies führt auf die parabolische Differentialgleichung (*Wärmeleitungsgleichung*)

$$\frac{\partial}{\partial t}u(x, t) - \frac{\partial^2}{\partial x^2}u(x, t) = 0 \quad \text{in } \Omega = \{(x, t) : 0 < x < \pi, t < 0\}.$$

mit der *Randbedingung*

$$u(0, t) = u(\pi, t) = 0 \quad \text{für } t > 0$$

und der *Anfangsbedingung*

$$u(x, 0) = f(x) \quad \text{für } 0 < x < \pi.$$

Wie man leicht einsieht, sind alle Funktionen

$$e^{-k^2 t} \sin(kx), \quad k \in \mathbb{N}^*,$$

Lösungen der Differentialgleichung und der Randbedingung. Daher gilt gleiches für jede Linearkombination

$$u_n(x, t) = \sum_{k=1}^n a_k e^{-k^2 t} \sin(kx), \quad n \in \mathbb{N}^*.$$

Die Anfangsbedingung führt auf die Forderung

$$u_n(x, 0) = \sum_{k=1}^n a_k \sin(kx) \stackrel{!}{=} f(x).$$

Man kann zeigen, dass wegen des Temperatúrausgleiches der Fehler zur Anfangszeit maximal ist, d.h.

$$\max_{(x,t) \in \Omega} |u(x, t) - u_n(x, t)| = \max_{0 \leq x \leq \pi} \left| f(x) - \sum_{k=1}^n a_k \sin(kx) \right|.$$

Dies führt auf das Approximationsproblem

$$\max_{0 \leq x \leq \pi} \left| f(x) - \sum_{k=1}^n a_k \sin(kx) \right| \rightarrow \min_{a_k}.$$

Wir werden zeigen, dass der Fehler der besten Approximation sich asymptotisch wie n^{-p} verhält, wenn f p -mal stetig differenzierbar ist. Andererseits werden wir auch zeigen, dass die Umkehrung gilt, so dass die p -malige stetige Differenzierbarkeit der Lösung u der Differentialgleichung folgt. Dieser Regularitätssatz folgt nicht direkt aus der Differentialgleichung.

Unsere drei Beispiele hatten folgende gemeinsame Struktur: Gegeben sind ein normierter Raum $(X, \|\cdot\|)$, eine nicht leere Teilmenge $V \subset X$ und ein Element $f \in X$. Gesucht ist eine *beste Approximation* $u^* \in V$ an f , so dass gilt

$$\|f - u^*\| = \inf_{u \in V} \|f - u\|.$$

In Beispiel I.1.1 ist $X = C([0, 1])$, $\|\cdot\| = \|\cdot\|_\infty$, $f = e^x$, $V = \mathbb{P}_n = \text{span}\{1, x, \dots, x^n\}$. In Beispiel I.1.2 ist $X = \mathbb{R}^n$, $\|\cdot\| = |\cdot|$, $V = \mathbb{R}^2$. In Beispiel I.1.3 ist $X = C_{2\pi} = \{f \in C([0, 2\pi]) : f \text{ ist } 2\pi\text{-periodisch}\}$, $\|\cdot\| = \|\cdot\|_\infty$, $V = \mathbb{T}_n^u = \text{span}\{\sin x, \sin(2x), \dots, \sin(nx)\}$.

Die grundlegenden Fragen sind:

- Gibt es eine beste Approximation?
- Ist die beste Approximation eindeutig?
- Wie groß ist der Fehler der besten Approximation und lässt er sich a priori, d.h. ohne Bestimmung der besten Approximation, abschätzen?
- Wie kann man die beste Approximation bestimmen?
- Gibt es eventuell eine einfacher berechenbare gute Approximation, die den gleichen asymptotischen Fehler hat wie die beste Approximation?

Wir werden im Folgenden sehen, dass man an diese Fragen unterschiedlich herangeht, indem man z. B. nur die abstrakten Eigenschaften der Räume X und V heranzieht oder konkret spezielle Eigenschaften des zu approximierenden Elementes f und der Menge V ausnutzt. Ebenso werden wir sehen, dass die verwendeten Methoden wesentlich von der gewählten Norm abhängen.

I.2. Allgemeine Existenzaussagen

Im Folgenden bezeichnet $(X, \|\cdot\|)$ stets einen normierten Vektorraum über \mathbb{R} oder \mathbb{C} . V ist stets eine nicht leere Teilmenge von X und f ist ein gegebenes Element von X . Man beachte, dass X nicht notwendig vollständig sein muss und dass V kein Unterraum sein muss.

DEFINITION I.2.1. Ein Element $u^* \in V$ heißt *beste Approximation*, kurz *b.A.*, an $f \in X$ in V , wenn gilt

$$\|f - u^*\| = \inf_{u \in V} \|f - u\| = \rho_V(f).$$

Die Größe $\rho_V(f)$ ist der *Abstand* zwischen dem Element f und der Menge V .

Die wichtigsten Beispiele für die Räume X sind der Raum $C(T)$ der stetigen Funktionen auf einer kompakten Menge T versehen mit der Supremum-Norm

$$\|f\|_\infty = \sup_{t \in T} |f(t)|$$

und die Räume $L^p(I)$, $1 \leq p < \infty$, der zur p -ten Potenz Lebesgue integrierbaren messbaren Funktionen auf einem Intervall I versehen mit der Norm

$$\|f\|_p = \left\{ \int_I |f(t)|^p dt \right\}^{\frac{1}{p}}.$$

In Definition 1.2.1 wird bewusst *nicht* gefordert, dass X ein Funktionenraum ist. Im Gegenteil: X darf auch endlich dimensional sein, und der \mathbb{R}^n versehen mit der Norm

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

bzw.

$$\|x\|_p = \left\{ \sum_{i=1}^n |x_i|^p \right\}^{\frac{1}{p}}$$

ist häufig ein guter Kandidat, um sich allgemeine Sätze geometrisch zu veranschaulichen.

BEISPIEL I.2.2. Man überlege sich Beispiele für Elemente $f \in \mathbb{R}^2$ und Mengen $V \subset \mathbb{R}^2$ mit zwei, drei, unendlich vielen bzw. keiner b.A..

DEFINITION I.2.3. Eine Menge V heißt *Existenzmenge* bzw. *Eindeutigkeitsmenge* bzw. *Čebyšev-Menge*, wenn es zu jedem $f \in X$ mindestens eine bzw. höchstens eine bzw. genau eine b.A. an f in V gibt.

Sei $f \in X$ beliebig, aber fest. Definiere die Abstandsfunktion $\Phi : V \rightarrow \mathbb{R}$ durch

$$\Phi(u) = \|f - u\|.$$

Aus der Dreiecksungleichung folgt

$$|\Phi(u_1) - \Phi(u_2)| = | \|f - u_1\| - \|f - u_2\| | \leq \|u_1 - u_2\|,$$

d.h., Φ ist Lipschitz-stetig. Falls V kompakt ist, folgt daher sofort, dass V eine Existenzmenge ist.

Leider ist in den meisten Anwendungen V nicht kompakt. Jedoch kann man dieses Problem durch einen Kunstgriff umgehen.

SATZ I.2.4 (Existenzsatz). *Jeder endlich dimensionale Unterraum eines normierten Raumes ist Existenzmenge.*

BEWEIS. Die Teilmenge

$$V_1 = \{u \in V : \|u\| \leq 2\|f\|\}$$

ist als beschränkte und abgeschlossene Teilmenge eines endlich dimensionalen Raumes kompakt. Daher nimmt die Abstandsfunktion Φ auf der Menge V_1 ihr Minimum bei einem Element $u^* \in V_1$ an. Insbesondere ist

$$\Phi(u^*) \leq \Phi(0) = \|f\|.$$

Für jedes $u \in V \setminus V_1$ folgt hieraus mit der Dreiecksungleichung

$$\begin{aligned}\Phi(u) &= \|f - u\| \geq \|u\| - \|f\| \geq 2\|f\| - \|f\| = \|f\| \\ &\geq \Phi(u^*).\end{aligned}$$

Also ist u^* sogar das Minimum von Φ auf ganz V . \square

Das folgende Beispiel zeigt, dass man in Satz I.2.4 nicht auf die Voraussetzung $\dim V < \infty$ verzichten kann.

BEISPIEL I.2.5. Sei X der Raum der Nullfolgen $x = (x_n)_{n \in \mathbb{N}^*}$ versehen mit der Supremum-Norm und V der Unterraum der Nullfolgen mit

$$\sum_{n=1}^{\infty} 2^{-n} x_n = 0.$$

Dann gibt es zu keinem $f \in X \setminus V$ eine b.A. in V .

Denn: Sei $f \in X \setminus V$ beliebig. Dann ist

$$\lambda = \sum_{n=1}^{\infty} 2^{-n} f_n \neq 0.$$

Wir behaupten, dass

$$\rho_V(f) \leq |\lambda|$$

ist. Definiere hierzu $u_m \in X$ durch

$$u_m = f - \lambda \frac{2^m}{2^m - 1} (\underbrace{1, \dots, 1}_{m\text{-mal}}, 0, \dots).$$

Dann ist

$$\begin{aligned}\sum_{n=1}^{\infty} 2^{-n} u_{m,n} &= \sum_{n=1}^m 2^{-n} (f_n - \lambda \frac{2^m}{2^m - 1}) + \sum_{n=m+1}^{\infty} 2^{-n} f_n \\ &= \underbrace{\sum_{n=1}^{\infty} 2^{-n} f_n}_{=\lambda} - \lambda \frac{2^m}{2^m - 1} \underbrace{\sum_{n=1}^m 2^{-n}}_{=\frac{\frac{1}{2} - 2^{-m-1}}{1 - \frac{1}{2}} = 1 - 2^{-m}} \\ &= 0.\end{aligned}$$

Also ist $u_m \in V$ für jedes m und somit

$$\rho_V(f) \leq \|f - u_m\| = |\lambda| \frac{2^m}{2^m - 1}$$

für jedes m . Also ist $\rho_V(f) \leq |\lambda|$.

Wir nehmen nun an, dass es zu f eine b.A. $u^* = (u_n)_{n \in \mathbb{N}^*}$ gibt. Dann ist natürlich

$$\|f - u^*\| \leq |\lambda|.$$

Da $f - u^*$ eine Nullfolge ist, gibt es ein $N \in \mathbb{N}^*$ mit

$$|f_n - u_n| \leq \frac{1}{2} |\lambda| \quad \text{für alle } n \geq N.$$

Damit folgt

$$\begin{aligned}
|\lambda| &= \left| \sum_{n=1}^{\infty} 2^{-n} f_n \right| \\
&= \left| \sum_{n=1}^{\infty} 2^{-n} f_n - \underbrace{\sum_{n=1}^{\infty} 2^{-n} u_n}_{=0} \right| \\
&\leq \sum_{n=1}^{\infty} 2^{-n} |f_n - u_n| \\
&= \sum_{n=1}^{N-1} 2^{-n} \underbrace{|f_n - u_n|}_{\leq \|f - u^*\| \leq |\lambda|} + \sum_{n=N}^{\infty} 2^{-n} \underbrace{|f_n - u_n|}_{\leq \frac{1}{2} |\lambda|} \\
&\leq \sum_{n=1}^{N-1} 2^{-n} |\lambda| + \sum_{n=N}^{\infty} 2^{-n} \frac{1}{2} |\lambda| \\
&= |\lambda| \left\{ \frac{\frac{1}{2} - 2^{-N}}{1 - \frac{1}{2}} + \frac{1}{2} \left[1 - \frac{\frac{1}{2} - 2^{-N}}{1 - \frac{1}{2}} \right] \right\} \\
&= \frac{1}{2} |\lambda| \left\{ 1 + \frac{\frac{1}{2} - 2^{-N}}{1 - \frac{1}{2}} \right\} \\
&= (1 - 2^{-N}) |\lambda| \\
&< |\lambda|.
\end{aligned}$$

Dies ist ein Widerspruch.

Wegen der grundsätzlichen Bedeutung von Kompaktheitsargumenten in der Approximationstheorie hat man spezielle Kompaktheitsbegriffe eingeführt.

DEFINITION I.2.6. (1) V heißt *beschränkt kompakt*, wenn für jedes $c > 0$ die Teilmenge $\{u \in V : \|u\| \leq c\}$ kompakt ist.

(2) Eine Folge $(u_n)_{n \in \mathbb{N}} \subset V$ heißt *Minimalfolge*, wenn für ein $f \in X$ gilt

$$\lim_{n \rightarrow \infty} \|f - u_n\| = \rho_V(f).$$

(3) V heißt *approximativ kompakt*, wenn jede Minimalfolge aus V eine konvergente Teilfolge mit Grenzwert in V besitzt.

Die folgende Bemerkung ist offensichtlich.

BEMERKUNG I.2.7. (1) Jede kompakte Menge ist beschränkt kompakt. Jede beschränkt kompakte Menge ist approximativ kompakt.

(2) Jeder endlich dimensionale Teilraum und jede abgeschlossene Teilmenge eines endlich dimensionalen Teilraumes ist beschränkt kompakt.

(3) Jede approximativ kompakte Menge ist Existenzmenge.

I.3. Allgemeine Eindeutigkeitsaussagen

Aus der Analysis ist bekannt, dass auf dem \mathbb{R}^n alle Normen äquivalent sind, also auf die gleichen Konvergenzaussagen führen. Dieser Abschnitt zeigt, dass verschiedene Normen aber zu sehr unterschiedlichen Eindeutigkeitsaussagen für b.A. führen. Hier sind die metrischen Eigenschaften einer Norm relevant und nicht ihre topologischen.

DEFINITION I.3.1. (1) Eine Menge M eines Vektorraumes heißt *konvex*, wenn für alle $x, y \in M$ und alle $\lambda \in (0, 1)$ gilt

$$\lambda x + (1 - \lambda)y \in M.$$

(2) Sei M eine konvexe Menge. Eine Funktion $\varphi : M \rightarrow \mathbb{R}$ heißt *konvex*, wenn für alle $x, y \in M$ und alle $\lambda \in (0, 1)$ gilt

$$\varphi(\lambda x + (1 - \lambda)y) \leq \lambda\varphi(x) + (1 - \lambda)\varphi(y).$$

BEMERKUNG I.3.2. (1) Seien M eine konvexe Menge und $\varphi : M \rightarrow \mathbb{R}$ eine Funktion. Dann ist φ genau dann konvex, wenn die Menge

$$G = \{(x, s) \in M \times \mathbb{R} : s \geq \varphi(x)\}$$

konvex ist.

(2) In normierten Räumen sind die offenen und abgeschlossenen Kugeln konvex.

BEMERKUNG I.3.3. In einer konvexen Menge V ist die Menge der b.A. stets konvex.

BEWEIS. Sei $f \in X$ beliebig. Dann ist die Menge der b.A. an f in V gleich $B_{\rho_V(f)}(f) \cap V$, wenn $B_r(g) = \{u \in X : \|u - g\| \leq r\}$ den abgeschlossenen Ball um g mit Radius r bezeichnet. Da der Durchschnitt beliebig vieler konvexer Mengen konvex ist, folgt hieraus die Behauptung. \square

Eindeutigkeitsaussagen erhalten wir, wenn wir schärfere Konvexitätseigenschaften für die Einheitskugel eines normierten Raumes fordern.

DEFINITION I.3.4. (1) Ein normierter Raum $(X, \|\cdot\|)$ heißt *strikt konvex*, wenn für alle $x, y \in X$ mit $\|x\| = \|y\| = 1$ und $x \neq y$ gilt

$$\left\| \frac{1}{2}(x + y) \right\| < 1.$$

(2) Ein normierter Raum $(X, \|\cdot\|)$ heißt *uniform konvex*, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x, y \in X$ mit $\|x\| = \|y\| = 1$ und $\|x - y\| > \varepsilon$ gilt

$$\left\| \frac{1}{2}(x + y) \right\| < 1 - \delta.$$

Anschaulich bedeutet strikte Konvexität, dass der Rand der Einheitskugel keine Geradenstücke positiver Länge enthalten darf. $(\mathbb{R}^2, \|\cdot\|_2)$ ist strikt konvex, $(\mathbb{R}^2, \|\cdot\|_\infty)$ ist nicht strikt konvex.

SATZ I.3.5. *Jede konvexe (und damit insb. jede lineare) Teilmenge eines strikt konvexen Raumes ist Eindeutigkeitsmenge.*

BEWEIS. Seien u_1 und u_2 zwei verschiedene b.A. an $f \in X$ in V . Da V konvex ist, ist $u_0 = \frac{1}{2}(u_1 + u_2) \in V$. Setze $\rho = \|f - u_1\|$. O.E. ist $\rho > 0$. Dann folgt aus der strikten Konvexität

$$\frac{1}{\rho} \|f - u_0\| = \left\| \frac{1}{2\rho}(f - u_1) + \frac{1}{2\rho}(f - u_2) \right\| < 1.$$

Dies ist aber ein Widerspruch zu $\rho = \rho_V(f)$. □

Das folgende Beispiel zeigt, dass man auf die strikte Konvexität in Satz I.3.5 nicht verzichten kann.

BEISPIEL I.3.6. Seien $X = \mathbb{R}^2$, $\|\cdot\| = \|\cdot\|_\infty$ und $V = \{(x, 0) : x \in \mathbb{R}\}$. Dann sind zu $f = (f_1, f_2) \in X$ alle Elemente $(x, 0) \in V$ mit $f_1 - |f_2| \leq x \leq f_1 + |f_2|$ b.A. an f .

BEISPIEL I.3.7. Seien $a, b \in \mathbb{R}$ mit $a < b$. $(C([a, b]), \|\cdot\|_\infty)$ ist nicht strikt konvex, denn für

$$f(x) = 1, \quad g(x) = \frac{x - a}{b - a}$$

gilt

$$\|f\|_\infty = \|g\|_\infty = \left\| \frac{1}{2}(f + g) \right\|_\infty = 1.$$

Allgemeiner ist für jede kompakte Teilmenge T eines normierten Raumes, die mehr als einen Punkt enthält, der Raum $(C(T), \|\cdot\|_\infty)$ nicht strikt konvex.

BEISPIEL I.3.8. Jeder Hilbert-Raum ist uniform konvex. Denn sind $f, g \in X$ und bezeichnen (\cdot, \cdot) und $\|\cdot\|$ das Skalarprodukt und die Norm von X , so gilt die *Parallelogrammgleichung*

$$\|f + g\|^2 + \|f - g\|^2 = 2\|f\|^2 + 2\|g\|^2.$$

Ist insbesondere $\|f\| = \|g\| = 1$, so folgt

$$\begin{aligned} \left\| \frac{1}{2}(f + g) \right\|^2 &= \frac{1}{4} \|f + g\|^2 \\ &= \frac{1}{4} [2\|f\|^2 + 2\|g\|^2 - \|f - g\|^2] \\ &= 1 - \frac{1}{4} \|f - g\|^2. \end{aligned}$$

Dies beweist die uniforme Konvexität mit $\delta = 1 - \sqrt{1 - \frac{1}{4}\varepsilon^2}$.

SATZ I.3.9. *Seien $1 < p < \infty$ und I ein perfektes Intervall. Dann ist $L^p(I)$ uniform konvex.*

BEWEIS. Für einen detaillierten Beweis verweisen wir auf F. Hirzebruch, W. Scharlau: Einführung in die Funktionalanalysis §17. Im Folgenden skizzieren wir die wesentlichen Schritte des Beweises.

Fall $p = 2$: Da $L^2(I)$ ein Hilbert-Raum ist, folgt die Behauptung aus Beispiel I.3.8.

Fall $p > 2$: Seien a, b zwei reelle Zahlen. Dann folgt aus der Jensenschen Ungleichung

$$\begin{aligned} \{|a + b|^p + |a - b|^p\}^{\frac{1}{p}} &\leq \{|a + b|^2 + |a - b|^2\}^{\frac{1}{2}} \\ &= \sqrt{2} \{|a|^2 + |b|^2\}^{\frac{1}{2}}. \end{aligned}$$

Definiere r durch $\frac{2}{p} + \frac{1}{r} = 1$, d.h. $r = \frac{p}{p-2}$. Dann folgt aus der Hölderschen Ungleichung

$$\begin{aligned} a^2 + b^2 &\leq \left\{ (a^2)^{\frac{p}{2}} + (b^2)^{\frac{p}{2}} \right\}^{\frac{2}{p}} 2^{\frac{1}{r}} \\ &= \{|a + b|^p + |a - b|^p\}^{\frac{2}{p}} 2^{\frac{p-2}{p}} \end{aligned}$$

und damit

$$\begin{aligned} \sqrt{2} \{|a|^2 + |b|^2\}^{\frac{1}{2}} &\leq \{|a + b|^p + |a - b|^p\}^{\frac{1}{p}} 2^{\frac{p-2}{2p} + \frac{1}{2}} \\ &= 2^{\frac{p-1}{p}} \{|a + b|^p + |a - b|^p\}^{\frac{1}{p}}. \end{aligned}$$

Diese beiden Abschätzungen ergeben zusammen

$$\left| \frac{1}{2}(a + b) \right|^p + \left| \frac{1}{2}(a - b) \right|^p \leq \frac{1}{2} \{|a|^p + |b|^p\}.$$

Setzen wir $a = f(t)$ und $b = g(t)$ ein und integrieren über I , erhalten wir

$$\left\| \frac{1}{2}(f + g) \right\|_p^p + \left\| \frac{1}{2}(f - g) \right\|_p^p \leq \frac{1}{2} \{ \|f\|_p^p + \|g\|_p^p \}.$$

Ist insbesondere $\|f\|_p = \|g\|_p = 1$, folgt

$$\left\| \frac{1}{2}(f + g) \right\|_p^p \leq 1 - 2^{-p} \left\| \frac{1}{2}(f - g) \right\|_p^p.$$

Wie in Beispiel I.3.8 folgt hieraus die uniforme Konvexität mit $\delta = 1 \{1 - 2^{-p} \varepsilon^p\}^{\frac{1}{p}}$.

Fall $1 < p < 2$: Definiere q durch $\frac{1}{p} + \frac{1}{q} = 1$, d.h. $q = \frac{p}{p-1}$. Durch Binomialentwicklung und Ausnutzen der Monotonie von $t \mapsto \frac{1}{t}(1 - z^t)$ für festes $z \in (0, 1)$, zeigt man zunächst, dass für alle $z \in [0, 1]$ gilt

$$(1 + z^q)^{p-1} \leq \frac{1}{2} [(1 - z)^p + (1 + z)^p].$$

Hieraus folgert man für beliebige reelle Zahlen a, b die Ungleichung

$$|a + b|^q + |a - b|^q \leq \{|a|^p + |b|^p\}^{q-1}.$$

Setzt man wieder $a = f(t)$, $b = g(t)$, integriert über I und benutzt die Höldersche Ungleichung, liefert dies

$$\|f + g\|_p^{\frac{p}{p-1}} + \|f - g\|_p^{\frac{p}{p-1}} \leq 2 \left[\|f\|_p^p + \|g\|_p^p \right]^{\frac{1}{p-1}}$$

bzw.

$$\left\| \frac{1}{2}(f + g) \right\|_p^{\frac{p}{p-1}} + \left\| \frac{1}{2}(f - g) \right\|_p^{\frac{p}{p-1}} \leq \left[\frac{1}{2} \|f\|_p^p + \frac{1}{2} \|g\|_p^p \right]^{\frac{1}{p-1}}.$$

Hieraus folgt wie in den anderen Fällen die uniforme Konvexität. \square

BEISPIEL I.3.10. Der Raum $L^1((0, 1))$ ist nicht uniform konvex. Definiere dazu die Funktionen f und g durch (vgl. Abbildung I.3.1)

$$f(x) = \begin{cases} 4 - 8x & \text{für } 0 \leq x \leq \frac{1}{2}, \\ 0 & \text{für } \frac{1}{2} \leq x \leq 1, \end{cases}$$

$$g(x) = \begin{cases} 0 & \text{für } 0 \leq x \leq \frac{1}{2}, \\ 8x - 4 & \text{für } \frac{1}{2} \leq x \leq 1, \end{cases}$$

Dann gilt

$$\|f\|_1 = \|g\|_1 = \left\| \frac{1}{2}(f + g) \right\|_1 = 1.$$

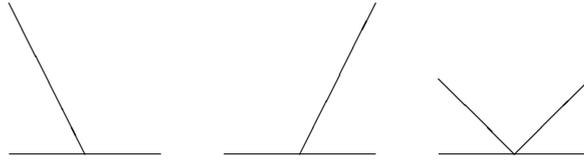


ABBILDUNG I.3.1. Funktionen f , g und $\frac{1}{2}(f + g)$ aus Beispiel I.3.10

SATZ I.3.11. Jede abgeschlossene konvexe Menge V in einem uniform konvexen Banach-Raum X ist Čebyšev-Menge.

BEWEIS. Wegen Satz I.3.5 ist nur die Existenz einer b.A. zu zeigen. Sei dazu $f \in X$ beliebig. O.E. ist $f = 0$, andernfalls transformieren wir alle Elemente durch $g \mapsto g - f$ und die approximierende Menge durch $V \mapsto V' = \{v - f : v \in V\}$. Setze $\rho = \rho_V(f)$.

Ist $\rho = 0$, liegt f im Abschluss von V und damit in V , da V als abgeschlossen vorausgesetzt ist. Daher ist in diesem Fall nichts zu zeigen. Sei also $\rho > 0$ und $(u_n)_{n \in \mathbb{N}}$ eine Minimalfolge. Sei $\varepsilon > 0$ beliebig und $\delta = \delta(\varepsilon)$ wie in Definition I.3.4 (2). Wegen

$$\lim_{n \rightarrow \infty} \|u_n\| = \lim_{n \rightarrow \infty} \|u_n - f\| = \rho$$

gibt es ein n_0 mit

$$\|u_n\| \leq \rho(1 + \delta)$$

für alle $n \geq n_0$. O.E. gilt $\|u_n\| > \rho$ für alle n , da sonst eine b.A. vorläge. Setze

$$\lambda_n = \|u_n\|^{-1}.$$

Für $n, m \geq n_0$ folgt dann

$$\begin{aligned} & \left\| \frac{1}{2}(\lambda_n u_n + \lambda_m u_m) \right\| \\ &= \left\| \frac{1}{2\rho}(u_n + u_m) - \frac{1}{2} \left(\frac{1}{\rho} - \lambda_n \right) u_n - \frac{1}{2} \left(\frac{1}{\rho} - \lambda_m \right) u_m \right\| \\ &\geq \frac{1}{\rho} \underbrace{\left\| \frac{1}{2}(u_n + u_m) \right\|}_{\substack{\in V \\ \geq \rho}} - \frac{1}{2} \left(\frac{1}{\rho} - \lambda_n \right) \|u_n\| - \frac{1}{2} \left(\frac{1}{\rho} - \lambda_m \right) \|u_m\| \\ &\geq 1 - \frac{1}{2\rho} \underbrace{\|u_n\|}_{\leq \rho(1+\delta)} - \frac{1}{2\rho} \underbrace{\|u_m\|}_{\leq \rho(1+\delta)} + \frac{1}{2} \underbrace{\lambda_n \|u_n\|}_{=1} + \frac{1}{2} \underbrace{\lambda_m \|u_m\|}_{=1} \\ &\geq 2 - (1 + \delta) \\ &= 1 - \delta. \end{aligned}$$

Wegen der uniformen Konvexität folgt hieraus

$$\|\lambda_n u_n - \lambda_m u_m\| < \varepsilon.$$

Also ist $(\lambda_n u_n)_{n \in \mathbb{N}}$ und damit auch $(\rho \lambda_n u_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge in X . Da X vollständig ist, konvergiert $\rho \lambda_n u_n$ gegen ein $u^* \in X$. Insbesondere ist

$$\|u^*\| = \lim_{n \rightarrow \infty} \underbrace{\|\rho \lambda_n u_n\|}_{= \rho} = \rho.$$

Da $(u_n)_{n \in \mathbb{N}}$ beschränkt ist und $(\lambda_n)_{n \in \mathbb{N}}$ gegen $\frac{1}{\rho}$ konvergiert, ist $(1 - \rho \lambda_n)u_n$ eine Nullfolge. Daher konvergiert $(u_n)_{n \in \mathbb{N}}$ auch gegen u^* . Da V abgeschlossen ist, ist $u^* \in V$. Wegen $\|u^*\| = \rho$ ist u^* die gesuchte b.A.. \square

DEFINITION I.3.12. Sei V eine nicht leere Teilmenge eines normierten Raumes X . Weiter sei \tilde{X} die Teilmenge derjenigen Element von X , die eine eindeutige b.A. in V besitzen. Die Abbildung $T : \tilde{X} \rightarrow V$, die jedem $f \in \tilde{X}$ die b.A. zuordnet, heißt *metrische Projektion*.

BEMERKUNG I.3.13. (1) Ist V eine Čebyšev-Menge, so ist die metrische Projektion auf ganz X definiert.

(2) Selbst wenn V ein Vektorraum ist, ist die metrische Projektion i.a. keine lineare Abbildung.

(3) Ist X ein Hilbert-Raum und V ein Unterraum von X , so ist die metrische Projektion linear.

BEWEIS. *ad* (1): Ist offensichtlich.

ad (2): Sei $X = C([0, 1])$ versehen mit der Supremum-Norm und V der Unterraum der konstanten Funktionen. Für

$$f = \begin{cases} -1 + 4x & \text{für } 0 \leq x \leq \frac{1}{2}, \\ 1 & \text{für } \frac{1}{2} \leq x \leq 1, \end{cases}$$

$$g = \begin{cases} 1 & \text{für } 0 \leq x \leq \frac{1}{2}, \\ 3 - 4x & \text{für } \frac{1}{2} \leq x \leq 1 \end{cases}$$

gilt $Tf = Tg = 0$ und $T(f + g) = 1$ (vgl. Abbildung I.3.2).

ad (3): Folgt aus der Charakterisierung

$$(f - Tf, u) = 0 \quad \text{für alle } u \in V$$

der metrischen Projektion in Hilbert-Räumen. □

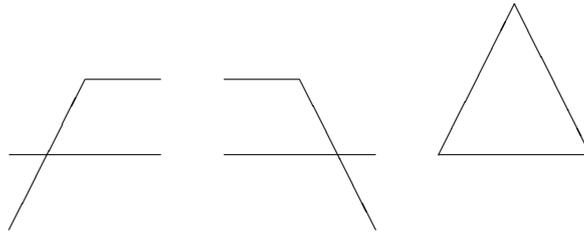


ABBILDUNG I.3.2. Funktionen f , g und $f + g$ aus dem Beweis von Bemerkung I.3.13 (2)

SATZ I.3.14. *Für jede approximativ kompakte Čebyšev-Menge ist die metrische Projektion stetig.*

BEWEIS. Es reicht, die Folgenstetigkeit der metrischen Projektion zu zeigen. Seien dazu $f \in X$ und $(f_n)_{n \in \mathbb{N}} \subset X$ mit $\lim_{n \rightarrow \infty} f_n = f$. Wegen der Optimalität von Tf_n bzgl. $\|f_n - u\|$ erhalten wir

$$\begin{aligned} \|f - Tf_n\| &\leq \|f - f_n\| + \underbrace{\|f_n - Tf_n\|}_{\leq \|f_n - Tf\|} \\ &\leq \|f - f_n\| + \|f_n - Tf\| \\ &\leq \|f - f_n\| + \|f_n - f\| + \underbrace{\|f - Tf\|}_{=\rho_V(f)} \\ &= 2\|f - f_n\| + \rho_V(f). \end{aligned}$$

Also ist $(Tf_n)_{n \in \mathbb{N}} \subset V$ eine Minimalfolge zu f . Da V approximativ kompakt ist, konvergiert eine Teilfolge gegen ein $u^* \in V$ mit

$$\|f - u^*\| = \rho_V(f).$$

Wegen der Eindeutigkeit der b.A. folgt $u^* = Tf$. Da dies für jede konvergente Teilfolge von $(Tf_n)_{n \in \mathbb{N}}$ gilt, folgt $\lim_{n \rightarrow \infty} Tf_n = Tf$. \square

BEMERKUNG I.3.15. Man kann auf die approximative Kompaktheit in Satz I.3.14 nicht verzichten.

Für die Eindeutigkeitsaussagen haben wir immer nur die metrischen Eigenschaften der Norm ausgenutzt. Es spielen aber auch topologische Eigenschaften eine Rolle. Dies zeigt die folgende Bemerkung.

BEMERKUNG I.3.16. Sei $V \subset \mathbb{R}^n$ eine beschränkte Čebyšev-Menge bzgl. einer beliebigen Norm. Dann besitzt jede stetige Selbstabbildung $\varphi : V \rightarrow V$ mindestens einen Fixpunkt. Insbesondere kann die Sphäre S^{n-1} bezüglich keiner Norm eine Čebyšev-Menge sein.

BEWEIS. V ist als Čebyšev-Menge abgeschlossen. Als beschränkte und abgeschlossene Teilmenge des \mathbb{R}^n ist V daher kompakt. Insbesondere ist V in einer Kugel $B_r(0)$ um Null mit geeignetem Radius r enthalten. Gemäß Satz I.3.14 ist die metrische Projektion $T : B_r(0) \rightarrow V$ stetig. Also ist die Abbildung $\varphi \circ T : B_r(0) \rightarrow V \subset B_r(0)$ stetig und besitzt nach dem Browserschen Fixpunktsatz einen Fixpunkt $u^* \in B_r(0)$. Wegen $u^* = \varphi \circ T(u^*)$ und $\varphi \circ T(B_r(0)) \subset V$ ist $u^* \in V$ und $Tu^* = u^*$. Also ist u^* auch ein Fixpunkt von φ .

Die Abbildung $S^{n-1} \rightarrow S^{n-1}$ mit $x \mapsto -x$ besitzt offensichtlich keinen Fixpunkt. \square

KAPITEL II

Lineare Čebyšev-Approximation

Im Mittelpunkt dieses Kapitels steht die Approximation durch endlich dimensionale Unterräume. Wegen Satz I.2.4 (S. 8) ist die Existenz einer b.A. gesichert. Wegen der Sätze I.3.9 (S. 13) und I.3.11 (S. 14) ist die Eindeutigkeit in L^p -Räumen mit $1 < p < \infty$ ebenfalls gesichert. Wegen der Beispiele I.3.6 (S. 12) und I.3.7 (S. 12) ist die Situation für die Supremum-Norm $\|\cdot\|_\infty$ unklar. Die Approximation bezüglich dieser Norm wurde 1859 erstmals von P. L. Čebyšev untersucht. Daher spricht man in diesem Zusammenhang von *Čebyšev-Approximation*. Sie steht im Mittelpunkt dieses und der nächsten Kapitel.

Ist $V = \text{span}\{v_1, \dots, v_n\}$ ein endlich dimensionaler Raum und T eine kompakte Menge, so kann man die Čebyšev-Approximation in $C(T)$ durch V auch wie folgt formulieren:

Gesucht sind Zahlen a_1, \dots, a_n und ρ , so dass ρ minimal wird unter den Nebenbedingungen

$$\begin{aligned} f(t) - \sum_{k=1}^n a_k v_k(t) &\leq \rho \quad \text{für alle } t \in T, \\ -f(t) + \sum_{k=1}^n a_k v_k(t) &\leq \rho \quad \text{für alle } t \in T. \end{aligned}$$

Da hier ein Optimierungsproblem mit endlich vielen Freiheitsgraden, aber unendlich vielen Nebenbedingungen vorliegt, spricht man von einem *semi-infiniten Programm*. Ein wesentliches Ergebnis der folgenden Abschnitte ist die Erkenntnis, dass bereits endlich viele Nebenbedingungen zur Lösung dieses Approximationsproblems ausreichen.

II.1. Kolmogoroff-Kriterium

Im Folgenden ist stets T eine kompakte Menge, $\|\cdot\|$ die Supremum-Norm auf $C(T)$, V ein endlich dimensionaler Unterraum von $C(T)$ mit Basis v_1, \dots, v_n , f ein beliebiges Element von $C(T)$ und u ein beliebiges Element von V .

DEFINITION II.1.1. (1) Zu $f \in C(T)$ bezeichnet

$$M[f] = \{t \in T : |f(t)| = \|f\|\}$$

die Menge der Extremalpunkte von f .

(2) Seien $f \in C(T)$, $u_0 \in V$ und $D \subset T$ abgeschlossen. Dann heißt D

eine *Referenz* zu $f - u_0$, wenn gilt

$$\inf_{t \in D} \operatorname{Re} \left[\overline{(f(t) - u_0(t))} u(t) \right] \leq 0$$

für alle $u \in V$.

BEMERKUNG II.1.2. Sind alle Funktionen reellwertig, so ist D eine Referenz zu $f - u_0$, wenn gilt

$$\inf_{t \in D} (f(t) - u_0(t)) u(t) \leq 0$$

für alle $u \in V$.

Obere Schranken für $\rho_V(f)$ gewinnt man leicht, indem man $\|f - u\|$ für ein beliebiges $u \in V$ betrachtet. Referenzen liefern untere Schranken für $\rho_V(f)$.

SATZ II.1.3. *Sei D eine Referenz zu $f - u_0$. Dann gilt*

$$\rho_V(f) \geq \inf_{t \in D} |f(t) - u_0(t)|.$$

BEWEIS. Setze

$$\eta = \inf_{t \in D} |f(t) - u_0(t)|.$$

O.E. ist $\eta > 0$, da die Behauptung sonst offensichtlich erfüllt ist. Seien $u_1 \in V$ und $t \in D$ beliebig. Dann gilt

$$\begin{aligned} & \operatorname{Re} \left[\overline{(f(t) - u_0(t))} (u_1(t) - u_0(t)) \right] \\ &= \operatorname{Re} \left[\overline{(f(t) - u_0(t))} (f(t) - u_0(t) - f(t) + u_1(t)) \right] \\ &= |f(t) - u_0(t)|^2 - \underbrace{\operatorname{Re} \left[\overline{(f(t) - u_0(t))} (f(t) - u_1(t)) \right]}_{\leq |f(t) - u_0(t)| |f(t) - u_1(t)|} \\ &\geq |f(t) - u_0(t)| \left\{ \underbrace{|f(t) - u_0(t)|}_{\geq \eta} - |f(t) - u_1(t)| \right\} \\ &\geq |f(t) - u_0(t)| \{ \eta - |f(t) - u_1(t)| \}. \end{aligned}$$

Wegen $u_1 - u_0 \in V$ gibt es ein $t_1 \in D$ mit

$$\operatorname{Re} \left[\overline{(f(t_1) - u_0(t_1))} (u_1(t_1) - u_0(t_1)) \right] \leq 0.$$

Für dieses t_1 folgt

$$\eta - |f(t_1) - u_1(t_1)| \leq 0$$

und somit

$$\eta \leq |f(t_1) - u_1(t_1)| \leq \|f - u_1\|.$$

Da $u_1 \in V$ beliebig war, folgt hieraus die Behauptung. \square

BEISPIEL II.1.4. Sei $T = [a, b]$ mit $a < b$ und $V = \mathbb{P}_n$ der Raum der Polynome vom Grad $\leq n$. Zu $f \in C(T)$ und $u_0 \in \mathbb{P}_n$ gebe es $n+2$ Punkte $a \leq t_0 < \dots < t_{n+1} \leq b$, in denen $f - u_0$ das Vorzeichen wechselt, d.h.

$$\operatorname{sgn} [f(t_i) - u_0(t_i)] = -\operatorname{sgn} [f(t_{i+1}) - u_0(t_{i+1})] \quad \text{für } 0 \leq i \leq n.$$

Dann ist $D = \{t_0, \dots, t_{n+1}\}$ eine Referenz. Denn andernfalls gäbe es ein Polynom $u_1 \in \mathbb{P}_n$ mit

$$(f(t_i) - u_0(t_i)) u_1(t_i) > 0 \quad \text{für } 0 \leq i \leq n+1.$$

Dann hätte u_1 aber $n+2$ Vorzeichenwechsel und damit $n+1$ Nullstellen. Dies ist ein Widerspruch.

SATZ II.1.5 (Kolmogoroff-Kriterium). *Es ist $u_0 \in V \subset C(T)$ genau dann eine b.A. an $f \in C(T)$, wenn $M[f - u_0]$ eine Referenz enthält, d.h. wenn*

$$\min_{t \in M[f - u_0]} \operatorname{Re} \left[\overline{(f(t) - u_0(t))} u(t) \right] \leq 0$$

ist für alle $u \in V$.

BEWEIS. „ \Leftarrow “: Satz II.1.3 angewandt auf $D = M[f - u_0]$ liefert

$$\begin{aligned} \rho_V(f) &\geq \min_{t \in M[f - u_0]} \underbrace{|f(t) - u_0(t)|}_{= \|f - u_0\|} \\ &= \|f - u_0\| \\ &\geq \rho_V(f). \end{aligned}$$

Also gilt überall die Gleichheit, und u_0 ist b.A. an f .

„ \Rightarrow “: Wir nehmen an, die Bedingung sei verletzt. Dann gibt es ein $u \in V$ mit

$$\alpha = \min_{t \in M[f - u_0]} \operatorname{Re} \left[\overline{(f(t) - u_0(t))} u(t) \right] > 0.$$

O.E. ist $\|u\| = 1$. Wir konstruieren nun eine bessere Approximation an f als u_0 .

Da $\operatorname{Re} \left[\overline{(f - u_0)u} \right]$ stetig ist, gibt es eine offene Umgebung U von $M[f - u_0]$ mit

$$\operatorname{Re} \left[\overline{(f(t) - u_0(t))} u(t) \right] \geq \frac{\alpha}{2}$$

für alle $t \in U$. Das Komplement $T \setminus U$ ist abgeschlossen und damit kompakt. Mithin nimmt $f - u_0$ sein Maximum in $T \setminus U$ an. Wegen der Definition von U und $M[f - u_0]$ folgt

$$\beta = \|f - u_0\| - \max_{t \in T \setminus U} |f(t) - u_0(t)| > 0.$$

Setze $\delta = \frac{1}{2} \min\{\alpha, \beta\}$. Wir behaupten, dass $u_1 = u_0 + \delta u$ eine bessere Approximation an f ist als u_0 . Sei dazu zunächst $t \in U$. Wegen

$$|a + b|^2 = |a|^2 + |b|^2 + 2 \operatorname{Re} \bar{a} b$$

folgt

$$\begin{aligned}
& |f(t) - u_1(t)|^2 \\
&= \left| \underbrace{f(t) - u_0(t)}_{=a} - \underbrace{\delta u(t)}_{=b} \right|^2 \\
&= \underbrace{|f(t) - u_0(t)|^2}_{\leq \|f - u_0\|^2} - 2\delta \operatorname{Re} \left[\underbrace{(f(t) - u_0(t))u(t)}_{\geq \frac{\alpha}{2}} \right] + \delta^2 \underbrace{|u(t)|^2}_{\leq 1} \\
&\leq \|f - u_0\|^2 - 2\delta \underbrace{\frac{\alpha}{2}}_{\geq \delta} + \delta^2 \\
&\leq \|f - u_0\|^2 - \delta^2.
\end{aligned}$$

Sei nun $t \in T \setminus U$. Mit der Dreiecksungleichung folgt

$$\begin{aligned}
|f(t) - u_1(t)| &= |f(t) - u_0(t) - \delta u(t)| \\
&\leq \underbrace{|f(t) - u_0(t)|}_{\leq \|f - u_0\| - \beta} + \delta \underbrace{|u(t)|}_{\leq 1} \\
&\leq |f(t) - u_0(t)| - \underbrace{\beta}_{\geq 2\delta} + \delta \\
&\leq \|f - u_0\| - \delta.
\end{aligned}$$

Also ist

$$\|f - u\| < \|f - u_0\|.$$

Dies ist ein Widerspruch. \square

BEISPIEL II.1.6. Betrachte die Approximation von $f(t) = \sin t$ auf $[0, \frac{\pi}{2}]$ durch $V = \{u = at : a \in \mathbb{R}\}$ und die Approximation $u_0(t) = \frac{3}{4}t$. Dann ist

$$\begin{aligned}
f\left(\frac{\pi}{4}\right) - u_0\left(\frac{\pi}{4}\right) &= \frac{\sqrt{2}}{2} - \frac{3\pi}{16} \approx 0.118, \\
f\left(\frac{\pi}{2}\right) - u_0\left(\frac{\pi}{2}\right) &= 1 - \frac{3\pi}{8} \approx -0.178.
\end{aligned}$$

Offensichtlich gibt es keine Funktion in V , die in den Punkten $\frac{\pi}{4}$ und $\frac{\pi}{2}$ verschiedenes Vorzeichen hat. Also ist $D = \{\frac{\pi}{4}, \frac{\pi}{2}\}$ eine Referenz. Aus Satz II.1.3 folgt $\rho_V(f) \geq 0.118$. Durch Verschieben des Punktes $\frac{\pi}{4}$ und gleichzeitiges Anpassen des Parameters a kann man die nach Satz II.1.5 existierende optimale Referenz finden.

BEMERKUNG II.1.7. Ist V nur konvex, so wird die b.A. durch die Bedingung

$$\min_{t \in M[f - u_0]} \operatorname{Re} \left[\overline{(f(t) - u_0(t))} (u(t) - u_0(t)) \right] \leq 0$$

für alle $u \in V$ charakterisiert.

II.2. Charakterisierungssatz

Wir benötigen einige Eigenschaften stetiger linearer Funktionale und einige Ergebnisse über Systeme von Ungleichungen.

SATZ II.2.1. *Sei X ein normierter Vektorraum. Eine lineares Funktional $\ell : X \rightarrow \mathbb{R}$ bzw. \mathbb{C} heißt beschränkt, wenn*

$$\|\ell\|_{X'} = \sup_{x \in X, \|x\|=1} |\ell(x)| < \infty$$

ist. Die Menge der beschränkten linearen Funktionale ist ein Vektorraum und wird mit X' bezeichnet. $\|\cdot\|_{X'}$ ist eine Norm auf X' .

BEMERKUNG II.2.2. Eine lineare Abbildung $\ell : X \rightarrow \mathbb{R}$ bzw. \mathbb{C} ist genau dann stetig, wenn sie beschränkt ist. Ist X vollständig, gilt gleiches für X' .

Da keine Verwechslungen möglich sind, lassen wir im Folgenden die Indizes X und X' bei den Normen weg.

BEISPIEL II.2.3. (1) Seien T kompakt, t_1, \dots, t_m verschiedene Punkte in T und $\alpha_1, \dots, \alpha_m$ reelle oder komplexe Zahlen. Dann definiert

$$\ell(f) = \sum_{i=1}^m \alpha_i f(t_i)$$

ein stetiges lineares Funktional auf $C(T)$ mit

$$\|\ell\| = \sum_{i=1}^m |\alpha_i|.$$

(2) Sei I ein kompaktes perfektes Intervall und $w \in C(I)$. Dann definiert

$$\ell(f) = \int_I w(t) f(t) dt$$

ein stetiges lineares Funktional auf $C(T)$ mit

$$\|\ell\| = \int_I |w(t)| dt.$$

(3) Sei X ein Hilbertraum und $g \in X$. Dann definiert

$$\ell(f) = (f, g)$$

ein stetiges lineares Funktional auf X mit

$$\|\ell\| = \|g\|.$$

Dabei sind (\cdot, \cdot) und $\|\cdot\|$ das Skalarprodukt und die Norm von X .

BEMERKUNG II.2.4. Sei V ein linearer Unterraum des normierten Raumes X . Wenn $\ell \in X'$ die Eigenschaften

$$\|\ell\| = 1$$

und

$$\ell(u) = 0 \quad \forall u \in V$$

hat, gilt für alle $f \in X$

$$\rho_V(f) \geq |\ell(f)|.$$

BEWEIS. Die Behauptung folgt aus

$$\begin{aligned} |\ell(f)| &= \left| \ell(f) - \underbrace{\ell(u)}_{=0} \right| = |\ell(f - u)| \leq \underbrace{\|\ell\|}_{=1} \|f - u\| \\ &= \|f - u\| \end{aligned}$$

für alle f und u . □

Eine konvexe Menge ist dadurch charakterisiert, dass für je zwei Elemente x_1 und x_2 und alle $\theta \in [0, 1]$ auch $\theta x_1 + (1 - \theta)x_2$ zu dieser Menge gehört. Durch Induktion folgt, dass für jedes $m \geq 2$, je m Elemente x_1, \dots, x_m und je m nicht negative Zahlen $\theta_1, \dots, \theta_m$ mit $\sum_{i=1}^m \theta_i = 1$ das Element $\sum_{i=1}^m \theta_i x_i$ auch zu der Menge gehört. Diese Beobachtung führt auf folgende Definition.

DEFINITION II.2.5. Sei M eine beliebige Teilmenge eines Vektorraumes X . Dann bezeichnet

$$\kappa(M) = \left\{ x \in X : x = \sum_{i=1}^m \theta_i x_i, x_i \in M, \theta_i \geq 0, \sum_{i=1}^m \theta_i = 1, m \in \mathbb{N}^* \right\}$$

die *konvexe Hülle* von M . Sie ist die kleinste konvexe Menge, die M enthält.

Die Zahl der Punkte, die in einer Konvexkombination auftreten, ist endlich, aber im Prinzip beliebig. In endlich dimensionalen Räumen gibt es aber eine von der Dimension abhängige obere Schranke für die Zahl dieser Punkte.

SATZ II.2.6 (Carathéodory). *Sei $M \subset \mathbb{R}^n$. Dann ist jeder Punkt in $\kappa(M)$ darstellbar als Konvexkombination von höchstens $n + 1$ Punkten.*

BEWEIS. Sei

$$y = \sum_{i=0}^m \theta_i x_i \in \kappa(M)$$

mit $\theta_i > 0$, $\sum \theta_i = 1$ und $x_i \in M$ und dies sei eine Darstellung mit einer minimalen Anzahl von Punkten.

Wir nehmen an, dass $m > n$ ist. Dann sind $x_1 - x_0, \dots, x_m - x_0$ linear abhängig. Also gibt es eine nicht triviale Relation

$$0 = \sum_{i=1}^m \beta_i (x_i - x_0) = \sum_{i=0}^m \alpha_i x_i,$$

bei der nicht alle β_i bzw. α_i verschwinden. Da $\sum \alpha_i = 0$ ist, gibt es mindestens einen positive Faktor α_i . Außerdem gilt für jedes $\lambda > 0$

$$y = \sum_{i=0}^m (\theta_i - \lambda \alpha_i) x_i.$$

Setze

$$\lambda = \min_{i, \alpha_i > 0} \frac{\theta_i}{\alpha_i}.$$

Dann gilt für alle i

$$d_i = \theta_i - \lambda \alpha_i \geq 0$$

und $d_i = 0$ für mindestens ein i . Wegen $\sum d_i = \sum \theta_i = 1$ ist also m in der Darstellung von y nicht minimal im Widerspruch zur Annahme. \square

SATZ II.2.7. *Die konvexe Hülle einer kompakten Menge im \mathbb{R}^n ist kompakt.*

BEWEIS. Der Simplex

$$\Delta = \left\{ \theta \in \mathbb{R}^{n+1} : \theta_i \geq 0, \sum \theta_i = 1 \right\}$$

ist abgeschlossen und beschränkt und somit kompakt. Ist $M \subset \mathbb{R}^n$ kompakt, ist somit auch $\Delta \times M^{n+1}$ kompakt. Die Abbildung

$$\begin{aligned} \Delta \times M^{n+1} &\rightarrow \kappa(M) \\ (\theta_1, \dots, \theta_{n+1}, x_1, \dots, x_{n+1}) &\mapsto \sum \theta_i x_i \end{aligned}$$

ist stetig und gemäß Satz II.2.6 surjektiv. Also ist $\kappa(M)$ kompakt. \square

SATZ II.2.8. *Sei $W \subset \mathbb{R}^n$ kompakt. Es existiert genau dann eine Lösung $z \in \mathbb{R}^n$ für das System von Ungleichungen*

$$(w, z) > 0 \quad \forall w \in W,$$

wenn $0 \notin \kappa(W)$ ist.

BEWEIS. (1) Sei $0 \in \kappa(W)$. Dann gibt es m positive Zahlen θ_i mit $\sum \theta_i = 1$ und m Elemente w_i in W mit $0 = \sum \theta_i w_i$. Angenommen z sei eine Lösung des Systems von Ungleichungen. Dann folgt

$$0 = (0, z) = \sum \underbrace{\theta_i}_{>0} \underbrace{(w_i, z)}_{>0} > 0.$$

Dies ist ein Widerspruch.

(2) Sei nun $0 \notin \kappa(W)$. Da gemäß Satz II.2.7 die konvexe Hülle $\kappa(W)$ kompakt ist, gibt es ein $z \in \kappa(W)$, das in der Euklidischen Norm den kleinsten Abstand zu 0 hat. Es ist

$$(z, z) = \|z\|^2 > 0.$$

Wir behaupten, dass z eine Lösung des Systems von Ungleichungen ist. Seien dazu $w \in W$ und $\theta \in (0, 1)$ beliebig. Wegen $\theta w + (1 - \theta)z \in \kappa(W)$ folgt

$$\begin{aligned} \|z\|^2 &\leq \|\theta w + (1 - \theta)z\|^2 = \|z + \theta(w - z)\|^2 \\ &= \|z\|^2 + 2\theta(w - z, z) + \theta^2 \|w - z\|^2 \end{aligned}$$

und somit

$$(w - z, z) \geq -\frac{\theta}{2} \|w - z\|^2.$$

Da dies für alle $\theta \in (0, 1)$ gilt, folgt

$$(w - z, z) \geq 0$$

und daher

$$(w, z) \geq (z, z) > 0. \quad \square$$

Aus den Sätzen II.2.6 und II.2.8 folgt:

KOROLLAR II.2.9. *Ist das System von Ungleichungen aus Satz II.2.8 nicht lösbar, gibt es $m \leq n + 1$ Punkte $w_i \in W$, so dass das System*

$$(w_i, z) > 0 \quad \forall 1 \leq i \leq m$$

bereits nicht lösbar ist.

Indem wir \mathbb{C}^n mit \mathbb{R}^{2n} identifizieren, erhalten wir:

KOROLLAR II.2.10. *Sei $W \subset \mathbb{C}^n$ kompakt. Das System von Ungleichungen*

$$\operatorname{Re}(\bar{w}, z) > 0 \quad \forall w \in W$$

ist genau dann lösbar, wenn $0 \notin \kappa(W)$ ist. Wenn es nicht lösbar ist, gibt es $m \leq 2n + 1$ Punkte $w_i \in W$, für die das System

$$\operatorname{Re}(\bar{w}_i, z) > 0 \quad \forall 1 \leq i \leq m$$

bereits nicht lösbar ist.

Nach diesen Vorbereitungen können wir den Hauptsatz dieses Abschnittes beweisen.

SATZ II.2.11 (Charakterisierungssatz). *Sei V ein n -dimensionaler Unterraum von $C(T)$ über \mathbb{R} bzw. \mathbb{C} und $f \in C(T)$ sowie $u_0 \in V$. Dann sind folgende Aussagen äquivalent:*

- (1) u_0 ist b.A. an f in V .
- (2) $M[f - u_0]$ enthält eine Referenz von höchstens $n + 1$ bzw. $2n + 1$ Punkten.
- (3) Es existieren m Punkte $t_i \in M[f - u_0]$ mit $m \leq n + 1$ bzw. $m \leq 2n + 1$ und m Zahlen $\theta_i > 0$ mit $\sum \theta_i = 1$, so dass gilt

$$\sum_{i=1}^m \theta_i \operatorname{Re} \overline{[f(t_i) - u_0(t_i)]} u(t_i) = 0 \quad \forall u \in V.$$

- (4) *Es existiert ein stetiges lineares Funktional $\ell : C(T) \rightarrow \mathbb{R}$ bzw. \mathbb{C} mit*

$$\begin{aligned}\|\ell\| &= 1, \\ \ell(f - u_0) &= \|f - u_0\|, \\ \ell(u) &= 0 \quad \forall u \in V.\end{aligned}$$

BEWEIS. Indem wir \mathbb{C} mit \mathbb{R}^2 identifizieren, können wir uns auf den reellen Fall beschränken.

(3) \Rightarrow (4): Definiere $\ell \in C(T)'$ durch

$$\ell(g) = \frac{1}{\|f - u_0\|} \sum_{i=1}^m \theta_i [f(t_i) - u_0(t_i)]g(t_i).$$

Wegen (3) ist $\ell(u) = 0$ für alle $u \in V$. Wegen $\sum \theta_i = 1$ ist, $\ell(f - u_0) = \|f - u_0\|$. Wegen $\theta_i > 0$ und $\sum \theta_i = 1$ sowie Beispiel II.2.3(1) ist $\|\ell\| = 1$.

(4) \Rightarrow (1): Wegen Bemerkung II.2.4 folgt

$$\begin{aligned}\rho_V(f) &\geq |\ell(f)| = |\ell(f - u_0)| \\ &= \|f - u_0\| \\ &\geq \rho_V(f).\end{aligned}$$

Also ist u_0 eine b.A. an f in V .

(2) \Rightarrow (1): Ist ein Spezialfall von Satz II.1.5 (S. 21).

(3) \Rightarrow (2): Sei $u \in V$ beliebig. Nach Voraussetzung ist

$$\sum_{i=1}^m \theta_i \operatorname{Re} [\overline{f(t_i) - u_0(t_i)}] u(t_i) = 0.$$

Also ist mindestens ein Summand nicht positiv. Wegen $\theta_i > 0$ für alle i , gibt es mindestens ein i mit

$$\operatorname{Re} [\overline{f(t_i) - u_0(t_i)}] u(t_i) \leq 0.$$

Also ist $\{t_1, \dots, t_m\}$ eine Referenz.

(1) \Rightarrow (3): Gemäß Satz II.1.5 (S. 21) ist das System von Ungleichungen

$$(f(t) - u_0(t), u(t)) > 0 \quad \forall t \in M[f - u_0]$$

nicht lösbar. Sei v_1, \dots, v_n eine Basis von V . Dann ist das System

$$\sum_{k=1}^n a_k (f(t) - u_0(t)) v_k(t) > 0 \quad \forall t \in M[f - u_0]$$

für die Koeffizienten a_1, \dots, a_n nicht lösbar. Mit

$$W = \{w \in \mathbb{R}^n : w_k = (f(t) - u_0(t))v_k(t), t \in M[f - u_0]\}$$

bedeutet dies, dass das System $(w, a) > 0$ nicht lösbar ist. Gemäß Korollar II.2.9 gibt es also $m \leq n + 1$ Punkte $t_i \in M[f - u_0]$, für die das System $(w^{(i)}, a) > 0$ mit $w^{(i)} = ((f(t_i) - u_0(t_i))v_k(t_i))_{1 \leq k \leq n}$

nicht lösbar ist. Wegen Satz II.2.8 ist daher $0 \in \kappa(w^{(1)}, \dots, w^{(m)})$. Dies bedeutet

$$\sum_{i=1}^m \theta_i (f(t_i) - u_0(t_i)) v_k(t_i) = 0 \quad \forall 1 \leq k \leq n$$

für geeignete $\theta_i > 0$ mit $\sum \theta_i = 1$. Da v_1, \dots, v_n eine Basis von V ist, folgt

$$\sum_{i=1}^m \theta_i (f(t_i) - u_0(t_i)) u(t_i) = 0 \quad \forall u \in V. \quad \square$$

II.3. Haarscher Eindeutigkeitsatz

Die Čebyšev-Approximation wird wesentlich einfacher, wenn die Familie der approximierenden Funktionen die sog. Haarsche Bedingung erfüllt. Polynome und andere wichtige Funktionenfamilien haben diese Eigenschaft. Im übernächsten Abschnitt werden wir zwar sehen, dass diese Theorie im wesentlichen auf Funktionen auf Intervallen beschränkt ist, dennoch ist ihre Bedeutung nicht zu unterschätzen.

DEFINITION II.3.1. Die kompakte Menge T enthalte mindestens $n + 1$ Punkte. Der n -dimensionale Unterraum V von $C(T)$ erfüllt die *Haarsche Bedingung* (kurz *HB*) und heißt ein *Haarscher Raum* (kurz *HR*), wenn eine der vier folgenden äquivalenten Bedingungen erfüllt ist:

- (1) Jede Funktion u in V hat höchstens $n - 1$ Nullstellen oder verschwindet identisch.
- (2) Zu je n paarweise verschiedenen Punkten t_1, \dots, t_n in T und n Daten y_1, \dots, y_n in \mathbb{R} ist die Interpolationsaufgabe

$$u(t_i) = y_i \quad 1 \leq i \leq n$$

stets eindeutig lösbar in V .

- (3) Wenn für n paarweise verschiedene Punkte t_1, \dots, t_n in T die Relation

$$\sum_{i=1}^n \beta_i u(t_i) = 0 \quad \forall u \in V$$

gilt, folgt $\beta_1 = \dots = \beta_n = 0$.

- (4) Sei v_1, \dots, v_n eine Basis von V . Dann verschwindet für je n paarweise verschiedene Punkte t_1, \dots, t_n in T die Determinante

$$D(t_1, \dots, t_n) = \begin{vmatrix} v_1(t_1) & v_1(t_2) & \dots & v_1(t_n) \\ \vdots & & & \vdots \\ v_n(t_1) & v_n(t_2) & \dots & v_n(t_n) \end{vmatrix}$$

nicht.

Die Äquivalenz obiger Bedingungen folgt sofort aus der Theorie linearer Gleichungssysteme. Die Interpolationsbedingung (2) ist nämlich äquivalent zu dem linearen Gleichungssystem

$$\sum_{k=1}^n a_k v_k(t_i) = y_i \quad 1 \leq i \leq n,$$

wenn v_1, \dots, v_n eine Basis von V ist. Dieses ist genau dann stets eindeutig lösbar, wenn das homogene Problem nur die triviale Lösung hat ((2) \iff (1)) bzw. wenn die Determinante der Matrix nicht verschwindet ((2) \iff (4)) bzw. wenn die Zeilen der Matrix linear unabhängig sind ((2) \iff (3)).

BEISPIEL II.3.2. (1) Sei I ein kompaktes, perfektes Intervall und $V = \mathbb{P}_n$ der Raum der algebraischen Polynome vom Grad $\leq n$. V ist ein Haarscher Raum der Dimension $n + 1$.

(2) Sei T das Intervall $[-\pi, \pi]$, wobei die Endpunkte identifiziert werden, und $V = \mathbb{T}_n = \{\cos(\ell\varphi), \sin(k\varphi) : 0 \leq \ell \leq n, 1 \leq k \leq n\}$ der Raum der trigonometrischen Polynome vom Grad $\leq n$. V ist ein Haarscher Raum der Dimension $2n + 1$.

(3) Der Raum $\text{span}\{\cos(\ell\varphi) : 0 \leq \ell \leq n\}$ der geraden trigonometrischen Polynome vom Grad n über dem Intervall $[0, \pi]$ ist ein Haarscher Raum der Dimension $n + 1$.

(4) Sei $T \subset \mathbb{R}^2$ das Einheitsquadrat und V der Raum der Polynome vom Grad $\leq n$, $n \geq 1$, in zwei Variablen. V ist kein Haarscher Raum.

BEWEIS. *ad* (1): Ein Polynom vom Grad $\leq n$, das nicht identisch verschwindet, hat höchstens n Nullstellen.

ad (2): Jedes $u \in \mathbb{T}_n$ läßt sich in der Form

$$u(\varphi) = \sum_{k=-n}^n a_k e^{ik\varphi} = e^{-in\varphi} \sum_{\ell=0}^{2n} a_{\ell-n} e^{i\ell\varphi}$$

mit komplexen Koeffizienten a_{-n}, \dots, a_n schreiben. Wegen $e^{-in\varphi} \neq 0$ für alle φ hat u höchstens $2n$ Nullstellen in $[-\pi, \pi]$. (Man beachte, dass die Endpunkte identifiziert und daher nur einfach gezählt werden.)

ad (3): Sei u ein gerades trigonometrisches Polynom vom Grad $\leq n$. Dann hat u höchstens $2n$ Nullstellen in $[-\pi, \pi]$. Wegen der Symmetrie liegt die Hälfte der Nullstellen in $[0, \pi]$.

ad (4): Die Funktion $u(x, y) = x - y$ ist ein lineares Polynom über T mit unendlich vielen Nullstellen. \square

Für den Eindeutigkeitssatz benötigen wir ein Ergebnis über die Fortsetzung stetiger Funktionen. Wir verdeutlichen dieses zunächst am Beispiel eines Intervalles.

BEISPIEL II.3.3. Seien I ein perfektes, kompaktes Intervall und $t_1 < \dots < t_n$ Punkte in I . Die Funktion f sei nur in diesen Punkten definiert.

Dann gibt es eine stetige Fortsetzung von f auf I mit gleicher Norm, d.h., es gibt ein $\tilde{f} \in C(I)$ mit

$$\begin{aligned}\tilde{f}(t_i) &= f(t_i) & 1 \leq i \leq n, \\ \|\tilde{f}\| &= \max_{1 \leq i \leq n} |f(t_i)|.\end{aligned}$$

Um dies einzusehen, setze man

$$\tilde{f}(t) = \begin{cases} f(t_1) & \text{für } t \leq t_1, \\ f(t_n) & \text{für } t \geq t_n, \\ \frac{1}{t_{i+1}-t_i} \{(t-t_i)f(t_{i+1}) + (t_{i+1}-t)f(t_i)\} & \text{für } t_i \leq t \leq t_{i+1}, \\ & 1 \leq i \leq n-1. \end{cases}$$

SATZ II.3.4. *Jede kompakte Menge T ist normal, d.h., jede auf einer abgeschlossenen Teilmenge D von T definierte stetige Funktion besitzt eine stetige Fortsetzung auf T mit gleicher Norm.*

BEWEIS. [8, S. 85] □

SATZ II.3.5 (Haarscher Eindeutigkeitsatz). *Sei T kompakt und V ein endlich dimensionaler Unterraum von $C(T)$. Die b.A. bzgl. der Maximumnorm in V ist genau dann für jedes $f \in C(T)$ eindeutig, wenn V ein Haarscher Raum ist.*

BEWEIS. „ \Leftarrow “: Die Existenz einer b.A. folgt aus Satz I.2.4 (S. 8). Die Eindeutigkeit ist für $f \in V$ offensichtlich. Sei also $f \notin V$. Sei $u \in V$ eine b.A. an f in V . Dann besteht $M[f-u]$ aus mindestens $n+1$ Punkten, wenn n die Dimension von V ist. Denn sonst gäbe es ein $u_1 \in V$ mit

$$u_1(t) = f(t) - u(t) \quad \forall t \in M[f-u].$$

Dann gilt aber

$$\operatorname{Re} \overline{(f(t) - u(t))} u_1(t) = |f(t) - u(t)|^2 > 0 \quad \forall t \in M[f-u].$$

Dies ist ein Widerspruch zum Kolmogoroff-Kriterium II.1.5 (S. 21).

Seien nun u_1 und u_2 zwei verschiedene b.A. an f in V . Dann ist $u = \frac{1}{2}(u_1 + u_2)$ ebenfalls eine b.A. an f . Wie soeben gezeigt, kann man n verschiedene Punkte t_1, \dots, t_n in $M[f-u]$ finden. Mit

$$\varepsilon = f - u, \quad \varepsilon_i = f - u_i, \quad \rho = \|\varepsilon\| = \|\varepsilon_1\| = \|\varepsilon_2\|$$

gilt dann für alle $1 \leq i \leq n$

$$\begin{aligned}\varepsilon(t_i) &= \frac{1}{2}(\varepsilon_1(t_i) + \varepsilon_2(t_i)), \\ \varepsilon(t_i) &= \rho, \\ \varepsilon_1(t_i) &\leq \rho, \\ \varepsilon_2(t_i) &\leq \rho.\end{aligned}$$

Wegen der strikten Konvexität von $(\mathbb{K}, |\cdot|)$ folgt hieraus

$$\varepsilon_1(t_i) = \varepsilon_2(t_i) = \varepsilon(t_i)$$

und somit

$$u_1(t_i) = u_2(t_i)$$

für alle i . Also hat $u_1 - u_2$ mindestens n Nullstellen im Widerspruch zur Haarschen Bedingung. Mithin ist die b.A. an f eindeutig.

„ \Rightarrow “: Die Haarsche Bedingung sei verletzt. Wir werden eine Funktion mit mehreren b.A. konstruieren.

Da die Haarsche Bedingung verletzt ist, gibt es ein $u_0 \in V$ mit $\|u_0\| = 1$ und n verschiedenen Nullstellen t_1, \dots, t_n . Sei v_1, \dots, v_n eine Basis von V . Dann ist $\det[v_k(t_i)]_{1 \leq i, k \leq n} = 0$. Also gibt es eine nicht-triviale Lösung des homogenen Gleichungssystems

$$\sum_{i=1}^n \beta_i v_k(t_i) = 0 \quad 1 \leq k \leq n.$$

Daraus folgt für alle $u \in V$

$$\sum_{i=1}^n \beta_i u(t_i) = 0.$$

O.E. ist $\sum |\beta_i| = 1$. Gemäß Satz II.3.4 gibt es eine stetige Funktion g auf T mit

$$\begin{aligned} \|g\| &= 1, \\ g(t_i) &= \begin{cases} \frac{\bar{\beta}_i}{|\beta_i|} & \text{falls } \beta_i \neq 0, \\ 0 & \text{sonst.} \end{cases} \end{aligned}$$

Setze

$$f_0 = g(1 - |u_0|).$$

Wir behaupten, dass alle Funktionen der Form au_0 mit $|a| \leq 1$ b.A. an f_0 sind. Für eine derartige Funktion gilt nämlich

$$\begin{aligned} |f_0(t) - au_0(t)| &\leq |f_0(t)| + |au_0(t)| \\ &= \underbrace{|g(t)|}_{\leq 1} \underbrace{|1 - u_0(t)|}_{\geq 0} + |au_0(t)| \\ &\leq 1 - |u_0(t)| + \underbrace{|a|}_{\leq 1} |u_0(t)| \\ &\leq 1 \end{aligned}$$

also

$$\|f_0 - au_0\| \leq 1.$$

Andererseits hat das lineare Funktional

$$\ell(f) = \sum_{i=1}^n \beta_i f(t_i)$$

wegen Beispiel II.2.3(1) (S. 23) und $\sum |\beta_i| = 1$ die Norm 1 und erfüllt

$$\ell(f_0) = \ell(g) = \sum_{i=1}^n \beta_i g(t_i) = \sum_{i=1}^n |\beta_i| = 1.$$

Wegen Bemerkung II.2.4 (S. 23) folgt $\rho_V(f_0) \geq 1$. Damit ist die Behauptung bewiesen. \square

Sehr elegant werden die Kriterien, wenn wir die Approximation durch einen reellen Haarschen Raum über einem reellen Intervall betrachten. Referenzen lassen sich dann gut geometrisch interpretieren. Wir erinnern daran, dass gemäß Satz II.2.11 (S. 26) im reellen Fall stets Referenzen mit höchstens $n + 1$ Punkten existieren. Im ersten Teil des Beweises von Satz II.3.5 haben wir gezeigt, dass in einem Haarschen Raum jede Referenz aus mindestens $n + 1$ Punkten besteht. Also haben für reelle Haarsche Räume Referenzen genau $n + 1$ Punkte.

SATZ II.3.6 (de la Vallée-Poussin). *Sei V ein reeller Haarscher Raum der Dimension n über dem Intervall I . Es gebe $n + 1$ Punkte $t_0 < \dots < t_n$ in I , in denen $f - u$ das Vorzeichen wechselt, d.h.*

$$\operatorname{sgn}[(f - u)(t_i)] = -\operatorname{sgn}[(f - u)(t_{i+1})] \neq 0 \quad 0 \leq i \leq n - 1.$$

Dann bilden diese Punkte eine Referenz, und es gilt

$$\rho_V(f) \geq \min_{0 \leq i \leq n} |f(t_i) - u(t_i)|.$$

BEWEIS. Der Beweis folgt wie in Beispiel II.1.4 (S. 21). (Dort wurde die Haarsche Bedingung für Polynome ausgenutzt.) \square

DEFINITION II.3.7. Zu $f - u$ gebe es m Punkte $t_1 < \dots < t_m$ in I , die Extrempunkte sind und in denen $f - u$ wechselndes Vorzeichen hat, d.h.

$$f(t_i) - u(t_i) = \sigma(-1)^i \|f - u\|$$

mit $\sigma \in \{-1, 1\}$. Dann heißen diese Punkte eine *Alternante* der Länge m . Gibt es keine längere Alternante, so sagt man, $f - u$ habe eine Alternante der genauen Länge m .

SATZ II.3.8 (Alternantenkriterium). *Sei V ein Haarscher Raum der Dimension n über dem Intervall I . Es ist $u \in V$ genau dann die b.A. zu $f \in C(I)$, wenn $f - u$ eine Alternante der Länge $n + 1$ besitzt.*

BEWEIS. „ \Leftarrow “ Folgt aus Satz II.3.6.

„ \Rightarrow “ Da es eine Referenz aus $n + 1$ Punkten gibt, müssen wir nur zeigen, dass die Vorzeichen auf den Referenzpunkten alternieren. Sei dazu u_0 die b.A. und $t_0 < \dots < t_n$ eine Referenz aus $M[f - u_0]$. Gemäß Satz II.2.11 (S. 26) gibt es Zahlen $\theta_i \geq 0$ mit $\sum \theta_i = 1$ und

$$\sum_{i=0}^n \theta_i (f(t_i) - u_0(t_i)) u(t_i) = 0$$

für alle $u \in V$. Wegen der Haarschen Bedingung (Punkt (3)) verschwindet keiner der Faktoren θ_i . Sei $1 \leq k \leq n$, Wegen der Haarschen Bedingung gibt es eine Funktion $u_1 \in V$ mit

$$u_1(t_k) = 1, \quad u_1(t_i) = 0 \quad 0 \leq i \leq n, i \neq k, i \neq k-1.$$

Da u_1 nur $n-1$ Nullstellen hat, liegt keine Nullstelle im Intervall $[t_{k-1}, t_k]$. Also ist insbesondere $u_1(t_{k-1}) > 0$. Setzen wir u_1 in obige Bedingung ein, erhalten wir

$$\underbrace{\theta_{k-1}u_1(t_{k-1})}_{>0}[f(t_{k-1}) - u_0(t_{k-1})] + \underbrace{\theta_k u_1(t_k)}_{>0}[f(t_k) - u_0(t_k)] = 0.$$

Also müssen $f(t_{k-1}) - u_0(t_{k-1})$ und $f(t_k) - u_0(t_k)$ entgegengesetztes Vorzeichen haben. \square

KOROLLAR II.3.9. *In einem reellen Haarschen Raum über einem Intervall gibt es eine positive Funktion.*

BEWEIS. O.E. ist die konstante Funktion $f = 1$ nicht in V enthalten. Sei u^* die b.A. an f in V . Wegen Satz II.3.8 ist u^* nicht die Nullfunktion. Also ist für jedes $t \in I$

$$|f(t) - u^*(t)| \leq \|f - u^*\| < \|f\| = 1$$

und somit $1 - u^*(t) < 1$, d.h. $u^*(t) > 0$. \square

KOROLLAR II.3.10. *Sei I ein kompaktes, perfektes Intervall und*

$$V = \left\{ u \in \mathbb{P}_n : u(t) = \sum_{k=0}^n a_k t^k, a_n \geq 0 \right\}.$$

Dann ist V eine Čebyšev-Menge.

BEWEIS. Sei $f \in C(I)$ und $u^* \in \mathbb{P}_n$ die b.A. an f in \mathbb{P}_n . Ist $u^* \in V$, so ist u^* offensichtlich auch die einzige b.A. an f in V . Wir nehmen daher an, dass u^* nicht in V ist, d.h. $a_n < 0$. Sei $u \in V$ beliebig. Für $0 < \theta < 1$ folgt

$$\|f - (1 - \theta)u - \theta u^*\| \leq (1 - \theta)\|f - u\| + \theta \underbrace{\|f - u^*\|}_{< \|f - u\|} < \|f - u\|.$$

Wenn u einen positiven Höchstkoeffizienten a_n hat, gilt gleiches für $(1 - \theta)u + \theta u^*$ mit hinreichend kleinem θ . Also kann eine b.A. an f nur auf dem Rand von V liegen. Da $\partial V = \mathbb{P}_{n-1}$ ein Haarscher Raum ist, folgt die Behauptung. \square

II.4. Strenge Eindeutigkeit

In diesem Abschnitt geht es um die Frage, wie sich die Güte der Approximation verschlechtert, wenn man von der b.A. weggeht. Das folgende Beispiel zeigt, dass die Čebyšev-Approximation eine Sonderrolle spielt.

BEISPIEL II.4.1. Wir betrachten die Approximation der Funktion $f(t) = t$ auf dem Intervall $[-1, 1]$ durch konstante Funktionen

$$V = \{u_a : u_a(t) = a \forall t, a \in \mathbb{R}\}$$

bzgl. der Maximum- und bzgl. der L^2 -Norm.

In der Maximumnorm erhalten wir

$$\begin{aligned} \Phi(a) &= \|f - u_a\|_\infty \\ &= \max_{-1 \leq t \leq 1} |t - a| \\ &= \max\{|1 - a|, |-1 - a|\} \\ &= 1 + |a|. \end{aligned}$$

Das Minimum ist $a^* = 0$. Die Funktion Φ ist in a^* nicht differenzierbar und wächst linear mit dem Abstand vom Minimum.

In der L^2 -Norm erhalten wir

$$\begin{aligned} \Phi(a) &= \|f - u_a\|_2 \\ &= \left\{ \int_{-1}^1 |t - a|^2 dt \right\}^{\frac{1}{2}} \\ &= \left\{ \frac{1}{3} [(1 - a)^3 + (1 + a)^3] \right\}^{\frac{1}{2}} \\ &= \left\{ \frac{1}{3} [2 + 6a^2] \right\}^{\frac{1}{2}}. \end{aligned}$$

Das Minimum ist wieder $a^* = 0$, aber die Funktion Φ ist in a^* differenzierbar und wächst quadratisch mit dem Abstand vom Minimum.

DEFINITION II.4.2. Sei V eine nicht leere Teilmenge des normierten Raumes X . Dann heißt $u^* \in V$ eine *streng eindeutige beste Approximation* zu $f \in X$, wenn es ein $c > 0$ gibt mit

$$\|f - u\| \geq \|f - u^*\| + c \|u^* - u\|$$

für alle $u \in V$.

SATZ II.4.3. *Bei der Čebyšev-Approximation in einem reellen Haarschen Raum ist die b.A. stets eine streng eindeutige beste Approximation.*

BEWEIS. Seien $f \in X$ und $u^* \in V$ die b.A. an f in V . O.E. ist $\|f - u^*\| > 0$. Sei t_0, \dots, t_n eine Referenz in $M[f - u^*]$. Für alle von Null verschiedenen $u \in V$ ist dann

$$\psi(u) = \min_{0 \leq i \leq n} [f(t_i) - u^*(t_i)]u(t_i) < 0.$$

Denn wegen der Haarschen Bedingung sind nicht alle $n + 1$ Terme gleich 0 und aus der Relation $\sum \theta_i [f(t_i) - u^*(t_i)]u(t_i) = 0$ folgt, dass mindestens ein Term negativ ist. Da wegen der endlichen Dimension

von V die Einheitskugel in V kompakt und die Abbildung ψ stetig ist, folgt

$$\max_{u \in V, \|u\|=1} \psi(u) = -c \|f - u^*\| < 0,$$

wobei die Konstante c gerade passend normiert ist. Weiter ist die Funktion ψ homogen. Daher gilt für alle von Null verschiedenen $u \in V$

$$\psi(u) = \|u\| \psi\left(\frac{u}{\|u\|}\right) \leq -c \|f - u^*\| \|u\|.$$

Dies bedeutet

$$\min_{0 \leq i \leq n} [f(t_i) - u^*(t_i)]u(t_i) \leq -c \|f - u^*\| \|u\|$$

bzw.

$$\max_{0 \leq i \leq n} \{-[f(t_i) - u^*(t_i)]u(t_i)\} \geq c \|f - u^*\| \|u\|.$$

Aus dieser Abschätzung folgt wegen $\|f - u^*\| = [f(t_i) - u^*(t_i)]$

$$\begin{aligned} & \|f - u^* - u\| \\ & \geq \max_{0 \leq i \leq n} |f(t_i) - u^*(t_i) - u(t_i)| \\ & \geq \frac{1}{\|f - u^*\|} \max_{0 \leq i \leq n} \{[f(t_i) - u^*(t_i)]^2 - [f(t_i) - u^*(t_i)]u(t_i)\} \\ & \geq \frac{1}{\|f - u^*\|} \{ \|f - u^*\|^2 + c \|f - u^*\| \|u\| \} \\ & = \|f - u^*\| + c \|u\|. \end{aligned}$$

Damit ist die Behauptung bewiesen. \square

SATZ II.4.4. *Die metrische Projektion ist auf jedem Haarschen Raum stetig. Insbesondere ist*

$$\|Tf - Tg\| \leq \frac{2}{c} \|f - g\|,$$

wobei c die Konstante aus der strengen Eindeutigkeit ist.

BEWEIS. Aus der Optimalität und der Dreiecksungleichung folgt

$$\begin{aligned} \|f - Tg\| & \leq \|f - g\| + \underbrace{\|g - Tg\|}_{\leq \|g - Tf\|} \\ & \leq \|f - g\| + \|g - Tf\| \\ & \leq 2 \|f - g\| + \|f - Tf\| \\ & = 2 \|f - g\| + \rho_V(f). \end{aligned}$$

Andererseits gilt nach Satz II.4.3

$$\|f - Tg\| \geq \rho_V(f) + c \|Tf - Tg\|.$$

Durch Vergleich dieser beiden Ungleichungen folgt die Behauptung. \square

BEISPIEL II.4.5. Wir wollen zeigen, dass man auf die Voraussetzungen von Satz II.4.3 nicht verzichten kann.

Für die Voraussetzung „reeller Vektorraum“ betrachten wir Beispiel II.4.1 in \mathbb{C} . Wir erhalten

$$\Phi(ia) = |1 - ia| = \sqrt{1 + a^2}$$

und sehen, dass der Fehler quadratisch mit der Entfernung vom Optimum wächst.

Für die Voraussetzung „Haarscher Raum“ betrachten wir die Approximation von $f(t) = t^2 - 1$ auf $[-1, 1]$ bzgl. der Maximumnorm in

$$V = \{u_a : u_a(t) = at, a \in \mathbb{R}\}.$$

Da die Funktion $-1 + t^2 - at$ ihren einzigen kritischen Punkt in $t^* = \frac{a}{2}$ hat, erhalten wir für $|a| \leq 2$

$$\begin{aligned} \Phi(a) &= \|f - u_a\| = \max \left\{ |a|, 1 + \frac{a^2}{4} \right\} \\ &= 1 + \frac{a^2}{4}. \end{aligned}$$

Dies zeigt, dass $a^* = 0$ das Optimum ist und dass der Fehler quadratisch mit der Entfernung zum Optimum wächst.

II.5. Satz von Mairhuber-Curtis

In diesem Abschnitt zeigen wir, dass Haar Systeme nur auf Intervallen möglich sind.

DEFINITION II.5.1. Die Menge A heißt *einbettbar* in die Menge B , wenn A homöomorph zu einer Teilmenge von B ist, d.h., wenn es eine injektive stetige Abbildung $\varphi : A \rightarrow B$ mit stetiger Umkehrabbildung gibt.

LEMMA II.5.2. *Sei T kompakt und für jede nicht leere, offene Teilmenge U von T sei $T \setminus U$ in S^1 einbettbar. Dann ist T selbst in $S^1 \cup \{t_0\}$ mit $t_0 \notin S^1$ einbettbar.*

BEWEIS. *Fall 1: T ist nicht zusammenhängend:* In diesem Fall gibt es nicht leere, offene, disjunkte Teilmengen A, B von T mit $T = A \cup B$. Jede dieser Mengen für sich ist nach Voraussetzung in S^1 einbettbar. Sind A und B homöomorph zu echten Teilmengen von S^1 , so auch zu Viertelkreisen, und es folgt die Behauptung. Ist dagegen eine dieser Mengen homöomorph zu ganz S^1 , können wir o.E. annehmen, dass dies die Menge A ist. Falls B nur aus einem Punkt besteht, folgt sofort die Behauptung. Andernfalls enthält B einen Punkt t_1 und eine dazu disjunkte, offene Menge U . Die Einbettung von $A \cup (B \setminus U)$ in S^1 liefert dann aber einen Widerspruch.

Fall 2: T ist zusammenhängend: O.E. besteht T aus mehr als einem Punkt. Wir wählen zwei verschiedene Punkte t_1 und t_2 in T . Zu diesen

gibt es offene, disjunkte Umgebungen U_1 und U_2 . Sei $\varphi_1 : (T \setminus U_2) \rightarrow S^1$ eine Einbettung. Da T zusammenhängend ist, enthält $\varphi_1(U_1)$ keinen isolierten Punkt, also wenigstens ein offenes Intervall. Auf diesem seien a, b, c, d geordnete Punkte und $I = \varphi_1^{-1}((b, c))$. Sei $\varphi_2 : (T \setminus I) \rightarrow S^1$ eine Einbettung. Seien a', b', c', d' die Bilder von a, b, c, d unter $\varphi_2 \circ \varphi_1^{-1}$. Weiter bilde ψ das Intervall $[a, d]$ stetig und injektiv auf die Sehne $[a', d']$ ab. Definiere die Abbildung $\varphi : T \rightarrow S^1 \cup [a', d']$ durch

$$\varphi(t) = \begin{cases} \psi \circ \varphi_1(t) & \text{falls } t \in U_1 \text{ und } \varphi_1(t) \in [a, d], \\ \varphi_2(t) & \text{sonst.} \end{cases}$$

Konstruktionsgemäß spart φ die Bögen (a', b') und (c', d') auf S^1 aus¹. Da T zusammenhängend ist, wird auch der Bogen (b', c') ausgespart. Also ist T homöomorph zu einer Figur, die selber homöomorph zu einem Teil von S^1 ist (vgl. Abbildung II.5.1). \square

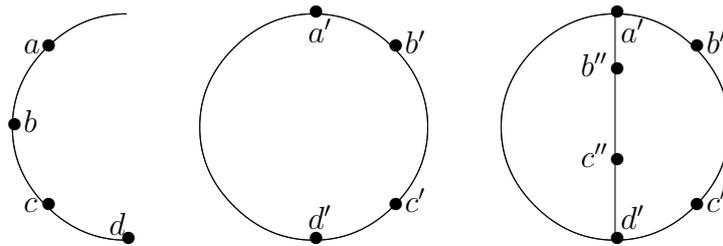


ABBILDUNG II.5.1. Abbildungen φ_1 , φ_2 und φ

LEMMA II.5.3. (1) Auf S^1 gibt es keinen reellen Haarschen Raum gerader Dimension.

(2) Auf $S^1 \cup \{t_0\}$ mit $t_0 \notin S^1$ gibt es keinen reellen Haarschen Raum der Dimension $n \geq 2$.

BEWEIS. ad (1): Sei $V \subset C(S^1)$ ein reeller Unterraum und $\dim V = n$ gerade. Wähle n Punkte t_1, \dots, t_n aus S^1 in zyklischer Anordnung. Falls V die Haarsche Bedingung erfüllt, gibt es ein $u \in V$ mit $u(t_i) = (-1)^i$ für alle i . Dann hat u aber n Nullstellen. Dies ist ein Widerspruch. ad (2): Sei $V \subset C(S^1 \cup \{t_0\})$ ein reeller Haarscher Raum der Dimension $n \geq 2$. Durch Einschränkung der Funktionen auf S^1 erhalten wir einen Haarschen Raum auf S^1 . Also ist n ungerade. Wähle $n - 1$ Punkte t_1, \dots, t_{n-1} aus S^1 in zyklischer Anordnung. Da V ein Haarscher Raum ist, gibt es ein $u \in V$ mit $u(t_0) = 0$ und $u(t_i) = (-1)^i$ für $1 \leq i \leq n - 1$. Diese Funktion hat n Nullstellen. Dies ist erneut ein Widerspruch. \square

¹Angenommen $x \in (a', b')$ liegt nahe bei a' und ist ein Bildpunkt von φ . Dann ist $x = \varphi_2(t)$ für ein t . Dann ist aber $t \in \varphi_1^{-1}([a, d])$ und daher $x = \psi \circ \varphi_1(t)$. Dies ist ein Widerspruch. Analog geht man einen Punkt in (c', d') nahe bei d' vor.

SATZ II.5.4 (Mairhuber-Curtis). *Auf der kompakten Menge T gebe es einen reellen Haarschen Unterraum von $C(T)$ der Dimension $n \geq 2$. Dann ist T einbettbar in S^1 . Das Bild von T unter dieser Einbettung ist eine echte Teilmenge von S^1 , falls die Dimension n gerade ist.*

BEWEIS. Wir beweisen die Behauptung durch Induktion über die Dimension n .

$n = 2$: Sei u_1, u_2 eine Basis von V . Die Abbildung

$$T \rightarrow S^1$$

$$t \mapsto \frac{(u_1(t), u_2(t))}{[u_1(t)^2 + u_2(t)^2]^{\frac{1}{2}}}$$

ist wegen $\det(u_k(t_i)) \neq 0$ stetig und injektiv. Da T kompakt ist, ist die Umkehrabbildung auch stetig.

$n \rightarrow n + 1$: Sei V ein $(n + 1)$ -dimensionaler Haarscher Raum auf T . Wähle eine offene Teilmenge U von T und ein $t^* \in U$. Setze $V_1 = \{u \in V : u(t^*) = 0\}$. Dann ist V_1 ein n -dimensionaler Haarscher Raum auf $T \setminus U$. Nach Induktionsannahme ist $T \setminus U$ in S^1 einbettbar. Wegen Lemma II.5.2 ist T in $S^1 \cup \{t_0\}$ mit $t_0 \notin S^1$ einbettbar. Wegen Lemma II.5.3 ist das Bild aber nicht ganz $S^1 \cup \{t_0\}$. Also ist T sogar in S^1 einbettbar. \square

KAPITEL III

Čebyšev-Approximation durch Polynome

In diesem Kapitel betrachten wir Approximationsprobleme, deren Lösung explizit bekannt ist. Diese Kenntnis ist wichtig für die asymptotischen Aussagen des folgenden Kapitels.

Im Folgenden bezeichnen wir mit \mathbb{P}_n den Raum der algebraischen Polynome vom Grad $\leq n$ und mit ∂p den Grad eines algebraischen Polynomes. Aus historischen Gründen wird der Abstand der b.A. an f in \mathbb{P}_n auf dem Intervall $[a, b]$ mit $E_{n,[a,b]}(f)$ bezeichnet. Ist speziell $[a, b] = [-1, 1]$, wird der Index $[a, b]$ fortgelassen.

III.1. Čebyšev-Polynome

Wir beginnen mit zwei Ergebnissen, die die Bedeutung des Intervalles $[-1, 1]$ unterstreichen.

SATZ III.1.1. *Seien $a < b$ und*

$$\begin{aligned}\psi : [a, b] &\rightarrow [-1, 1] \\ x &\mapsto -1 + 2\frac{x-a}{b-a}\end{aligned}$$

sowie $f \in C([a, b])$ und $g = f \circ \psi^{-1} \in C([-1, 1])$. Dann ist

$$E_{n,[a,b]}(f) = E_{n,[-1,1]}(g)$$

und $p \in \mathbb{P}_n$ ist genau dann b.A. an f , wenn $q = p \circ \psi^{-1} \in \mathbb{P}_n$ b.A. an g ist.

BEWEIS. Die Transformationen

$$p \rightarrow p \circ \psi^{-1}, \quad q \rightarrow q \circ \psi$$

lassen den Raum \mathbb{P}_n invariant. Außerdem ist

$$\begin{aligned}\|f - p\|_{C([a,b])} &= \max_{a \leq x \leq b} |f(x) - p(x)| \\ &= \max_{-1 \leq t \leq 1} |f \circ \psi^{-1}(t) - p \circ \psi^{-1}(t)| \\ &= \|f \circ \psi^{-1} - p \circ \psi^{-1}\|_{C([-1,1])}\end{aligned}$$

und

$$\begin{aligned}\|g - q\|_{C([-1,1])} &= \max_{-1 \leq t \leq 1} |g(t) - q(t)| \\ &= \max_{a \leq x \leq b} |g \circ \psi(x) - q \circ \psi(x)| \\ &= \|g \circ \psi - q \circ \psi\|_{C([a,b])}.\end{aligned}$$

Hieraus folgt die Behauptung. \square

SATZ III.1.2 (Invarianzprinzip). Sei $f \in C([-1, 1])$ gerade bzw. ungerade. Dann ist die b.A. an f in \mathbb{P}_n ebenfalls gerade bzw. ungerade.

BEWEIS. Nach Voraussetzung ist $f(-x) = \pm f(x)$. Sei $p \in \mathbb{P}_n$ die b.A. an f in \mathbb{P}_n und $q(x) = \pm p(-x)$. Dann ist $q \in \mathbb{P}_n$ und

$$\begin{aligned} \max_{-1 \leq x \leq 1} |f(x) - q(x)| &= \max_{-1 \leq x \leq 1} |\pm f(-x) - \pm p(-x)| \\ &= \max_{-1 \leq x \leq 1} |f(-x) - p(-x)| \\ &= \max_{-1 \leq t \leq 1} |f(t) - p(t)|. \end{aligned}$$

Wegen der Eindeutigkeit der b.A. ist $q = p$, d.h. $p(x) = \pm p(-x)$. \square

SATZ III.1.3 (Čebyšev-Polynome). Durch

$$T_n(x) = \cos(n \arccos x) \quad \forall x \in [-1, 1]$$

ist für jedes $n \in \mathbb{N}$ ein Polynom vom Grade n definiert. Die T_n heißen Čebyšev-Polynome (erster Art). Für $n \geq 1$ ist

$$T_n(x) = 2^{n-1}x^n + p \quad \text{mit } p \in \mathbb{P}_{n-2}.$$

Die Čebyšev-Polynome genügen der Rekursionsformel

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_{n+1}(x) &= 2xT_n(x) - T_{n-1}(x), \quad n \geq 1, \end{aligned}$$

und der Symmetrierelation

$$T_n(-x) = (-1)^n T_n(x).$$

BEWEIS. Aus den Additionstheoremen für den Cosinus folgt

$$\cos(a \pm b) = \cos a \cos b \mp \sin a \sin b$$

und somit

$$\cos(a + b) + \cos(a - b) = 2 \cos a \cos b.$$

Setzen wir speziell $a = n\varphi$ und $b = \varphi$, erhalten wir

$$\cos((n+1)\varphi) + \cos((n-1)\varphi) = 2 \cos(n\varphi) \cos \varphi.$$

Mit $\varphi = \arccos x$ folgt hieraus die Rekursionsformel für T_{n-1} . Die Formeln für T_0 und T_1 folgen direkt aus der Definition. Aus der Rekursionsformel folgen die restlichen Eigenschaften durch Induktion über den Grad n . \square

BEMERKUNG III.1.4. (1) Für alle $x \in \mathbb{R}$ gilt

$$T_n(x) = \frac{1}{2} \left\{ \left(x + \sqrt{x^2 - 1} \right)^n + \left(x - \sqrt{x^2 - 1} \right)^n \right\}.$$

(2) Die Čebyšev-Polynome genügen der Orthogonalitätsbeziehung

$$\frac{2}{\pi} \int_{-1}^1 \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}} dx = \begin{cases} 2 & \text{für } n = m = 0, \\ 1 & \text{für } n = m \neq 0, \\ 0 & \text{für } n \neq m. \end{cases}$$

(3) Die Polynome

$$U_n(x) = \frac{1}{n+1} T'_{n+1}(x) = \frac{\sin((n+1) \arccos x)}{\sin(\arccos x)}$$

heißen *Čebyšev-Polynome zweiter Art*. Sie spielen bei der L^1 -Approximation eine ähnliche Rolle wie die Čebyšev-Polynome erster Art bei der Čebyšev-Approximation.

BEWEIS. *ad* (1): Setze

$$Q_n(x) = \frac{1}{2} \left\{ \left(x + \sqrt{x^2 - 1} \right)^n + \left(x - \sqrt{x^2 - 1} \right)^n \right\}.$$

Aus der Binomischen Formel folgt $Q_n \in \mathbb{P}_n$ für alle n . Für $x \in [-1, 1]$ sei $\varphi = \arccos x$. Dann folgt

$$\begin{aligned} Q_n(x) &= \frac{1}{2} \{ (\cos \varphi + i \sin \varphi)^n + (\cos \varphi - i \sin \varphi)^n \} \\ &= \frac{1}{2} \{ e^{in\varphi} + e^{-in\varphi} \} \\ &= \cos(n\varphi) \\ &= T_n(x). \end{aligned}$$

Hieraus folgt die Behauptung.

ad (2): Mit der Transformation $x = \cos \varphi$ folgt wegen $\frac{dx}{d\varphi} = -\sin \varphi = \sqrt{1-x^2}$

$$\begin{aligned} \frac{2}{\pi} \int_{-1}^1 \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}} dx &= \frac{2}{\pi} \int_0^\pi \cos(n\varphi) \cos(m\varphi) d\varphi \\ &= \frac{1}{\pi} \int_0^\pi \{ \cos((n+m)\varphi) + \cos((n-m)\varphi) \} d\varphi \\ &= \frac{1}{2\pi} \int_{-\pi}^\pi \{ \cos((n+m)\varphi) + \cos((n-m)\varphi) \} d\varphi. \end{aligned}$$

Wegen

$$\int_{-\pi}^\pi \cos(k\varphi) d\varphi = 0 \quad \forall k \in \mathbb{Z} \setminus \{0\}$$

folgt hieraus die Behauptung.

ad (3): Ist offensichtlich. \square

SATZ III.1.5. Sei $n \geq 1$ und $p \in \mathbb{P}_{n-1}$ die b.A. an x^n auf $[-1, 1]$ in \mathbb{P}_{n-1} . Dann ist

$$x^n - p(x) = 2^{-n+1} T_n(x)$$

und

$$E_{n-1}(x^n) = 2^{-n+1}.$$

BEWEIS. Gemäß Satz III.1.3 ist $x^n - 2^{-n+1}T_n(x) \in \mathbb{P}_{n-1}$. Für $x_{n,k} = \cos(\frac{k\pi}{n})$, $0 \leq k \leq n$, gilt außerdem

$$T_n(x_{n,k}) = \cos(k\pi) = (-1)^k = (-1)^k \|T_n\|.$$

Also bilden die Punkte $x_{n,0}, \dots, x_{n,n}$ eine Alternante, und die Behauptung folgt aus dem Alternantenkriterium Satz II.3.8 (S. 32). \square

KOROLLAR III.1.6. *Es ist*

$$E_{n-1,[a,b]}(x^n) = \frac{(b-a)^n}{2^{2n-1}}.$$

BEWEIS. Wir wenden Satz III.1.1 mit $f(x) = x^n$ an und erhalten

$$\begin{aligned} g(t) &= f(\psi^{-1}(t)) \\ &= f\left(a + \frac{1}{2}(t+1)(b-a)\right) \\ &= \left[a + \frac{1}{2}(t+1)(b-a)\right]^n \\ &= \left(\frac{b-a}{2}\right)^n t^n + r(t) \end{aligned}$$

mit $r \in \mathbb{P}_{n-1}$. Also ist die b.A. an g in \mathbb{P}_{n-1} von der Form $\left(\frac{b-a}{2}\right)^n q(t) + r(t)$, wenn q die b.A. an t^n in \mathbb{P}_{n-1} ist. Insbesondere folgt

$$\begin{aligned} E_{n-1}(g) &= \left\| \left(\frac{b-a}{2}\right)^n [t^n - q] \right\| \\ &= \left(\frac{b-a}{2}\right)^n E_{n-1}(t^n). \end{aligned}$$

Damit folgt die Behauptung aus Satz III.1.5 und Satz III.1.1. \square

III.2. Approximation von $(x-a)^{-1}$

Während die Approximation einer höheren Potenz charakteristisch ist für die Approximation analytischer Funktionen, ist die Approximation von $(x-a)^{-1}$ typisch für die Approximation von Funktionen mit Singularitäten.

SATZ III.2.1. *Für $a > 1$ ist*

$$E_n((x-a)^{-1}) = \frac{(a - \sqrt{a^2 - 1})^n}{a^2 - 1}.$$

BEWEIS. Setze $\alpha = a - \sqrt{a^2 - 1}$. Dann ist

$$\begin{aligned} \frac{1}{2} \left(\alpha + \frac{1}{\alpha} \right) &= \frac{1}{2} \left(a - \sqrt{a^2 - 1} + \frac{1}{a - \sqrt{a^2 - 1}} \right) \\ &= \frac{(a - \sqrt{a^2 - 1})^2 + 1}{2(a - \sqrt{a^2 - 1})} \\ &= \frac{a^2 - a\sqrt{a^2 - 1}}{a - \sqrt{a^2 - 1}} \\ &= a. \end{aligned}$$

Für $x \in [-1, 1]$ definieren wir $v \in \mathbb{C}$ mit $|v| = 1$ durch $v = e^{i\varphi}$ mit $\varphi = \arccos x$. Dann ist

$$\frac{1}{2} \left(v + \frac{1}{v} \right) = \frac{1}{2} (e^{i\varphi} + e^{-i\varphi}) = \cos \varphi = x.$$

Setze zur Abkürzung

$$M = \frac{(a - \sqrt{a^2 - 1})^n}{a^2 - 1}.$$

Wegen

$$\begin{aligned} a^2 - 1 &= \frac{1}{4} \left(\alpha + \frac{1}{\alpha} \right)^2 - 1 \\ &= \frac{1}{4} \left(\alpha^2 + 2 + \frac{1}{\alpha^2} \right) - 1 \\ &= \frac{1}{4} \left(\alpha - \frac{1}{\alpha} \right)^2 \\ &= \frac{(1 - \alpha^2)^2}{4\alpha^2} \end{aligned}$$

ist

$$M = \frac{4\alpha^{n+2}}{(1 - \alpha^2)^2}.$$

Schließlich definieren wir zwei Funktionen g und p durch

$$g(v) = \frac{1}{2} \left(v^n \frac{\alpha - v}{1 - \alpha v} + v^{-n} \frac{1 - \alpha v}{\alpha - v} \right)$$

und

$$p(x) = \frac{1}{x - a} - Mg(v).$$

1. Schritt: Wir zeigen $p \in \mathbb{P}_n$. Setze

$$\Phi(v) = 1 - (x - a)p(x) = M(x - a)g(v).$$

Es genügt dann zu zeigen, dass $\Phi \in \mathbb{P}_{n+1}$ ist und an der Stelle $x = a$ entsprechend $v = \alpha$ den Wert 1 annimmt. Wegen

$$T_n(x) = \cos(n\varphi) = \frac{1}{2} (e^{in\varphi} + e^{-in\varphi}) = \frac{1}{2} (v^n + v^{-n})$$

erhalten wir

$$\begin{aligned} \Phi(v) &= \frac{M}{2} \left(v + \frac{1}{v} - \alpha - \frac{1}{\alpha} \right) g(v) \\ &= \frac{M}{4} \frac{\alpha v^2 + \alpha - \alpha^2 v - v}{\alpha v} \left(v^n \frac{\alpha - v}{1 - \alpha v} + v^{-n} \frac{1 - \alpha v}{\alpha - v} \right) \\ &= \frac{M}{4} \frac{(\alpha - v)(1 - \alpha v)}{\alpha v} \left(v^n \frac{\alpha - v}{1 - \alpha v} + v^{-n} \frac{1 - \alpha v}{\alpha - v} \right) \\ &\stackrel{(*)}{=} \frac{M}{4\alpha} \left((\alpha - v)^2 v^{n-1} + v^{-n-1} (1 - \alpha v)^2 \right) \\ &= \frac{M}{4\alpha} \left(\alpha^2 v^{n-1} - 2\alpha v^n + v^{n+1} + \alpha^2 v^{-n+1} - 2\alpha v^{-n} + v^{-n-1} \right) \\ &= \frac{M}{4\alpha} \left(\alpha^2 [v^{n-1} + v^{-n+1}] - 2\alpha [v^n + v^{-n}] + [v^{n+1} + v^{-n-1}] \right) \\ &= \frac{M}{2\alpha} \left(\alpha^2 T_{n-1}(x) - 2\alpha T_n(x) + T_{n+1}(x) \right). \end{aligned}$$

Also ist $\Phi \in \mathbb{P}_{n+1}$. Einsetzen von $v = \alpha$ an der Stelle $(*)$ liefert

$$\begin{aligned} \Phi(\alpha) &= \frac{M}{4\alpha} \alpha^{-n-1} (1 - \alpha^2)^2 = M \frac{(1 - \alpha^2)^2}{4\alpha^{n+2}} \\ &= 1. \end{aligned}$$

Also ist $p \in \mathbb{P}_n$.

2. Schritt: Wir zeigen $\sup_{v \in \mathbb{C}, |v|=1} |g(v)| = 1$. Für $v \in \mathbb{C}$ mit $|v| = 1$ haben wir

$$\begin{aligned} \left| v^n \frac{\alpha - v}{1 - \alpha v} \right| &= \left| v^{n-1} \frac{\alpha - v}{\alpha - \frac{1}{v}} \right| \\ &= \left| \frac{\alpha - v}{\alpha - \frac{1}{v}} \right| \\ &= \left| \frac{\alpha - v}{\alpha - \bar{v}} \right| \\ &= \left| \frac{\alpha - v}{\overline{\alpha - v}} \right| \\ &= 1 \end{aligned}$$

und

$$\left| v^{-n} \frac{1 - \alpha v}{\alpha - v} \right| = \left| v^n \frac{\alpha - v}{1 - \alpha v} \right|^{-1} = 1.$$

Also ist

$$\sup_{v \in \mathbb{C}, |v|=1} |g(v)| \leq 1.$$

Für $x = 1$ bzw. $v = 1$ gilt andererseits

$$g(1) = \frac{1}{2} \left(1 \frac{\alpha - 1}{1 - \alpha} + 1 \frac{1 - \alpha}{\alpha - 1} \right) = -1.$$

Also ist

$$\sup_{v \in \mathbb{C}, |v|=1} |g(v)| = 1.$$

3. Schritt: Wir zeigen, dass g an $n + 2$ Punkten sein Betragsmaximum mit alternierendem Vorzeichen annimmt. Da für $v \in \mathbb{C}$ mit $|v| = 1$ gilt $\frac{1}{v} = \bar{v}$ und da α reell ist, folgt

$$\overline{v^n \frac{\alpha - v}{1 - \alpha v}} = v^{-n} \frac{1 - \alpha v}{\alpha - v}.$$

Also ist

$$g(v) = \operatorname{Re} \left(v^n \frac{\alpha - v}{1 - \alpha v} \right) = \cos \left(\arg \left(v^n \frac{\alpha - v}{1 - \alpha v} \right) \right).$$

Mithin gilt

$$\begin{aligned} \arg \left(v^n \frac{\alpha - v}{1 - \alpha v} \right) &= 0 \pmod{2\pi} \implies g(v) = 1, \\ \arg \left(v^n \frac{\alpha - v}{1 - \alpha v} \right) &= \pi \pmod{2\pi} \implies g(v) = -1. \end{aligned}$$

Die Funktion $v^n(\alpha - v)(1 - \alpha v)^{-1}$ ist wegen $\alpha \in (0, 1)$ im Einheitskreis analytisch und hat dort α als einfache und 0 als n -fache Nullstelle. Nach dem Satz über die Umkehrzahl aus der Funktionentheorie nimmt daher $\arg \left(v^n \frac{\alpha - v}{1 - \alpha v} \right)$ um $(n + 1)2\pi$ zu, wenn man einmal den Einheitskreis im positiven Sinn durchläuft. Aus Symmetriegründen nimmt dieser Wert um $(n + 1)\pi$ zu, wenn man den oberen Halbkreis durchläuft. Wegen

$$\arg \left(1 \frac{\alpha - 1}{1 - \alpha} \right) = \arg(-1) = \pi$$

folgt, dass $(n + 2)$ -mal der Extremalwert mit alternierendem Vorzeichen angenommen wird. \square

SATZ III.2.2. Für $b > 0$ ist

$$E_{2n}((x^2 + b^2)^{-1}) = E_{2n+1}((x^2 + b^2)^{-1}) = \frac{(\sqrt{b^2 + 1} - b)^{2n}}{2b^2(1 + b^2)}.$$

BEWEIS. Da die Funktion $(x^2 + b^2)^{-1}$ gerade ist, ist ihre b.A. gemäß Satz III.1.2 (S. 40) auch gerade. Daher

$$E_{2n}((x^2 + b^2)^{-1}) = E_{2n+1}((x^2 + b^2)^{-1}).$$

Sei $p \in \mathbb{P}_{2n}$ die b.A. an $(x^2 + b^2)^{-1}$ auf $[-1, 1]$. Da p gerade ist, ist es ein Polynom vom Grad $\leq n$ in x^2 . Mit der Transformation $z = x^2$ folgt daher

$$E_{2n}((x^2 + b^2)^{-1}) = E_{n,[0,1]}((z + b^2)^{-1}).$$

Aus Satz III.1.1 (S. 39) erhalten wir mit der Transformation $z = \frac{1}{2}(1-t)$

$$\begin{aligned} E_{n,[0,1]}((z+b^2)^{-1}) &= E_n(-2(t-(2b^2+1))^{-1}) \\ &= 2E_n((t-(2b^2+1))^{-1}). \end{aligned}$$

Für den Parameter α aus dem Beweis von Satz III.2.1 ergibt sich nun

$$\begin{aligned} \alpha &= 2b^2 + 1 - \sqrt{(2b^2 + 1)^2 - 1} \\ &= 2b^2 + 1 - \sqrt{4b^4 + 4b^2} \\ &= 2b^2 + 1 - 2b\sqrt{b^2 + 1} \\ &= \left(\sqrt{b^2 + 1} - b\right)^2. \end{aligned}$$

Insgesamt folgt

$$\begin{aligned} E_{2n}((x^2 + b^2)^{-1}) &= 2E_n((t - (2b^2 + 1))^{-1}) \\ &= \frac{2(\sqrt{b^2 + 1} - b)^{2n}}{(2b^2 + 1)^2 - 1} \\ &= \frac{(\sqrt{b^2 + 1} - b)^{2n}}{2b^2(b^2 + 1)}. \quad \square \end{aligned}$$

III.3. Vergleichssatz von Bernstein

In der Einführung haben wir gesagt, dass man häufig eine „gute“ Approximation zur Abschätzung der Güte der b.A. heranzieht. In diesem Abschnitt wählen wir den umgekehrten Weg. Das Ergebnis ist wichtig für die asymptotischen Aussagen des nächsten Kapitels.

SATZ III.3.1 (Vergleichssatz der Interpolation). *Sei I ein Intervall. Die Funktionen $f, g \in C^{n+1}(I)$ mögen für alle $x \in I$ die Relation*

$$|f^{(n+1)}(x)| \leq g^{(n+1)}(x)$$

erfüllen. Zu den $n+1$ Punkten $x_0 < \dots < x_n \in I$ seien p und q die Interpolationspolynome zu f bzw. g in diesen Punkten. Dann gilt für alle $x \in I$ die Abschätzung

$$|f(x) - p(x)| \leq |g(x) - q(x)|.$$

BEWEIS. Aus der Voraussetzung folgt offensichtlich $g^{(n+1)}(x) \geq 0$ für alle $x \in I$. Wir nehmen zunächst an, dass sogar $g^{(n+1)}(x) > 0$ für alle $x \in I$ gilt.

Angenommen, die Behauptung ist für einen Punkt $z \in I$ nicht erfüllt. Dann gilt für $\lambda = \frac{g(z)-q(z)}{f(z)-p(z)}$ die Abschätzung $|\lambda| < 1$. Außerdem fällt z mit keinem der Punkte x_0, \dots, x_n zusammen. Die Hilfsfunktion $h = g - q - \lambda(f - p)$ hat die $n+2$ Nullstellen x_0, \dots, x_n und z . Nach dem

Satz von Rolle besitzt $h^{(n+1)}$ dann eine Nullstelle in I . Andererseits ist für jedes $x \in I$

$$\begin{aligned} h^{(n+1)}(x) &= g^{(n+1)}(x) - \lambda f^{(n+1)}(x) \\ &\geq g^{(n+1)}(x) - |\lambda| |f^{(n+1)}(x)| \\ &\geq (1 - |\lambda|)g^{(n+1)}(x) \\ &> 0. \end{aligned}$$

Dies ist ein Widerspruch. Daher ist die Aussage in diesem Fall gezeigt. Für den allgemeinen Fall betrachten wir die Funktion $g_\varepsilon = g + \varepsilon x^{n+1}$ mit beliebig kleinem $\varepsilon > 0$. Sei q_ε das entsprechende Interpolationspolynom. Aus dem ersten Teil des Beweises folgt für alle $x \in I$ und alle $\varepsilon > 0$

$$|f(x) - p(x)| \leq |g_\varepsilon(x) - q_\varepsilon(x)|.$$

Da die rechte Seite dieser Ungleichung wegen der Lagrangeschen Interpolationsformel stetig von ε abhängt, folgt die Behauptung. \square

SATZ III.3.2 (Vergleichssatz von Bernstein). *Sei I ein Intervall. Für die Funktionen $f, g \in C^{n+1}(I)$ gelte die Relation*

$$|f^{(n+1)}(x)| \leq g^{(n+1)}(x)$$

für alle $x \in I$. Dann gilt für die Čebyšev-Approximation auf I

$$E_{n,I}(f) \leq E_{n,I}(g).$$

BEWEIS. Sei q die b.A. an g in \mathbb{P}_n . Wegen der Existenz einer Alternante der Länge $n+2$ (vgl. Satz II.3.8 (S. 32) und Beispiel II.3.2(1) (S. 29)) interpoliert q die Funktion g in $n+1$ Punkten. Sei p das Interpolationspolynom von f in diesen Punkten. Dann folgt mit Satz III.3.1

$$E_{n,I}(f) \leq \|f - p\| \leq \|g - q\| = E_{n,I}(g). \quad \square$$

Durch die Wahl passender Vergleichsfunktionen erhalten wir:

SATZ III.3.3. *Sei $f \in C^{n+1}([-1, 1])$. Dann gibt es einen Punkt $\eta \in [-1, 1]$ mit*

$$E_n(f) = \frac{1}{2^n(n+1)!} |f^{(n+1)}(\eta)|.$$

BEWEIS. Sei

$$a = \min_{-1 \leq x \leq 1} f^{(n+1)}(x), \quad b = \max_{-1 \leq x \leq 1} f^{(n+1)}(x).$$

Fall 1: $a < 0 < b$: Setze

$$c = \max\{-a, b\}$$

und

$$g = \frac{c}{(n+1)!} x^{n+1}.$$

Dann folgt für alle $x \in [-1, 1]$

$$|f^{(n+1)}(x)| \leq c = g^{(n+1)}(x).$$

Wegen Satz III.3.2 und Satz III.1.5 (S. 41) folgt

$$E_n(f) \leq \frac{c}{2^n(n+1)!}.$$

Da $|f^{(n+1)}|$ alle Werte zwischen 0 und c annimmt, folgt die Behauptung.

Fall 2: $0 \leq a \leq b$: Wegen der Einschließung

$$\left(\frac{a}{(n+1)!}x^{n+1}\right)^{(n+1)} \leq f^{(n+1)} \leq \left(\frac{b}{(n+1)!}x^{n+1}\right)^{(n+1)}$$

liefern Satz III.3.2 und Satz III.1.5 (S. 41) die Einschließung

$$\frac{a}{2^n(n+1)!} \leq E_n(f) \leq \frac{b}{2^n(n+1)!}.$$

Damit folgt die Behauptung aus dem Zwischenwertsatz.

Fall 3: $a \leq b \leq 0$: Wegen $E_n(f) = E_n(-f)$ kann dieser Fall durch Übergang zu $-f$ auf den zweiten Fall zurückgeführt werden. \square

BEISPIEL III.3.4. (1) Wegen $(e^x)^{(n+1)} = e^x$ und $e^{-1} \leq e^x \leq e$ auf $[-1, 1]$ folgt aus Satz III.3.3

$$E_n(e^x) \in \left[\frac{1}{e2^n(n+1)!}, \frac{e}{2^n(n+1)!} \right].$$

Mit Satz III.1.1 (S. 39) ergibt sich sogar

$$E_{n,[0,1]}(e^x) = E_n(e^{\frac{1+t}{2}}) \in \left[\frac{1}{e2^{2n+1}(n+1)!}, \frac{e}{2^{2n+1}(n+1)!} \right].$$

Man gewinnt also einen Faktor 2^{2n+1} im Vergleich zur Taylorreihe. Einen weiteren Faktor 2^{n+1} kann man gewinnen, wenn man die Beziehung $e^x = (e^{\frac{x}{2}})^2$ ausnutzt und sich auf das Intervall $[0, \frac{1}{2}]$ einschränkt.

(2) Wir wollen $E_{n,[1,2]}(\ln x)$ abschätzen. Wegen

$$(\ln x)^{(n+1)} = -\frac{1}{n+1}x \left(\frac{1}{x}\right)^{(n+1)}$$

ist auf $[1, 2]$

$$\begin{aligned} 0 &\leq \frac{1}{n+1} \left(\frac{(-1)^{n+1}}{x}\right)^{(n+1)} \leq ((-1)^n \ln x)^{(n+1)} \\ &\leq \frac{2}{n+1} \left(\frac{(-1)^{n+1}}{x}\right)^{(n+1)}. \end{aligned}$$

Also ist gemäß Satz III.3.2

$$\frac{1}{n+1} E_{n,[1,2]} \left(\frac{1}{x}\right) \leq E_{n,[1,2]}(\ln x) \leq \frac{2}{n+1} E_{n,[1,2]} \left(\frac{1}{x}\right).$$

Wegen Satz III.1.1 (S. 39) ist

$$\begin{aligned} E_{n,[1,2]} \left(\frac{1}{x} \right) &= E_n \left(\frac{2}{t+3} \right) & x &= 1 + \frac{1}{2}(t+1) \\ &= E_n \left(\frac{2}{3-z} \right) & t &= -z \\ &= 2E_n \left(\frac{1}{3-z} \right). \end{aligned}$$

Aus Satz III.2.1 (S. 42) folgt

$$E_n \left(\frac{1}{3-z} \right) = \frac{(3-\sqrt{8})^n}{8} = \frac{1}{8(3+\sqrt{8})^n}.$$

Insgesamt erhalten wir somit

$$\frac{1}{4(n+1)(3+\sqrt{8})^n} \leq E_{n,[1,2]}(\ln x) \leq \frac{1}{2(n+1)(3+\sqrt{8})^n}.$$

KAPITEL IV

Asymptotische Aussagen

Im Mittelpunkt dieses Kapitels steht die Beziehung zwischen der Glattheit einer Funktion und der Güte ihrer b.A. durch algebraische oder trigonometrische Polynome.

IV.1. Satz von Weierstraß und Lethargietheorem

Wir wollen zeigen, dass jede stetige Funktion beliebig gut durch Polynome approximiert werden kann, dass aber die Konvergenzgeschwindigkeit als Funktion des Polynomgrades beliebig schlecht sein kann.

Im Folgenden ist stets I ein kompaktes, perfektes Intervall.

DEFINITION IV.1.1. (1) Seien I ein Intervall und f, g stetige Funktionen auf I . Wir sagen $f \geq 0$ bzw. $f \geq g$, wenn für alle $x \in I$ gilt $f(x) \geq 0$ bzw. $f(x) \geq g(x)$.

(2) Ein linearer Operator $L : C(I) \rightarrow C(I)$ heißt *positiv*, wenn für alle $f \in C(I)$ mit $f \geq 0$ gilt $Lf \geq 0$.

BEMERKUNG IV.1.2. Seien $L : C(I) \rightarrow C(I)$ ein positiver linearer Operator und $f, g \in C(I)$ mit $f \geq g$. Dann ist $Lf \geq Lg$.

BEWEIS. Die Beziehung $f \geq g$ ist äquivalent zu $f - g \geq 0$. Also folgt aus der Positivität und Linearität von L

$$0 \leq L(f - g) = Lf - Lg \iff Lf \geq Lg. \quad \square$$

SATZ IV.1.3 (Bohmann-Korovkin). Sei $(L_n)_{n \in \mathbb{N}}$ eine Folge positiver linearer Operatoren, die $C(I)$ in sich abbilden. Mit $u_\nu(x) = x^\nu$ gelte für $\nu = 0, 1, 2$ in der Supremumsnorm

$$\lim_{n \rightarrow \infty} \|L_n u_\nu - u_\nu\| = 0.$$

Dann gilt für alle $f \in C(I)$

$$\lim_{n \rightarrow \infty} \|L_n f - f\| = 0.$$

BEWEIS. Seien $f \in C(I)$ und $\varepsilon > 0$ beliebig. Da I kompakt ist, ist f gleichmäßig stetig. Also gibt es ein $\delta = \delta(\varepsilon) > 0$ mit $|f(x) - f(y)| < \varepsilon$ für alle $x, y \in I$ mit $|x - y| < \delta$. Für $x, y \in I$ mit $|x - y| \geq \delta$ gilt andererseits

$$|f(x) - f(y)| \leq 2\|f\| \leq 2\delta^{-2}\|f\| |x - y|^2.$$

Also gilt für alle $x, y \in I$

$$|f(x) - f(y)| \leq \max \{ \varepsilon, 2\delta^{-2} \|f\| |x - y|^2 \} \leq \varepsilon + 2\delta^{-2} \|f\| |x - y|^2$$

d.h.

$$-\varepsilon - 2\delta^{-2} \|f\| |x - y|^2 \leq f(x) - f(y) \leq \varepsilon + 2\delta^{-2} \|f\| |x - y|^2.$$

Wir halten nun $y \in I$ fest und fassen es als Parameter auf. Dann können wir obige Abschätzung schreiben als

$$\begin{aligned} & -\varepsilon u_0 - 2\delta^{-2} \|f\| \{u_2 - 2yu_1 + y^2 u_0\} \\ & \leq f - f(y)u_0 \\ & \leq \varepsilon u_0 + 2\delta^{-2} \|f\| \{u_2 - 2yu_1 + y^2 u_0\} \end{aligned}$$

Aus der Positivität und Linearität der L_n folgt dann

$$\begin{aligned} & -\varepsilon L_n u_0 - 2\delta^{-2} \|f\| \{L_n u_2 - 2yL_n u_1 + y^2 L_n u_0\} \\ & \leq L_n f - f(y)L_n u_0 \\ & \leq \varepsilon L_n u_0 + 2\delta^{-2} \|f\| \{L_n u_2 - 2yL_n u_1 + y^2 L_n u_0\} \end{aligned}$$

und somit

$$\begin{aligned} & |L_n f(y) - f(y)L_n u_0(y)| \\ & \leq \varepsilon L_n u_0(y) + 2\delta^{-2} \|f\| \{L_n u_2(y) - 2yL_n u_1(y) + y^2 L_n u_0(y)\} \\ & = \varepsilon L_n u_0(y) \\ & \quad + 2\delta^{-2} \|f\| \left\{ L_n u_2(y) - \underbrace{u_2(y)}_{=y^2} - 2y \left[L_n u_1(y) - \underbrace{u_1(y)}_{=y} \right] \right. \\ & \quad \left. + y^2 \left[L_n u_0(y) - \underbrace{u_0(y)}_{=1} \right] \right\} \\ & \leq \varepsilon \|L_n u_0\| \\ & \quad + 2\delta^{-2} \|f\| \{ \|L_n u_2 - u_2\| + 2 \|u_1\| \|L_n u_1 - u_1\| \\ & \quad + \|u_2\| \|L_n u_0 - u_0\| \}. \end{aligned}$$

Mittels Dreiecksungleichung und Übergang zur Norm folgt hieraus

$$\begin{aligned} & \|L_n f - f\| \\ & = \max_{y \in I} |L_n f(y) - f(y)| \\ & \leq \max_{y \in I} \{ |L_n f(y) - f(y)L_n u_0(y)| + |f(y)[u_0(y) - L_n u_0(y)]| \} \\ & \leq \varepsilon \|L_n u_0\| \\ & \quad + 2\delta^{-2} \|f\| \{ \|L_n u_2 - u_2\| + 2 \|u_1\| \|L_n u_1 - u_1\| \\ & \quad + \|u_2\| \|L_n u_0 - u_0\| \} \\ & \quad + \|f\| \|L_n u_0 - u_0\|. \end{aligned}$$

Nach Voraussetzung gibt es ein $N = N(\varepsilon)$, so dass die rechte Seite dieser Abschätzung für alle $n \geq N$ kleiner ist als 4ε ist. Da ε und f beliebig waren, ist damit die Behauptung bewiesen. \square

DEFINITION IV.1.4. Sei $f \in C([0, 1])$. Dann definieren wir die *Bernstein-Polynome* $B_n f \in \mathbb{P}_n$ durch

$$B_0 f(x) = 1 \quad \forall x \in [0, 1],$$

$$B_n f(x) = \sum_{k=0}^n \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k} \quad \forall x \in [0, 1], n \in \mathbb{N}^*.$$

SATZ IV.1.5. Für alle $f \in C([0, 1])$ gilt

$$\lim_{n \rightarrow \infty} \|B_n f - f\| = 0.$$

BEWEIS. Offensichtlich ist jedes B_n ein linearer Operator von $C([0, 1])$ in sich. Da die Funktionen $\binom{n}{k} x^k (1-x)^{n-k}$ für alle n und k positiv sind, sind die B_n auch positiv. Aus der Binomischen Formel folgt außerdem für alle $n \geq 2$ und alle $x \in [0, 1]$

$$\begin{aligned} B_n u_0(x) &= \sum_{k=0}^n \binom{n}{k} x^k (1-x)^{n-k} = (x + 1 - x)^n \\ &= 1, \\ B_n u_1(x) &= \sum_{k=0}^n \binom{n}{k} \frac{k}{n} x^k (1-x)^{n-k} \\ &= x \sum_{k=1}^n \binom{n-1}{k-1} x^{k-1} (1-x)^{n-k} \\ &= x \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} x^\ell (1-x)^{n-1-\ell} \\ &= x (x + 1 - x)^{n-1} \\ &= x, \\ B_n u_2(x) &= \sum_{k=0}^n \binom{n}{k} \left(\frac{k}{n}\right)^2 x^k (1-x)^{n-k} \\ &= \sum_{k=0}^n \binom{n}{k} \left(\frac{k(k-1)}{n^2} + \frac{k}{n^2}\right) x^k (1-x)^{n-k} \\ &= \frac{n-1}{n} x^2 \sum_{k=2}^n \binom{n-2}{k-2} x^{k-2} (1-x)^{n-k} + \frac{1}{n} B_n u_1(x) \\ &= \frac{n-1}{n} x^2 + \frac{1}{n} x \\ &= x^2 + \frac{1}{n} x(1-x). \end{aligned}$$

Damit folgt die Behauptung aus Satz IV.1.3. \square

KOROLLAR IV.1.6 (Satz von Weierstraß). *Für jedes kompakte, perfekte Intervall I und jedes $f \in C(I)$ gilt $\lim_{n \rightarrow \infty} E_{n,I}(f) = 0$.*

BEWEIS. Die Behauptung folgt unmittelbar aus Satz IV.1.5, da jedes kompakte, perfekte Intervall durch eine affine Transformation homöomorph auf das Intervall $[0, 1]$ abgebildet werden kann. \square

BEMERKUNG IV.1.7. Sei $f \in C([0, 1])$ eine nicht-affine Funktion. Dann gibt es ein $\gamma > 0$ mit $\|B_n f - f\| \geq \frac{\gamma}{n}$. Dieses Phänomen nennt man *Saturation*. Die Funktionen, die mit $\|B_n f - f\| = O(\frac{1}{n})$ die optimale Konvergenzordnung erreichen, bezeichnet man als *Saturationsklasse*. Die affinen Funktionen, die durch $\|B_n f - f\| = o(\frac{1}{n})$ gekennzeichnet sind, bezeichnet man als *Saturationskern*.

BEWEIS. Sei $f \in C([0, 1])$ eine nicht-affine Funktion. Dann gibt es ein $t_0 \in (0, 1)$, so dass f in den Punkten $0, t_0$ und 1 durch eine Parabel $q_0(x) = ax^2 + bx + c$ mit $a \neq 0$ interpoliert wird. O.E. ist $a < 0$. Setze

$$q(x) = q_0(x) + \max_{0 \leq z \leq 1} \{f(z) - q_0(z)\}.$$

Dann ist $q(x) \geq f(x)$ für alle $x \in [0, 1]$ und $q(t) = f(t)$ für ein $t \in (0, 1)$. Aus dem Beweis von Satz IV.1.5 wissen wir, dass

$$B_n q(x) = q(x) + \frac{a}{n} x(1-x)$$

ist für alle $x \in [0, 1]$. Damit folgt für obiges t wegen der Positivität der Operatoren B_n

$$B_n f(t) \leq B_n q(t) = q(t) + \frac{a}{n} t(1-t) = f(t) + \frac{a}{n} t(1-t).$$

Also ist

$$\begin{aligned} \|f - B_n f\| &\geq |f(t) - B_n f(t)| \geq f(t) - B_n f(t) \geq -\frac{a}{n} t(1-t) \\ &= \frac{|a|}{n} t(1-t). \end{aligned}$$

Hieraus folgt die Behauptung. \square

Gemäß Korollar IV.1.6 können stetige Funktionen beliebig gut durch Polynome approximiert werden. Die Konvergenzgeschwindigkeit dieser Approximation kann aber beliebig schlecht sein, d.h., zu jeder Nullfolge $(\varepsilon_n)_{n \in \mathbb{N}}$ gibt es ein $f \in C(I)$ mit $E_{n,I}(f) \geq \varepsilon_n$ für alle n . Wir beweisen dies in einem allgemeineren Rahmen.

SATZ IV.1.8. *Sei V ein endlich dimensionaler Unterraum eines normierten Raumes X und $V \neq X$. Dann gibt es ein $f \in X$ mit $\rho_V(f) = \|f\| = 1$.*

BEWEIS. Sei $f_0 \in X \setminus V$ und $u_0 \in V$ eine b.A. an f in V . Dann ist

$$\rho_V(f_0 - u_0) = \rho_V(f_0) = \|f_0 - u_0\| > 0.$$

Daher leistet $f = \|f_0 - u_0\|^{-1}(f_0 - u_0)$ das Gewünschte. \square

SATZ IV.1.9 (Lethargietheorem). Sei V_1, V_2, \dots eine Folge endlich dimensionaler Unterräume eines normierten Raumes X . Für alle $n \geq 1$ sei V_n ein echter Unterraum von V_{n+1} . Dann gibt es zu jeder monotonen Nullfolge $(\varepsilon_n)_{n \in \mathbb{N}^*}$ in \mathbb{R}_+^* ein Element f in X mit $\rho_{V_n}(f) \geq \varepsilon_n$ für alle $n \in \mathbb{N}^*$.

BEWEIS. O.E. ist $\varepsilon_1 \leq \frac{1}{18}$, da jeder Fall wegen $\rho_V(\lambda f) = |\lambda| \rho_V(f)$ auf diesen Fall zurückgeführt werden kann.

Wähle eine streng monoton wachsende Folge $(n_k)_{k \in \mathbb{N}^*}$ in \mathbb{N}^* mit $n_1 = 1$ und $\varepsilon_{n_k} \leq \frac{1}{2}3^{-k-1}$ für alle k . Setze $W_k = V_{n_k}$. Gemäß Satz IV.1.8 gibt es zu jedem k ein $f_k \in W_{k+1}$ mit $\rho_{W_k}(f_k) = \|f_k\| = 3^{-k}$. Dann ist $f = \sum_{k=1}^{\infty} f_k \in X$ wohldefiniert und erfüllt für alle k und alle $n > n_k$ die Abschätzung

$$\begin{aligned} \rho_{W_k}(f) &\geq \rho_{W_k} \left(\sum_{j=1}^k f_j \right) - \left\| \sum_{j=k+1}^{\infty} f_j \right\| \\ &\geq \rho_{W_k} \left(\underbrace{f_k + \sum_{j=1}^{k-1} f_j}_{\in W_k} \right) - \sum_{j=k+1}^{\infty} \underbrace{\|f_j\|}_{\leq 3^{-j}} \\ &\geq \rho_{W_k}(f_k) - \sum_{j=k+1}^{\infty} 3^{-j} \\ &= 3^{-k} - \frac{1}{1 - \frac{1}{3}} 3^{-k-1} \\ &= 3^{-k} \left(1 - \frac{1}{2} \right) \\ &\geq \varepsilon_{n_k} \\ &> \varepsilon_n. \end{aligned}$$

Zu gegebenem n wähle die Zahl k so, dass $n_{k-1} \leq n < n_k$ ist. Wegen $V_n \subset W_k$ folgt dann

$$\rho_{V_n}(f) \geq \rho_{W_k}(f) > \varepsilon_n. \quad \square$$

BEMERKUNG IV.1.10. Seien V_1, V_2, \dots und $(\varepsilon_n)_{n \in \mathbb{N}^*}$ wie in Satz IV.1.9. Dann ist die Menge aller $f \in X$ mit $\|f\| \leq 1$ und $\rho_{V_n}(f) \leq \varepsilon_n$ für alle $n \in \mathbb{N}^*$ kompakt in X . Also ist jede Menge mit vorgegebenem Approximationsverhalten kompakt und damit sehr dünn in dem unendlich dimensionalen Raum X . Insbesondere enthält sie keine noch so kleine Kugel.

BEWEISIDEE. Im normierten Raum reicht es, zu zeigen, dass es zu jedem $\delta > 0$ eine endliche Überdeckung durch Kugeln mit Radius δ gibt. Wähle zu gegebenem δ ein n , so dass $\varepsilon_n < \frac{\delta}{2}$ ist. Dann gibt es zu jedem Element f der Menge eines in V_n mit Abstand kleiner als $\frac{\delta}{2}$. Die Bildmenge in V_n lässt sich wegen der endlichen Dimension von V_n durch endlich viele Kugeln mit Radius $\frac{\delta}{2}$ überdecken. Verdoppelt man die Radien, erhält man die gesuchte Überdeckung. \square

IV.2. Glattheitsmaße

DEFINITION IV.2.1. Sei I ein Intervall (offen oder abgeschlossen, beschränkt oder unbeschränkt) und f eine Funktion auf I . Die für $h \geq 0$ durch

$$\omega(f; h) = \sup_{x, y \in I, |x-y| \leq h} |f(x) - f(y)|$$

definierte Funktion $\omega(f; \cdot)$ heißt der *Stetigkeitsmodul* von f .

BEMERKUNG IV.2.2. (1) Es ist $\lim_{h \rightarrow 0} \omega(f; h) = 0$ genau dann, wenn f gleichmäßig stetig ist.

(2) Ist f periodisch, so betrachtet man den Stetigkeitsmodul der periodisch auf ganz \mathbb{R} fortgesetzten Funktion. Ist also z.B. f 2π -periodisch, muss man bei der Berechnung von $\omega(f; h)$ auch die Werte $f(-\pi + \frac{h}{2})$ und $f(\pi - \frac{h}{2}) = f(-\pi - \frac{h}{2})$ vergleichen.

BEISPIEL IV.2.3. (1) Für $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(t) = t$ ist $\omega(f; h) = h$.
 (2) Für $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(t) = t^2$ ist $f(t+h) - f(t) = h(2t+h)$. Da dieser Ausdruck nicht beschränkt ist, existiert der Stetigkeitsmodul nicht. Beschränken wir uns aber auf das Intervall $[-1, 1]$ erhalten wir

$$\omega(f; h) = \begin{cases} h(2-h) & \text{für } 0 \leq h \leq 1, \\ 1 & \text{für } h \geq 1. \end{cases}$$

(3) Ist $f \in C^1(I)$ und $|f'|$ auf I durch M beschränkt, folgt aus dem Mittelwertsatz $\omega(f; h) \leq Mh$.

(4) Sei $f(x) = \sqrt{x}$ auf $[0, 1]$. Für $x, y \in [0, 1]$ mit $x \neq y$ erhalten wir

$$\begin{aligned} |f(x) - f(y)| &= \left| \frac{x-y}{\sqrt{x} + \sqrt{y}} \right| \\ &\leq \frac{|x-y|}{\max\{\sqrt{x}, \sqrt{y}\}} \\ &\leq \frac{|x-y|}{\sqrt{|x-y|}}. \end{aligned}$$

Wegen $|f(h) - f(0)| = \sqrt{h}$ folgt hieraus $\omega(f; h) = \sqrt{h}$ für $h \leq 1$.

(5) Sei $f(\varphi) = \cos \varphi$ für alle $\varphi \in \mathbb{R}$. Wegen $f(\varphi + \frac{h}{2}) - f(\varphi - \frac{h}{2}) =$

$2 \sin \varphi \sin \frac{h}{2}$ folgt

$$\omega(f; h) = \begin{cases} 2 \sin \frac{h}{2} & \text{für } h \leq \pi, \\ 1 & \text{für } h \geq \pi. \end{cases}$$

SATZ IV.2.4. *Der Stetigkeitsmodul hat folgende Eigenschaften:*

- (1) $\omega(f; \cdot)$ ist monoton wachsend.
- (2) $\omega(f; \cdot)$ ist sublinear, d.h.

$$\omega(f; h_1 + h_2) \leq \omega(f; h_1) + \omega(f; h_2).$$

- (3) $\omega(f; nh) \leq n\omega(f; h)$ für alle $n \in \mathbb{N}^*$.
- (4) $\omega(f; \lambda h) \leq (1 + \lambda)\omega(f; h)$ für alle $\lambda \in \mathbb{R}_+$.
- (5) $\lim_{h \rightarrow 0} \frac{\omega(f; h)}{h} = 0 \iff f$ ist konstant.

BEWEIS. *ad (1):* Folgt direkt aus der Definition.

ad (2): Wegen der Dreiecksungleichung gilt für alle $x \in I$ mit $x + h_1 + h_2 \in I$

$$\begin{aligned} & |f(x + h_1 + h_2) - f(x)| \\ & \leq |f(x + h_1 + h_2) - f(x + h_1)| + |f(x + h_1) - f(x)| \\ & \leq \omega(f; h_1) + \omega(f; h_2). \end{aligned}$$

Da die rechte Seite dieser Abschätzung nicht von x abhängt, folgt die Behauptung.

ad (3): Folgt mit Induktion aus (2).

ad (4): Für $\lambda \geq 0$ wähle $n \in \mathbb{N}^*$ so, dass $n \leq \lambda \leq n + 1$ ist. Aus (1) und (3) folgt dann

$$\begin{aligned} \omega(f; \lambda h) & \leq \omega(f; (n + 1)h) \\ & \leq (n + 1)\omega(f; h) \\ & \leq (1 + \lambda)\omega(f; h). \end{aligned}$$

ad (5): Die Richtung „ \Leftarrow “ ist offensichtlich. Für $h > 0$ und $n \in \mathbb{N}^*$ erhalten wir aus (3)

$$\omega(f; h) = \omega(f; n \frac{h}{n}) \leq n \omega(f; \frac{h}{n}) = h \underbrace{\frac{\omega(f; \frac{h}{n})}{\frac{h}{n}}}_{\xrightarrow{n \rightarrow \infty} 0}.$$

Also ist $\omega(f; h) = 0$ für alle $h > 0$. Dies beweist die Richtung „ \Rightarrow “. \square

DEFINITION IV.2.5. Sei I ein kompaktes, perfektes Intervall.

(1) Eine Funktion $f : I \rightarrow \mathbb{R}$ genügt einer *Lipschitz-Bedingung* der Ordnung $\alpha > 0$ mit Konstanter M , kurz $f \in \text{Lip}_M \alpha$, wenn gilt $\omega(f; h) \leq Mh^\alpha$.

(2) Die Menge

$$\text{Lip } \alpha = \bigcup_{M > 0} \{f : f \in \text{Lip}_M \alpha\} = \{f : \omega(f; h) = O(h^\alpha)\}$$

heißt *Lipschitz-Klasse* der Ordnung α .

(3) Für $k \in \mathbb{N}^*$ und $\alpha \in (0, 1)$ bezeichnet $C^{k+\alpha}(I)$ die Menge aller k -mal stetig differenzierbaren Funktionen auf I mit $f^{(k)} \in \text{Lip } \alpha$.

Zur späteren Behandlung eines Grenzfalles benötigen wir:

DEFINITION IV.2.6. Sei I ein Intervall (offen oder abgeschlossen, beschränkt oder unbeschränkt) und f eine Funktion auf I . Die für $h \geq 0$ durch

$$\omega_2(f; h) = \sup_{|\delta| \leq h, t \pm \delta \in I} |f(t + \delta) - 2f(t) + f(t - \delta)|$$

definierte Funktion $\omega_2(f; \cdot)$ heißt der *zweite Stetigkeitsmodul* von f .

BEMERKUNG IV.2.7. (1) Es ist $\omega_2(f; h) \leq 2\omega(f; h)$.

(2) Ist f differenzierbar, so ist $\omega_2(f; h) \leq h\omega(f'; h)$.

BEWEIS. *ad* (1): Folgt aus der Dreiecksungleichung.

ad (2): Folgt aus dem Mittelwertsatz. □

SATZ IV.2.8. *Der zweite Stetigkeitsmodul hat folgende Eigenschaften:*

(1) $\omega_2(f; \cdot)$ ist monoton wachsend.

(2) $\omega_2(f; nh) \leq n^2\omega_2(f; h)$ für alle $n \in \mathbb{N}^*$.

(3) $\omega_2(f; \lambda h) \leq (1 + \lambda)^2\omega_2(f; h)$ für alle $\lambda \in \mathbb{R}_+$.

(4) $\lim_{h \rightarrow 0} \frac{\omega_2(f; h)}{h^2} = 0 \iff f$ ist linear.

BEWEIS. Übungsaufgabe. □

IV.3. Jackson-Sätze

In diesem Abschnitt beweisen wir Sätze von der Form „Glattheit von $f \implies$ Konvergenzgeschwindigkeit von $E_n(f)$ “.

Wir beginnen mit der Approximation 2π -periodischer Funktionen durch trigonometrische Polynome. Hierzu bezeichnen wir mit $C_{2\pi}$ den Raum der stetigen 2π -periodischen Funktionen und mit

$$\mathbb{T}_n = \text{span}\{1, \cos(k\varphi), \sin(k\varphi) : 1 \leq k \leq n\}$$

den Raum der trigonometrischen Polynome vom Grad $\leq n$. Wir erinnern daran, dass \mathbb{T}_n ein Haarscher Raum ist.

Ist f 2π -periodisch, so ist

$$f(\varphi) \sim \frac{1}{2}a_0 + \sum_{k \geq 1} \{a_k \cos(k\varphi) + b_k \sin(k\varphi)\}$$

die formale Fourier-Reihe von f . Aus den Orthogonalitätsbeziehungen

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \cos(n\varphi) \cos(m\varphi) d\varphi = \begin{cases} 2 & \text{für } n = m = 0, \\ 1 & \text{für } n = m \neq 0, \\ 0 & \text{für } n \neq m, \end{cases}$$

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \sin(n\varphi) \sin(m\varphi) d\varphi = \begin{cases} 1 & \text{für } n = m, \\ 0 & \text{für } n \neq m, \end{cases}$$

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \cos(n\varphi) \sin(m\varphi) d\varphi = 0 \quad \text{für alle } n, m$$

folgt, dass für die Fourier-Koeffizienten a_n, b_n gilt

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\varphi) \cos(k\varphi) d\varphi, \quad k \geq 0,$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\varphi) \sin(k\varphi) d\varphi, \quad k \geq 1.$$

Hieraus und aus den Additionstheoremen ergibt sich

$$\begin{aligned} & a_k \cos(k\varphi) + b_k \sin(k\varphi) \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \{ \cos(kt) \cos(k\varphi) + \sin(kt) \sin(k\varphi) \} dt \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(k(\varphi - t)) dt. \end{aligned}$$

Ist

$$K(t) = \frac{1}{2} w_0 + \sum_{k=1}^n w_k \cos(k\varphi)$$

ein gerades trigonometrisches Polynom, so folgt hieraus

$$\begin{aligned} \frac{1}{\pi} \int_{-\pi}^{\pi} K(\varphi - t) f(t) dt &= \frac{1}{\pi} \int_{-\pi}^{\pi} K(t) f(\varphi - t) dt \\ &= \frac{1}{2} a_0 w_0 + \sum_{k=1}^n w_k \{ a_k \cos(k\varphi) + b_k \sin(k\varphi) \}. \end{aligned}$$

Eine besondere Rolle wird der Kern

$$\overline{K}_n(t) = \left[\frac{\sin\left(\frac{nt}{2}\right)}{\sin\frac{t}{2}} \right]^2$$

spielen. Wegen $2 \sin^2 \varphi = 1 - \cos(2\varphi)$ ergibt sich mit $u = \cos t$ die Darstellung

$$\overline{K}_n(t) = \frac{1 - \cos(nt)}{1 - \cos t} = \frac{1 - T_n(u)}{1 - u},$$

wobei T_n das n -te Čebyšev-Polynom ist. Da $T_n(1) = 1$ ist, folgt dass $\frac{1 - T_n(u)}{1 - u} \in \mathbb{P}_{n-1}$ ist. Also ist $\overline{K}_n(t) \in \mathbb{T}_{n-1}$. Diese Vorüberlegungen werden im Folgenden benutzt.

DEFINITION IV.3.1. Für $m \in \mathbb{N}^*$ sei

$$\lambda_m = \int_{-\pi}^{\pi} \left[\frac{\sin\left(\frac{mt}{2}\right)}{\sin\frac{t}{2}} \right]^4 dt$$

und

$$K_m(t) = \lambda_m^{-1} \left[\frac{\sin\left(\frac{mt}{2}\right)}{\sin\frac{t}{2}} \right]^4.$$

K_m ist ein trigonometrisches Polynom vom Grad $2m - 2$ und heißt *Jackson-Kern*. Mit seiner Hilfe definieren wir für jedes $n \in \mathbb{N}$ einen linearen Operator $J_n : C_{2\pi} \rightarrow \mathbb{T}_n$ durch

$$J_n f(\varphi) = \int_{-\pi}^{\pi} K_{n'}(t) f(\varphi - t) dt$$

mit

$$n' = \begin{cases} \frac{n+1}{2} & \text{falls } n \text{ ungerade,} \\ \frac{n+2}{2} & \text{falls } n \text{ gerade.} \end{cases}$$

BEMERKUNG IV.3.2. (1) Für alle $t \in [0, \frac{\pi}{2}]$ gilt

$$\frac{2}{\pi}t \leq \sin t \leq t.$$

Für alle $t \in [0, \frac{\pi}{4}]$ gilt

$$\frac{2\sqrt{2}}{\pi}t \leq \sin t \leq t.$$

(2) Für alle $n \in \mathbb{N}$ ist

$$\lambda_n \geq \frac{5}{2}n^3.$$

(3) Für alle $n \in \mathbb{N}$ gilt

$$\int_0^{\pi} t K_n(t) dt \leq \frac{2}{5}\pi^2 n^{-1}$$

und

$$\int_0^{\pi} t^2 K_n(t) dt \leq \frac{8}{15}\pi^3 n^{-2}.$$

BEWEIS. *ad* (1): Aus dem Mittelwertsatz folgt für $0 \leq t \leq \frac{\pi}{2}$

$$\sin t = t \cos \theta \leq t \quad \text{mit } 0 \leq \theta \leq t.$$

Wegen $(\sin t)'' = -\sin t$ ist der Sinus auf $[0, \frac{\pi}{2}]$ konkav und liegt oberhalb der Sehne. Hieraus folgen die unteren Schranken.

ad (2): Wegen

$$\begin{aligned} \int_{-\pi}^{\pi} \left[\frac{\sin\left(\frac{nt}{2}\right)}{\sin\frac{t}{2}} \right]^4 dt &= 2 \int_0^{\pi} \left[\frac{\sin\left(\frac{nt}{2}\right)}{\sin\frac{t}{2}} \right]^4 dt \\ &\geq 2 \int_0^{\pi} \left[\frac{\sin\left(\frac{nt}{2}\right)}{\frac{t}{2}} \right]^4 dt \end{aligned}$$

folgt mit der Substitution $t = \frac{2}{n}\varphi$ aus der Definition von λ_n und Teil (1)

$$\begin{aligned}
\lambda_n &= \int_{-\pi}^{\pi} \left[\frac{\sin\left(\frac{nt}{2}\right)}{\sin\frac{t}{2}} \right]^4 dt \\
&\geq 4n^3 \int_0^{\frac{n\pi}{2}} \frac{\sin^4 \varphi}{\varphi^4} d\varphi \\
&\geq 4n^3 \left\{ \int_0^{\frac{\pi}{4}} \frac{\sin^4 \varphi}{\varphi^4} d\varphi + \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{\sin^4 \varphi}{\varphi^4} d\varphi \right\} \\
&\geq 4n^3 \left\{ \int_0^{\frac{\pi}{4}} \left(\frac{2\sqrt{2}}{\pi} \right)^4 d\varphi + \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \left(\frac{2}{\pi} \right)^4 d\varphi \right\} \\
&= 4n^3 \left\{ \frac{\pi}{4} \frac{64}{\pi^4} + \frac{\pi}{4} \frac{16}{\pi^4} \right\} \\
&= \frac{80}{\pi^3} n^3 \\
&\geq \frac{5}{2} n^3.
\end{aligned}$$

ad (3): Aus den Additionstheoremen folgt

$$\sin((n+1)\varphi) = \sin(n\varphi) \cos \varphi + \cos(n\varphi) \sin \varphi$$

und daher

$$\begin{aligned}
|\sin((n+1)\varphi)| &= |\sin(n\varphi)| |\cos \varphi| + |\cos(n\varphi)| |\sin \varphi| \\
&\leq |\sin(n\varphi)| + |\sin \varphi|.
\end{aligned}$$

Hieraus folgt durch Induktion für alle n und alle φ die Abschätzung

$$|\sin(n\varphi)| \leq n |\sin \varphi|.$$

Aus dieser Abschätzung, Bemerkung IV.3.2 (1) und Teil (2) folgt

$$\begin{aligned}
\int_0^{\pi} t K_n(t) dt &= \underbrace{\lambda_n^{-1}}_{\leq \frac{2}{5} n^{-3}} \left\{ \int_0^{\frac{\pi}{n}} t \underbrace{\left[\frac{\sin\left(\frac{nt}{2}\right)}{\sin\frac{t}{2}} \right]^4}_{\leq n^4} dt + \int_{\frac{\pi}{n}}^{\pi} t \underbrace{\left[\frac{\sin\left(\frac{nt}{2}\right)}{\sin\frac{t}{2}} \right]^4}_{\leq \left(\frac{1}{t/\pi}\right)^4} dt \right\} \\
&\leq \frac{2}{5} n^{-3} \left\{ \int_0^{\frac{\pi}{n}} t n^4 dt + \int_{\frac{\pi}{n}}^{\pi} \pi^4 t^{-3} dt \right\} \\
&= \frac{2}{5} n^{-3} \left\{ \frac{1}{2} n^4 \frac{\pi^2}{n^2} + \frac{\pi^4}{2} \left[\left(\frac{\pi}{n}\right)^{-2} - \pi^{-2} \right] \right\} \\
&\leq \frac{2}{5} \pi^2 n^{-1}.
\end{aligned}$$

Mit denselben Argumenten erhalten wir

$$\begin{aligned}
\int_0^\pi t^2 K_n(t) dt &\leq \frac{2}{5} n^{-3} \left\{ \int_0^{\frac{\pi}{n}} n^4 t^2 dt + \int_{\frac{\pi}{n}}^\pi \pi^4 t^{-2} dt \right\} \\
&\leq \frac{2}{5} n^{-3} \left\{ \frac{1}{3} n^4 \left(\frac{\pi}{n} \right)^3 + \pi^4 \left[\frac{n}{\pi} - \frac{1}{\pi} \right] \right\} \\
&\leq \frac{2}{5} n^{-2} \pi^3 \left\{ \frac{1}{3} + 1 \right\} \\
&= \frac{8}{15} \pi^3 n^{-2}. \quad \square
\end{aligned}$$

SATZ IV.3.3. Für $f \in C_{2\pi}$ ist bei Approximation durch trigonometrische Polynome vom Grad $\leq n$

$$E_n(f) \leq 7\omega \left(f; \frac{\pi}{n+1} \right).$$

BEWEIS. Offensichtlich ist

$$E_n(f) \leq \|f - J_n f\|.$$

Wegen

$$\int_{-\pi}^\pi K_n(t) dt = 1$$

für alle n und der Symmetrie der K_n erhalten wir für alle $\varphi \in [\pi, \pi]$

$$\begin{aligned}
&|f(\varphi) - J_n f(\varphi)| \\
&= \left| \int_{-\pi}^\pi [f(\varphi) - f(\varphi - t)] K_n'(t) dt \right| \\
&= \left| \int_0^\pi [2f(\varphi) - f(\varphi - t) - f(\varphi + t)] K_n'(t) dt \right| \\
&\leq \int_0^\pi \omega_2(f; t) K_n'(t) dt.
\end{aligned}$$

Aus den Abschätzungen $\omega_2(f; h) \leq 2\omega(f; h)$ und

$$\begin{aligned}
\omega(f; h) &= \omega \left(f; \frac{\pi}{n+1} \frac{n+1}{\pi} h \right) \\
&\leq \left(1 + \frac{n+1}{\pi} h \right) \omega \left(f; \frac{\pi}{n+1} \right)
\end{aligned}$$

sowie $n' \geq \frac{n+1}{2}$ folgt

$$\begin{aligned}
\|f - J_n f\| &\leq \int_0^\pi \omega_2(f; t) K_{n'}(t) dt \\
&\leq \int_0^\pi 2K_{n'}(t) dt \\
&\leq 2\omega\left(f; \frac{\pi}{n+1}\right) \int_0^\pi \left(1 + \frac{n+1}{\pi}t\right) K_{n'}(t) dt \\
&\leq 2\omega\left(f; \frac{\pi}{n+1}\right) \left\{ \frac{1}{2} + \frac{n+1}{\pi} \frac{2}{5} \pi^2 (n')^{-1} \right\} \\
&\leq \omega\left(f; \frac{\pi}{n+1}\right) \left\{ 1 + \frac{8}{5} \pi \right\} \\
&\leq 7\omega\left(f; \frac{\pi}{n+1}\right). \quad \square
\end{aligned}$$

Aus Satz IV.3.3, Definition IV.2.5 (S. 57) und Beispiel IV.2.3(3) (S. 56) folgt:

KOROLLAR IV.3.4. Für $f \in C_{2\pi} \cap \text{Lip}_M 1$ ist

$$E_n(f) \leq \frac{7\pi M}{n+1}.$$

Für $f \in C_{2\pi}^1$ ist

$$E_n(f) \leq \frac{7\pi}{n+1} \|f'\|.$$

Bei der Untersuchung der Approximationsgüte differenzierbarer Funktionen ist zu beachten, dass nicht jede Funktion $f \in C_{2\pi}$ darstellbar ist als Ableitung einer Funktion $g \in C_{2\pi}$. Ist nämlich $f = g'$ mit $g \in C_{2\pi}$, so folgt

$$\int_{-\pi}^{\pi} f(t) dt = \int_{-\pi}^{\pi} g'(t) dt = g(\pi) - g(-\pi) = 0.$$

Wie man leicht einsieht, ist diese Bedingung auch hinreichend für die 2π -Periodizität der Stammfunktion $\int f(t) dt$. Diese Bedingung ist offensichtlich äquivalent dazu, dass der Fourier-Koeffizient a_0 von f verschwindet. Daher betrachten wir die Approximation durch trigonometrische Polynome mit verschwindendem Fourier-Koeffizienten a_0 :

$$E_n^0(f) = \inf_{a_k, b_k} \left\| f - \sum_{k=1}^n (a_k \cos(k\varphi) + b_k \sin(k\varphi)) \right\|.$$

SATZ IV.3.5. (1) Für alle $f \in C_{2\pi}^1$ ist

$$E_n(f) \leq \frac{7\pi}{n+1} E_n^0(f').$$

(2) Verschwindet der Fourier-Koeffizient a_0 von $f \in C_{2\pi}$, so gelten die Abschätzungen von Satz IV.3.3 und Korollar IV.3.4 mit gleichen Konstanten auch für $E_n^0(f)$.

BEWEIS. *ad (1)*: Sei $q = \sum (a_k \cos(k\varphi) + b_k \sin(k\varphi))$ die b.A. an f' durch Polynome mit verschwindendem Fourier-Koeffizienten a_0 . Setze

$$p = \sum_{k=1}^n \frac{1}{k} (-b_k \cos(k\varphi) + a_k \sin(k\varphi)).$$

Dann ist

$$p' = q$$

und

$$\begin{aligned} E_n(f) &= E_n(f - p) \\ &\leq \frac{7\pi}{n+1} \|(f - p)'\| \\ &= \frac{7\pi}{n+1} \|f' - q\| \\ &= \frac{7\pi}{n+1} E_n^0(f'). \end{aligned}$$

ad (2): Verschwindet der Fourier-Koeffizient a_0 von f , so gilt gleiches für den Fourier-Koeffizienten a_0 von $J_n f$. Damit ist in diesem Fall

$$E_n^0(f) \leq \|f - J_n f\|.$$

Damit folgt die Behauptung aus dem Beweis von Satz IV.3.3. \square

Aus Satz IV.3.5 folgt durch Induktion:

KOROLLAR IV.3.6. Seien $k \in \mathbb{N}$, $0 < \alpha \leq 1$ und $M \in \mathbb{R}_+$. Für alle $f \in C_{2\pi}^k$ ist

$$E_n(f) \leq \left(\frac{7\pi}{n+1} \right)^k \|f^{(k)}\|.$$

Für alle $f \in C_{2\pi}^{k+\alpha}$ mit $f^{(k)} \in \text{Lip}_M \alpha$ ist

$$E_n(f) \leq \left(\frac{7\pi}{n+1} \right)^{k+\alpha} M.$$

BEMERKUNG IV.3.7. Verschwindet der Fourier-Koeffizient a_0 von f , so verschwindet nicht notwendig der entsprechende Fourier-Koeffizient der b.A. an f . Betrachte z.B. die in Abbildung IV.3.1 skizzierte Funktion. Die gleichförmigen positiven und negativen Spitzen sind so gewählt, dass das Integral verschwindet und die absoluten Maxima verschieden sind. Die b.A. in \mathbb{T}_0 ist dann eine von Null verschiedene Konstante.

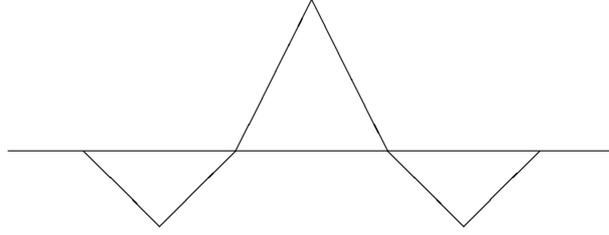


ABBILDUNG IV.3.1. Funktion mit verschwindendem Fourier-Koeffizienten a_0 aber nicht verschwindender b.A. durch konstante Funktionen

Wir kommen nun zur Approximation durch algebraische Polynome in $C([-1, 1])$ zurück. Zu gegebenem $f \in C([-1, 1])$ wird durch $g(\varphi) = f(\cos \varphi)$ eine gerade Funktion $g \in C_{2\pi}$ definiert. Die b.A. q an g in \mathbb{T}_n ist dann auch gerade und lässt sich daher in der Form $q(\varphi) = p(\cos \varphi)$ mit einem algebraischen Polynom $p \in \mathbb{P}_n$ schreiben. Also ist

$$\begin{aligned} E_n(f) &\leq \|f - p\| \\ &= \|f \circ \cos - p \circ \cos\| \\ &= \|g - q\| \\ &= E_n(g). \end{aligned}$$

Man kann sogar zeigen, dass diese Abschätzung scharf ist. Zur Abschätzung von $\omega(g; \cdot)$ nutzen wir aus, dass nach dem Mittelwertsatz gilt

$$\begin{aligned} |\cos \varphi - \cos \psi| &= |\sin(\theta\varphi + (1 - \theta)\psi)(\varphi - \psi)| \quad \text{für ein } \theta \in [0, 1] \\ &\leq |\varphi - \psi|. \end{aligned}$$

Damit folgt

$$\begin{aligned} \omega(g; h) &= \sup_{|\varphi - \psi| \leq h \pmod{2\pi}} |g(\varphi) - g(\psi)| \\ &\leq \sup_{|\cos \varphi - \cos \psi| \leq h} |g(\varphi) - g(\psi)| \\ &\leq \sup_{|\cos \varphi - \cos \psi| \leq h} \left| \underbrace{f(\cos \varphi)}_{=x} - \underbrace{f(\cos \psi)}_{=y} \right| \\ &\leq \sup_{|x - y| \leq h} |f(x) - f(y)| \\ &= \omega(f; h). \end{aligned}$$

SATZ IV.3.8. Für die Approximation durch algebraische Polynome gilt:

(1) Ist $f \in C([-1, 1])$, so ist

$$E_n(f) \leq 7\omega\left(f; \frac{\pi}{n+1}\right).$$

(2) Ist $f \in C^k([-1, 1])$, so ist für $n \geq k$

$$E_n(f) \leq \frac{(7\pi)^k}{(n+1) \cdot n \cdot \dots \cdot (n+2-k)} \|f^{(k)}\|.$$

(3) Ist $f \in C^{k+\alpha}([-1, 1])$ mit $f^{(k)} \in \text{Lip}_M \alpha$, so ist für $n \geq k$

$$E_n(f) \leq \frac{(7\pi)^{k+\alpha}}{(n+1) \cdot n \cdot \dots \cdot (n+2-k)(n+1-k)^\alpha} M.$$

BEWEIS. *ad* (1): Folgt wegen der Vorüberlegungen aus Satz IV.3.3.

ad (2): Sei $q \in \mathbb{P}_{n-1}$ die b.A. an f' . Setze

$$p(x) = \int_0^x q(t) dt \in \mathbb{P}_n.$$

Dann folgt

$$\begin{aligned} E_n(f) &= E_n(f - p) \\ &\leq 7\omega\left(f - p; \frac{\pi}{n+1}\right) \\ &\leq \frac{7\pi}{n+1} \|(f - p)'\| \\ &= \frac{7\pi}{n+1} \|f' - q\| \\ &= \frac{7\pi}{n+1} E_{n-1}(f'). \end{aligned}$$

Hieraus folgt die Behauptung mittels Induktion.

ad (3): Folgt aus den Teilen (1) und (2). \square

BEMERKUNG IV.3.9. Aus Satz IV.3.8(1) folgt der Satz von Weierstraß. Denn ist $f \in C([-1, 1])$, so ist f gleichmäßig stetig und somit $\lim_{n \rightarrow \infty} \omega(f; \frac{\pi}{n+1}) = 0$.

IV.4. Ungleichungen von Bernstein und Markov

Wir erinnern zunächst an die Lagrange-Interpolation. Seien $x_0 < \dots < x_n$ gegebene Punkte. Setze

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

Dann ist

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

das eindeutig bestimmte Polynom in \mathbb{P}_n , das f in den Punkten x_0, \dots, x_n interpoliert. Wegen

$$\omega'(x) = \sum_{i=0}^n \prod_{\substack{j=0 \\ j \neq i}}^n (x - x_j) = \sum_{i=0}^n \frac{\omega(x)}{x - x_i}$$

ist

$$\omega'(x_i) = \prod_{\substack{j=0 \\ j \neq i}}^n (x_i - x_j)$$

und

$$\ell_i(x) = \frac{\omega(x)}{(x - x_i)\omega'(x_i)}.$$

SATZ IV.4.1 (Lemma von Schur). Für jedes $p \in \mathbb{P}_{n-1}$ gilt in $C([-1, 1])$

$$\|p\| \leq n \left\| \sqrt{1 - x^2} p \right\|.$$

BEWEIS. Setze

$$M = \left\| \sqrt{1 - x^2} p \right\|.$$

Bezeichne mit

$$\eta_i = \cos\left(\frac{2i-1}{2n}\pi\right), \quad 1 \leq i \leq n$$

die Nullstellen des n -ten Čebyšev-Polynoms T_n . Für $\eta_n \leq x \leq \eta_1$ folgt wegen $\sin \varphi \geq \frac{2}{\pi} \varphi$

$$\begin{aligned} \sqrt{1 - x^2} &\geq \sqrt{1 - \eta_1^2} \\ &= \sqrt{1 - \cos^2\left(\frac{\pi}{2n}\right)} \\ &= \sin\left(\frac{\pi}{2n}\right) \\ &\geq \frac{1}{n} \end{aligned}$$

und somit

$$\begin{aligned} |p(x)| &= \frac{1}{\sqrt{1 - x^2}} \left| \sqrt{1 - x^2} p(x) \right| \\ &\leq n \left| \sqrt{1 - x^2} p(x) \right| \\ &\leq nM. \end{aligned}$$

Da p sich selber interpoliert, ist

$$p(x) = \sum_{i=1}^n p(\eta_i) \frac{T_n(x)}{(x - \eta_i)T_n'(\eta_i)}.$$

Mit der Kettenregel folgt

$$\begin{aligned} T'_n(x) &= \frac{d}{dx} \cos(n \arccos x) \\ &= \frac{d}{d\varphi} \cos(n\varphi) \frac{d}{dx} (\arccos x) \\ &= n \frac{\sin(n\varphi)}{\sin \varphi}. \end{aligned}$$

Insbesondere ist

$$\begin{aligned} T'_n(\eta_i) &= n \frac{\sin\left(n \frac{2i-1}{2n} \pi\right)}{\sin\left(\frac{2i-1}{2n} \pi\right)} \\ &= n(-1)^{i-1} (1 - \eta_i^2)^{-\frac{1}{2}} \end{aligned}$$

und

$$T'_n(1) = \lim_{\varphi \rightarrow 0} n \frac{\sin(n\varphi)}{\sin \varphi} = n^2.$$

Hieraus und aus $|p(\eta_i)| \sqrt{1 - \eta_i^2} \leq M$ für alle i folgt für alle x

$$\begin{aligned} |p(x)| &= \left| \sum_{i=1}^n p(\eta_i) \frac{T_n(x)}{(x - \eta_i) T'_n(\eta_i)} \right| \\ &\leq \sum_{i=1}^n |p(\eta_i)| \frac{|T_n(x)|}{|x - \eta_i| |T'_n(\eta_i)|} \\ &\leq \sum_{i=1}^n |p(\eta_i)| \frac{|T_n(x)|}{|x - \eta_i|} \frac{1}{n \sqrt{1 - \eta_i^2}} \\ &\leq \frac{M}{n} \sum_{i=1}^n \frac{|T_n(x)|}{|x - \eta_i|}. \end{aligned}$$

Sei nun speziell $x \geq \eta_1$ oder $x \leq \eta_n$. Dann haben alle Brüche in obiger Summe gleiches Vorzeichen und der Betrag kann aus der Summe gezogen werden. Wegen

$$T_n(x) = 2^{n-1} \prod_{i=1}^n (x - \eta_i)$$

und unserer Vorüberlegungen ist

$$\sum_{i=1}^n \frac{T_n(x)}{x - \eta_i} = T'_n(x).$$

Wegen

$$|\sin(n\varphi)| \leq n |\sin \varphi|$$

ist

$$|T'_n(x)| = n \left| \frac{\sin(n\varphi)}{\sin \varphi} \right| \leq n^2.$$

Wegen $T'_n(1) = n^2$ folgt insbesondere

$$\|T'_n\| = n^2.$$

Also ist für $x \geq \eta_1$ oder $x \leq \eta_n$

$$\begin{aligned} \sum_{i=1}^n \frac{|T_n(x)|}{|x - \eta_i|} &= \left| \sum_{i=1}^n \frac{T_n(x)}{x - \eta_i} \right| \\ &= |T'_n(x)| \\ &\leq n^2. \end{aligned}$$

Hieraus folgt die Behauptung. \square

BEMERKUNG IV.4.2. Die Abschätzung von Satz IV.4.1 ist scharf. Denn für $p = T'_n$ folgt aus seinem Beweis

$$\|p\| = \|T'_n\| = n^2$$

und

$$n \left\| \sqrt{1-x^2} p \right\| = n \left\| \sqrt{1-x^2} T'_n \right\| = n \left\| (\sin \varphi) n \frac{\sin(n\varphi)}{\sin \varphi} \right\| = n^2.$$

KOROLLAR IV.4.3. Für jedes ungerade trigonometrische Polynom S vom Grad $\leq n$ ist

$$\left\| \frac{S(\varphi)}{\sin \varphi} \right\| \leq n \|S\|.$$

BEWEIS. Es ist

$$\begin{aligned} S(\varphi) &= \sum_{k=1}^n b_k \sin(k\varphi) \\ &= \sum_{k=1}^n b_k \frac{1}{k} T'_k(\cos \varphi) \sin \varphi. \end{aligned}$$

Also ist S in der Form

$$S(\varphi) = p(\cos \varphi) \sin \varphi$$

mit $p \in \mathbb{P}_{n-1}$ darstellbar. Mit der Transformation $x = \cos \varphi$ folgt hieraus und aus Satz IV.4.1 die Behauptung. \square

SATZ IV.4.4 (Bernstein-Ungleichung). Für jedes trigonometrische Polynom S vom Grad $\leq n$ gilt

$$\|S'\| \leq n \|S\|.$$

Diese Abschätzung ist scharf.

BEWEIS. Wegen

$$(\sin(n\varphi))' = n \cos(n\varphi)$$

kann keine bessere Abschätzung gelten. Sei nun S ein trigonometrisches Polynom vom Grad $\leq n$ und $\alpha \in \mathbb{R}$ beliebig. Definiere die Hilfsfunktion

$$f(\varphi) = \frac{1}{2}[S(\alpha + \varphi) - S(\alpha - \varphi)].$$

Wegen der Additionstheoreme ist f ein trigonometrische Polynom vom Grad $\leq n$. Offensichtlich ist f ungerade. Ferner ist

$$\|f\| \leq \|S\|$$

und

$$\begin{aligned} S'(\alpha) &= \lim_{\varphi \rightarrow 0} \frac{S(\alpha + \varphi) - S(\alpha - \varphi)}{2\varphi} \\ &= \lim_{\varphi \rightarrow 0} \frac{f(\varphi) \sin \varphi}{\sin \varphi \varphi} \\ &= \lim_{\varphi \rightarrow 0} \frac{f(\varphi)}{\sin \varphi}. \end{aligned}$$

Gemäß Korollar IV.4.3 gilt aber für alle φ

$$\left| \frac{f(\varphi)}{\sin \varphi} \right| \leq \left\| \frac{f}{\sin} \right\| \leq n \|f\| \leq n \|S\|.$$

Also ist auch

$$|S'(\alpha)| \leq n \|S\|.$$

Da α beliebig war, folgt die Behauptung. \square

SATZ IV.4.5 (Markov-Ungleichung). *Für jedes Polynom $p \in \mathbb{P}_n$ gilt in $C([-1, 1])$*

$$\|p'\| \leq n \left\| \sqrt{1-x^2} p' \right\| \leq n^2 \|p\|.$$

Diese Abschätzung ist scharf.

BEWEIS. Wegen

$$\|T_n'\| = n^2 = n^2 \|T_n\|$$

kann keine bessere Abschätzung gelten. Wegen Satz IV.4.1 ist nur die zweite Ungleichung zu zeigen. Setze hierzu

$$S(\varphi) = p(\cos \varphi).$$

Dann ist nach der Kettenregel

$$S' = \frac{dS}{d\varphi} = -\sin \varphi \frac{dp}{dx}(\cos \varphi) = -\sqrt{1-x^2} p'(x).$$

Damit ergibt sich aus Satz IV.4.4

$$\left\| \sqrt{1-x^2} p' \right\| = \|S'\| \leq n \|S\| = n \|p\|. \quad \square$$

IV.5. Umkehrsätze von Bernstein und Zygmund

In diesem Abschnitt beweisen wir die Umkehrung der Jackson-Sätze aus § IV.3, d.h. Aussagen der Form „Konvergenzgeschwindigkeit von $E_n(f) \Rightarrow$ Glattheit von f “.

SATZ IV.5.1 (Bernstein). *Sei f eine 2π -periodische Funktion. Es gebe Zahlen $A \in \mathbb{R}_+$ und $\alpha \in (0, 1)$, so dass für die Approximation durch trigonometrische Polynome gilt $E_n(f) \leq An^{-\alpha}$. Dann ist f in $C_{2\pi}^\alpha$.*

BEWEIS. Für $n \in \mathbb{N}$ sei p_n ein trigonometrisches Polynom vom Grade $\leq n$ mit

$$\|f - p_n\| \leq An^{-\alpha}.$$

Setze $v_0 = p_1$, $v_n = p_{2^n} - p_{2^{n-1}}$. Dann ist $f = \sum_{n=0}^{\infty} v_n$. Wegen

$$\begin{aligned} \|v_n\| &\leq \|f - p_{2^n}\| + \|f - p_{2^{n-1}}\| \\ &\leq A2^{-\alpha n} + A2^{-\alpha(n-1)} \\ &= \underbrace{A(1 + 2^\alpha)}_{=A_1} 2^{-\alpha n} \end{aligned}$$

konvergiert diese Reihe gleichmäßig und kann durch geometrische Reihen abgeschätzt werden. Seien φ und ψ beliebig. Dann ist

$$\begin{aligned} |v_n(\varphi) - v_n(\psi)| &\leq |v_n(\varphi)| + |v_n(\psi)| \\ &\leq 2\|v_n\| \end{aligned}$$

und nach Satz IV.4.4 (S. 69)

$$\begin{aligned} |v_n(\varphi) - v_n(\psi)| &= |v_n'(\eta)(\varphi - \psi)| \\ &\leq \|v_n'\| |\varphi - \psi| \\ &\leq 2^n \|v_n\| |\varphi - \psi|. \end{aligned}$$

Für beliebiges $m \in \mathbb{N}$ folgt aus diesen Abschätzungen

$$\begin{aligned} |f(\varphi) - f(\psi)| &= \left| \sum_{n=0}^{\infty} (v_n(\varphi) - v_n(\psi)) \right| \\ &\leq \sum_{n=0}^{\infty} |v_n(\varphi) - v_n(\psi)| \\ &\leq \sum_{n=0}^{m-1} 2^n \|v_n\| |\varphi - \psi| + \sum_{n=m}^{\infty} 2 \|v_n\| \\ &\leq A_1 |\varphi - \psi| \sum_{n=0}^{m-1} 2^{n-\alpha n} + 2A_1 \sum_{n=m}^{\infty} 2^{-\alpha n} \\ &= A_1 |\varphi - \psi| \frac{2^{m(1-\alpha)} - 1}{2^{1-\alpha} - 1} + 2A_1 \frac{2^{-m\alpha}}{1 - 2^{-\alpha}} \\ &= A_2 \{ |\varphi - \psi| 2^{m(1-\alpha)} + 2^{-m\alpha} \} \end{aligned}$$

mit

$$A_2 = A_1 \left\{ \frac{1}{2^{1-\alpha} - 1} + \frac{2}{1 - 2^{-\alpha}} \right\}.$$

Sei nun $|\varphi - \psi| \leq h < 1$ und $m \in \mathbb{N}$ so gewählt, dass $\frac{1}{2h} \leq 2^m < \frac{1}{h}$ ist. Dann folgt

$$\begin{aligned} |f(\varphi) - f(\psi)| &\leq A_2 \{h2^{m(1-\alpha)} + 2^{-m\alpha}\} \\ &\leq A_2 \{hh^{-(1-\alpha)} + (2h)^\alpha\} \\ &= A_2(1 + 2^\alpha)h^\alpha. \end{aligned}$$

Also ist

$$\omega(f; h) \leq A_2(1 + 2^\alpha)h^\alpha$$

und somit $f \in C_{2\pi}^\alpha$. \square

Wegen der Sätze IV.3.3 (S. 62) und IV.5.1 ist für $\alpha \in (0, 1)$ eine Funktion f in $C_{2\pi}^\alpha$ genau dann, wenn $E_n(f) = O(n^{-\alpha})$ ist. Das folgende Beispiel zeigt, dass sich diese Äquivalenz nicht auf den Fall $\alpha = 1$ übertragen lässt.

BEISPIEL IV.5.2. Durch die gleichmäßig konvergente Reihe

$$\sum_{k \geq 1} \frac{\sin(k\varphi)}{k^2}$$

wird eine stetige 2π -periodische Funktion g definiert. Offensichtlich ist

$$\begin{aligned} E_n(g) &\leq \left\| \sum_{k=n+1}^{\infty} \frac{\sin(k\varphi)}{k^2} \right\| \leq \sum_{k=n+1}^{\infty} \frac{1}{k^2} \leq \sum_{k=n+1}^{\infty} \left(\frac{1}{k-1} - \frac{1}{k} \right) \\ &= \frac{1}{n}. \end{aligned}$$

Andererseits erhalten wir für $\frac{\pi}{4m} < \varphi < \frac{\pi}{2m}$

$$\begin{aligned} g(\varphi) &= \sum_{k=1}^{m-1} \underbrace{\frac{\sin(k\varphi)}{k^2}}_{\geq \frac{1}{k^2} \frac{2}{\pi} k\varphi} + \sum_{k=m}^{\infty} \underbrace{\frac{\sin(k\varphi)}{k^2}}_{\leq \frac{1}{k^2}} \\ &\geq \sum_{k=1}^{m-1} \frac{2}{\pi} \frac{\varphi}{k} - \sum_{k=m}^{\infty} \frac{1}{k^2} \\ &\geq \sum_{k=1}^{m-1} \frac{2\varphi}{\pi} \frac{1}{k} - \frac{1}{m} \\ &\geq \frac{2}{\pi} \varphi \ln m - \frac{1}{m} \\ &\geq \frac{2}{\pi} \varphi [\ln \varphi^{-1} - 2]. \end{aligned}$$

Also ist $\frac{g(\varphi) - g(0)}{\varphi - 0} = \frac{g(\varphi)}{\varphi}$ für $\varphi \rightarrow 0$ nicht beschränkt und somit $g \notin \text{Lip } 1$.

BEMERKUNG IV.5.3. Für $f \in C_{2\pi}$ gelte $E_n(f) = O\left(\frac{1}{n}\right)$. Dann ist $\omega(f; h) = O\left(h \ln \frac{1}{h}\right)$.

BEWEIS. Im Beweis von Satz IV.5.1 haben wir

$$|f(\varphi) - f(\psi)| \leq A_1 |\varphi - \psi| \sum_{n=0}^{m-1} 2^{n-\alpha n} + 2A_1 \sum_{n=m}^{\infty} 2^{-\alpha n}$$

gezeigt. Diese Abschätzung bleibt auch für $\alpha = 1$ gültig und liefert

$$|f(\varphi) - f(\psi)| \leq A_1 |\varphi - \psi| m + 2A_1 2^{1-m}.$$

Für $|\varphi - \psi| \leq h < 1$ und $m \in \mathbb{N}$ mit $\frac{1}{2h} \leq 2^m < \frac{1}{h}$ folgt

$$|f(\varphi) - f(\psi)| \leq A_1 h \ln h^{-1} [\ln 2]^{-1} + 8A_1 h \leq 8A_1 h [\ln h^{-1} + 1].$$

Dies beweist die Behauptung. \square

Von Zygmund stammt die Erkenntnis, dass man für den Fall $\alpha = 1$ auf den zweiten Stetigkeitsmodul zurückgreifen muss.

SATZ IV.5.4 (Zygmund). Für $f \in C_{2\pi}$ ist $E_n(f) = O\left(\frac{1}{n}\right)$ genau dann, wenn $\omega_2(f; h) = O(h)$ ist.

BEWEIS. „ \Rightarrow “: Wir benutzen die Notationen des Beweises von Satz IV.5.1. Aus dem Mittelwertsatz und Satz IV.4.4 (S. 69) folgt für jedes n

$$\begin{aligned} |v_n(\varphi + h) - 2v_n(\varphi) + v_n(\varphi - h)| &= h^2 |v_n''(\eta)| \\ &\leq h^2 \|v_n''\| \\ &\leq h^2 2^{2n} \|v_n\|. \end{aligned}$$

Da andererseits

$$|v_n(\varphi + h) - 2v_n(\varphi) + v_n(\varphi - h)| \leq 4 \|v_n\|$$

ist, erhalten wir wegen $A_1 = A(1 + 2^\alpha)$

$$\begin{aligned} &|f(\varphi + h) - 2f(\varphi) + f(\varphi - h)| \\ &\leq \sum_{n=0}^{\infty} |v_n(\varphi + h) - 2v_n(\varphi) + v_n(\varphi - h)| \\ &\leq \sum_{n=0}^{m-1} h^2 2^{2n} \underbrace{\|v_n\|}_{\leq 3A 2^{-n}} + \sum_{n=m}^{\infty} 4 \underbrace{\|v_n\|}_{\leq 3A 2^{-n}} \\ &\leq 3Ah^2 \sum_{n=0}^{m-1} 2^n + 12A \sum_{n=m}^{\infty} 2^{-n} \\ &\leq 3A \{h^2 2^m + 8 \cdot 2^{-m}\}. \end{aligned}$$

Bestimmen wir m wieder so, dass $\frac{1}{2h} \leq 2^m < \frac{1}{h}$ ist, folgt

$$|f(\varphi + h) - 2f(\varphi) + v(\varphi - h)| \leq 3A \{h^2 h^{-1} + 8 \cdot 2h\} \leq 51Ah.$$

Also ist $\omega_2(f; h) = O(h)$.

„ \Leftarrow “: Im Beweis von Satz IV.3.3 (S. 62) haben wir bereits die Abschätzung

$$|f(\varphi) - J_n f(\varphi)| \leq \int_0^\pi \omega_2(f; t) K_{n'}(t) dt$$

bewiesen. Aus Satz IV.2.8(3) (S. 58) und Bemerkung IV.3.2(3) (S. 60) folgt

$$\begin{aligned} & |f(\varphi) - J_n f(\varphi)| \\ & \leq \int_0^\pi \omega_2(f; \frac{1}{n} nt) K_{n'}(t) dt \\ & \leq \omega_2(f; \frac{1}{n}) \int_0^\pi (1 + nt)^2 K_{n'}(t) dt \\ & \leq \omega_2(f; \frac{1}{n}) \left\{ \frac{1}{2} + n \frac{2}{5} \pi^2 \underbrace{\frac{1}{n'}}_{\leq \frac{2}{n}} + n^2 \frac{8}{15} \pi^3 \underbrace{\frac{1}{(n')^2}}_{\leq \frac{4}{n^2}} \right\} \\ & \leq \omega_2(f; \frac{1}{n}) \left\{ \frac{1}{2} + \frac{4}{5} \pi^2 + \frac{32}{15} \pi^3 \right\}. \end{aligned}$$

Also ist $E_n(f) = O(\frac{1}{n})$. \square

SATZ IV.5.5. *Seien $k \in \mathbb{N}$, $\alpha \in (0, 1)$ und $f \in C_{2\pi}$. Gilt $E_n(f) = O(n^{-k-\alpha})$, so ist $f \in C_{2\pi}^{k+\alpha}$.*

BEWEIS. Mit den Notationen des Beweises von Satz IV.5.1 gilt

$$\|v_n\| \leq A_1 2^{-n(k+\alpha)} \quad \text{mit} \quad A_1 = A(1 + 2^{k+\alpha}).$$

Aus Satz IV.4.4 (S. 69) folgt

$$\|v_n^{(\ell)}\| \leq A_1 2^{-n(k+\alpha-\ell)}.$$

Also konvergieren die Reihen $\sum v_n^{(\ell)}$ für $\ell = 0, 1, \dots, k$ gleichmäßig und absolut. Also ist $f \in C_{2\pi}^{k+\alpha}$ und

$$f^{(\ell)} = \sum_{n=0}^{\infty} v_n^{(\ell)} \quad \text{für } \ell = 0, 1, \dots, k.$$

Mit dem Beweis von Satz IV.5.1 mit $f^{(k)}$ und $v_n^{(k)}$ an Stelle von f und v_n folgt sofort $f^{(k)} \in C_{2\pi}^\alpha$. \square

BEMERKUNG IV.5.6. Sei $f \in C_{2\pi}$ und $\kappa > 1$. Man kann zeigen, dass f genau dann in den Streifen $\{z \in \mathbb{C} : |\operatorname{Im} z| < \ln \kappa\}$ analytisch fortsetzbar ist, wenn gilt $\limsup_{n \rightarrow \infty} \sqrt[n]{E_n(f)} \leq \frac{1}{\kappa}$.

Wir wollen die Umkehrsätze nun auf algebraische Polynome ausdehnen. Das folgende Beispiel zeigt, dass wir dabei nur Aussagen im Innern des Intervalls gewinnen können.

BEISPIEL IV.5.7. Betrachte $f(x) = \sqrt{\frac{1+x}{2}}$ in $[-1, 1]$ und die Transformation $\frac{1+x}{2} = \eta^2$ von $[-1, 1]$ auf $[0, 1]$. Jedem Polynom in x entspricht ein Polynom in η^2 durch $p(x) = p(-1 + 2\eta^2)$. Also ist die Approximation von f in \mathbb{P}_n auf $[-1, 1]$ äquivalent zur Approximation von $g(\eta) = f(-1 + 2\eta^2) = \eta$ auf $[0, 1]$ durch Polynome vom Grad $\leq n$ in η^2 . Wegen des Invarianzprinzips ist dieses äquivalent zur Approximation von $g(\eta) = |\eta|$ auf $[-1, 1]$ in \mathbb{P}_{2n} . Wegen $g \in \text{Lip } 1$ folgt aus den Jackson-Sätzen $E_n(f) = E_{2n}(g) = O\left(\frac{1}{n}\right)$. Andererseits ist $f \in \text{Lip } \frac{1}{2}$ aber $f \notin \text{Lip } \alpha$ für $\alpha > \frac{1}{2}$. Also kann man aus $E_n(f) = O(n^{-\alpha})$ nicht auf $f \in C^\alpha([-1, 1])$ schließen, und das Problem liegt offensichtlich an möglichen Randsingularitäten.

SATZ IV.5.8. Seien $k \in \mathbb{N}$, $\alpha \in (0, 1)$ und $f \in C([-1, 1])$ mit $E_n(f) = O(n^{-k-\alpha})$. Für jedes Teilintervall $[a, b]$ mit $-1 < a < b < 1$ gilt dann $f \in C^{k+\alpha}([a, b])$.

BEWEIS. Für die Ableitung eines Polynomes vom Grad n gilt gemäß Satz IV.4.5

$$\left| \sqrt{1-x^2} p'(x) \right| \leq n \|p\|.$$

Mit

$$c = \max \left\{ (1-a^2)^{-\frac{1}{2}}, (1-b^2)^{-\frac{1}{2}} \right\} = \left\{ \min \left\{ \sqrt{1-a^2}, \sqrt{1-b^2} \right\} \right\}^{-1}$$

folgt hieraus

$$|p'(x)| \leq cn \|p\| \quad \forall x \in [a, b].$$

Da c unabhängig von n ist, lassen sich alle vorherigen Beweise von der trigonometrischen Approximation übertragen. \square

BEMERKUNG IV.5.9. (1) Wenn man bei der Approximation durch algebraische Polynome zugleich notwendige und hinreichende Bedingungen erhalten will, muss man das Konzept abändern und Approximationen betrachten, die zum Rand hin besser sind. Setze dazu

$$\Delta_n(t) = \max \left\{ \frac{1}{n^2}, \frac{1}{n} \sqrt{1-t^2} \right\}.$$

Dann kann man zeigen, dass folgende Aussagen äquivalent sind:

- Zu $f \in C([-1, 1])$ existiert eine Folge $(p_n)_{n \in \mathbb{N}}$ mit $p_n \in \mathbb{P}_n$ und

$$|f(x) - p_n(x)| \leq A \Delta_n(x)^\alpha \quad \forall x \in [-1, 1]$$

mit $A \in \mathbb{R}_+$ und $\alpha \in (0, 1]$.

- Es ist $f \in \text{Lip } \alpha$, falls $\alpha < 1$ ist, und $\omega_2(f; h) = O(h)$, falls $\alpha = 1$ ist.

(2) Statt des Stetigkeitsmoduls kann man das *Petree-Funktional*

$$K(f; h) = \inf_{g \in C^1} \{ \|f - g\| + h \|g'\| \}$$

verwenden. Man kann zeigen, dass für alle $f \in C_{2\pi}$ gilt

$$\frac{1}{2}\omega(f; h) \leq K(f; h) \leq 2\omega(f; h).$$

In diesem Sinne sind Petree-Funktional und Stetigkeitsmodul äquivalent.

IV.6. Resultate negativen Charakters

Statt approximierende Polynome mit Bernstein- oder Jackson-Operatoren zu konstruieren, könnte man versuchen, solche Polynome durch Interpolation zu gewinnen. Dazu seien für jedes $n \in \mathbb{N}$ paarweise verschiedene Punkte $t_{n,0}, \dots, t_{n,n}$ in $I = [a, b]$ gegeben. Zu $f \in C([a, b])$ bezeichne dann $P_n f \in \mathbb{P}_n$ das Lagrangesche Interpolationspolynom zu den Knoten $t_{n,0}, \dots, t_{n,n}$. Nach dem Satz von Faber gibt es zu jeder solchen Knotenmatrix ein $f \in C([a, b])$, für das $P_n f$ nicht gegen f konvergiert. Also ist die Lagrangesche Interpolation zur Funktionsapproximation weniger geeignet als die Čebyšev-Approximation.

Wir wollen dieses negative Resultat im vorliegenden Abschnitt näher untersuchen. Die Beobachtung, dass jedes P_n ein stetiger linearer Operator von $C([a, b])$ in \mathbb{P}_n ist, der \mathbb{P}_n invariant lässt, führt auf folgende Begriffsbildung. Dabei ist $\|\cdot\|$ stets die Maximumnorm auf $C(I)$ oder $C_{2\pi}$ und $\|\!\|\!\|\cdot\|\!\|\!\|$ die zugehörige Operatornorm.

DEFINITION IV.6.1. Sei V ein linearer Unterraum des normierten Raumes X . Ein linearer Operator $L : X \rightarrow V$ heißt *Projektion* oder *Projektionsoperator*, wenn er folgende Eigenschaften hat:

- (1) L ist stetig.
- (2) Es ist $Lf = f$ für alle $f \in V$.

Projektionen von $C(I)$ bzw. $C_{2\pi}$ auf \mathbb{P}_n bzw. \mathbb{T}_n werden auch als *Polynomoperatoren* vom Grad n bezeichnet.

BEISPIEL IV.6.2. (1) Sei P_n der Lagrangesche Interpolationsoperator zu den Knoten t_0, \dots, t_n . P_n ist ein Projektionsoperator.

(2) Sei $S_n : C_{2\pi} \rightarrow \mathbb{T}_n$ der *Fourier-Projektor*, der durch

$$S_n f(\varphi) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(k\varphi) + b_k \sin(k\varphi))$$

mit

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt) dt, \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt) dt$$

definiert ist. Wegen der Orthogonalitätsrelationen ist dies ein Projektionsoperator.

(3) Die Bernstein-Operatoren B_n aus § IV.1 sind keine Projektionsoperatoren, da sie die quadratischen Polynome nicht invariant lassen.

SATZ IV.6.3. Für jeden Polynomoperator $L : C_{2\pi} \rightarrow \mathbb{T}_n$ gilt

$$\|\!\|\!\|L\|\!\|\!\| \geq \|\!\|\!\|S_n\|\!\|\!\|.$$

BEWEIS. Für festes s definiert

$$\begin{aligned} T_s : C_{2\pi} &\rightarrow C_{2\pi} \\ f(\cdot) &\mapsto f(\cdot + s) \end{aligned}$$

einen Translationsoperator. Für jedes φ und jedes $f \in C_{2\pi}$ ist die Abbildung

$$s \mapsto (T_{-s}LT_s f)(\varphi)$$

stetig. Daher existiert

$$\Phi f(\varphi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (T_{-s}LT_s f)(\varphi) ds.$$

Offensichtlich ist $\Phi : C_{2\pi} \rightarrow C_{2\pi}$ ein linearer Operator. Wegen $\|T_s\| = 1$ folgt

$$\|\Phi\| \leq \|L\|.$$

Wir wollen zeigen, dass $\Phi = S_n$ ist. Definiere dazu für $k \in \mathbb{Z}$

$$u_k(\varphi) = e^{ik\varphi}.$$

Es ist

$$T_s u_k(\varphi) = e^{ik(\varphi+s)} = e^{iks} u_k(\varphi)$$

und somit

$$\begin{aligned} (LT_s u_k)(\varphi) &= e^{iks} (Lu_k)(\varphi), \\ (T_{-s}LT_s u_k)(\varphi) &= e^{iks} (Lu_k)(\varphi - s). \end{aligned}$$

Ist $|k| \leq n$, so ist $Lu_k = u_k$ und somit

$$(T_{-s}LT_s u_k)(\varphi) = e^{iks} u_k(\varphi - s) = u_k(\varphi).$$

Also ist für alle $k \in \mathbb{Z}$ mit $|k| \leq n$

$$\Phi u_k = u_k = S_n u_k.$$

Ist $|k| > n$, so ist $Lu_k(\varphi - s)$ ein trigonometrisches Polynom q vom Grad $\leq n$ in s . Aus den Orthogonalitätsrelationen folgt

$$\int_{-\pi}^{\pi} (T_{-s}LT_s u_k)(\varphi) ds = \int_{-\pi}^{\pi} e^{iks} q(s) ds = 0.$$

Also ist für $k \in \mathbb{Z}$ mit $|k| > n$

$$\Phi u_k = 0 = S_n u_k.$$

Also gilt

$$\Phi p = S_n p$$

für jedes trigonometrische Polynom beliebigen Grades. Da diese in $C_{2\pi}$ dicht sind, folgt $\Phi = S_n$ und damit die Behauptung. \square

Als nächstes wollen wir $\|S_n\|$ abschätzen. Zu Beginn von §IV.3 haben wir bereits gezeigt, dass

$$\begin{aligned} a_k \cos(k\varphi) + b_k \sin(k\varphi) &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(k(\varphi - t)) dt \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(\varphi - t) \cos(kt) dt \end{aligned}$$

ist. Hieraus folgt

$$S_n f(\varphi) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\varphi - t) \left\{ \frac{1}{2} + \sum_{k=1}^n \cos(kt) \right\} dt.$$

Für festes t ist aber

$$\begin{aligned} \frac{1}{2} + \sum_{k=1}^n \cos(kt) &= \frac{1}{2} \left[1 + \sum_{k=1}^n (e^{ikt} + e^{-ikt}) \right] \\ &= \frac{1}{2} \sum_{k=-n}^n e^{ikt} \\ &= \frac{1}{2} e^{-int} \sum_{k=0}^{2n} e^{ikt} \\ &= \frac{1}{2} e^{-int} \frac{e^{i(2n+1)t} - 1}{e^{it} - 1} \\ &= \frac{1}{2} e^{-i(n+\frac{1}{2})t} \frac{e^{i2(n+\frac{1}{2})t} - 1}{e^{i\frac{t}{2}} - e^{-i\frac{t}{2}}} \\ &= \frac{1}{2} \frac{\sin(n + \frac{1}{2})t}{\sin \frac{t}{2}}. \end{aligned}$$

Also ist

$$S_n f(\varphi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\varphi - t) D_n(t) dt$$

mit

$$D_n(t) = \frac{\sin(n + \frac{1}{2})t}{\sin \frac{t}{2}}.$$

SATZ IV.6.4. Für die Norm des Fourier-Projektors gilt

$$\lambda_n = \|S_n\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)| dt = \frac{4}{\pi^2} \ln n + O(1).$$

BEWEIS. Aus der Darstellung von S_n folgt sofort

$$\|S_n\| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)| dt.$$

Durch Glätten der Funktion

$$f(t) = \operatorname{sgn} D_n(-t)$$

und Abschätzung von $S_n f(0)$ erkennt man, dass $\|S_n\|$ nicht kleiner als $\frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)| dt$ sein kann.

Da D_n eine gerade Funktion ist, erhalten wir

$$\begin{aligned}\lambda_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_n(t)| dt \\ &= \frac{1}{\pi} \int_0^{\pi} |D_n(t)| dt & z = \frac{t}{2} \\ &= \frac{2}{\pi} \int_0^{\frac{\pi}{2}} |D_n(2z)| dz \\ &= \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{|\sin(2n+1)z|}{\sin z} dz.\end{aligned}$$

Aus der Taylorschen Formel folgt

$$\sin z - z = -\frac{z^2}{2} \sin \eta$$

für ein η mit $0 < \eta < z \leq \frac{\pi}{2}$. Also ist

$$|\sin z - z| \leq \frac{z^2}{2} \sin z \quad \text{bzw.} \quad \left| \frac{1}{z} - \frac{1}{\sin z} \right| \leq \frac{z}{2}.$$

Hieraus folgt wegen $\sin z \leq z$

$$\begin{aligned}\lambda_n &= \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{|\sin(2n+1)z|}{\sin z} dz \\ &\geq \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{|\sin(2n+1)z|}{z} dz\end{aligned}$$

und

$$\begin{aligned}\lambda_n &= \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{|\sin(2n+1)z|}{\sin z} dz \\ &\leq \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{|\sin(2n+1)z|}{z} dz + \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{z}{2} \underbrace{|\sin(2n+1)z|}_{\leq 1} dz \\ &\leq \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{|\sin(2n+1)z|}{z} dz + \frac{\pi}{8}.\end{aligned}$$

Als nächstes schätzen wir

$$\mu_n = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{|\sin(2n+1)z|}{z} dz$$

ab. Aus der Monotonie der Funktion $x \mapsto \frac{1}{x}$ erhalten wir

$$\sum_{k=2}^m \frac{1}{k} \leq \int_1^m \frac{1}{x} dx = \ln m \leq \sum_{k=1}^m \frac{1}{k}.$$

Mit der Transformation $(2n+1)z = x$ führt dies auf die Abschätzung

$$\begin{aligned}
\mu_n &= \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \frac{|\sin(2n+1)z|}{z} dz \\
&= \frac{2}{\pi} \int_0^{(2n+1)\frac{\pi}{2}} \frac{|\sin x|}{x} dx \\
&\geq \frac{2}{\pi} \sum_{k=1}^n \int_{(k-1)\pi}^{k\pi} \frac{|\sin x|}{x} dx \\
&\geq \frac{2}{\pi} \sum_{k=1}^n \frac{1}{k\pi} \int_{(k-1)\pi}^{k\pi} |\sin x| dx \\
&= \frac{4}{\pi^2} \sum_{k=1}^n \frac{1}{k} \\
&\geq \frac{4}{\pi^2} \ln n
\end{aligned}$$

und

$$\begin{aligned}
\mu_n &= \frac{2}{\pi} \int_0^{(2n+1)\frac{\pi}{2}} \frac{|\sin x|}{x} dx \\
&\leq \frac{2}{\pi} \int_0^{\pi} \underbrace{\frac{|\sin x|}{x}}_{\leq 1} dx + \frac{2}{\pi} \sum_{k=2}^{n+1} \int_{(k-1)\pi}^{k\pi} \frac{|\sin x|}{x} dx \\
&\leq 2 + \frac{2}{\pi} \sum_{k=2}^{n+1} \frac{1}{(k-1)\pi} \int_{(k-1)\pi}^{k\pi} |\sin x| dx \\
&= 2 + \frac{4}{\pi^2} \sum_{k=1}^n \frac{1}{k} \\
&\leq 2 + \frac{4}{\pi^2} + \frac{4}{\pi^2} \ln n.
\end{aligned}$$

Dies beweist die Behauptung. \square

Um aus $\lim_{n \rightarrow \infty} \|S_n\| = \infty$ auf die Resultate negativen Charakters schließen zu können, benötigen wir das Prinzip der gleichmäßigen Beschränktheit.

SATZ IV.6.5 (Lemma von Baire). *Der Banach-Raum X sei darstellbar als Vereinigung von abzählbar vielen abgeschlossenen Mengen A_n , d.h. $X = \bigcup_{n \in \mathbb{N}} A_n$. Dann enthält mindestens ein A_n eine abgeschlossene Kugel mit positivem Radius.*

BEWEIS. Die Mengen $O_n = X \setminus A_n$ sind offen und haben nach Voraussetzung einen leeren Durchschnitt. Wir nehmen an, die Behauptung sei falsch. Dann ist der Durchschnitt von jedem O_n mit jeder abgeschlossenen Kugel mit positivem Radius nicht leer. Wir konstruieren

nun eine Folge abgeschlossener Kugeln $B_n = \overline{B(x_n; r_n)}$ mit folgenden Eigenschaften:

$$\begin{aligned} B_1 &\subset O_1, \\ B_n &\subset O_n \cap B_{n-1} \quad \forall n \geq 1, \\ 0 < r_n &< 2^{-n} \quad \forall n \geq 0. \end{aligned}$$

Da O_1 offen und nicht leer ist, existiert ein B_1 mit den gewünschten Eigenschaften. Seien B_1, \dots, B_{n-1} konstruiert. Da $O_n \cap \overset{\circ}{B}_{n-1}$ offen und nicht leer ist, existiert ein B_n mit den gewünschten Eigenschaften.

Für die Mittelpunkte der Kugeln B_n gilt $x_m \in B_m \subset B_n$ für alle $m \geq n$ also $\|x_m - x_n\| \leq r_n$ für alle $m \geq n$. Also ist $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge. Da X vollständig ist, konvergiert $(x_n)_{n \in \mathbb{N}}$ gegen ein $x^* \in X$. Für dieses x^* gilt aber $x^* \in B_n \subset O_n$ für alle n und somit $x^* \in \bigcap_{n \in \mathbb{N}} O_n$. Dies ist ein Widerspruch. \square

SATZ IV.6.6 (Prinzip der gleichmäßigen Beschränktheit). *Sei $(L_n)_{n \in \mathbb{N}}$ eine Folge stetiger linearer Operatoren eines Banach-Raumes X in einen normierten Raum Y . Für jedes $f \in X$ gelte $\sup_n \|L_n f\| < \infty$. Dann ist auch $\sup_n \|L_n\| < \infty$.*

BEWEIS. Für $k \in \mathbb{N}^*$ sei

$$A_k = \left\{ f \in X : \sup_n \|L_n f\| \leq k \right\}.$$

Die Mengen A_k sind abgeschlossen und nach Voraussetzung ist $X = \bigcup_k A_k$. Gemäß Satz IV.6.5 gibt es ein m , so dass A_m eine Kugel $\overline{B}(g; r)$ mit $r > 0$ enthält. Sei $f \in X$ mit $\|f\| = 1$ beliebig. Dann ist $g + rf \in \overline{B}(g; r) \subset A_m$ und daher für jedes $n \in \mathbb{N}$

$$\begin{aligned} \|L_n f\| &= \left\| L_n \left(\frac{1}{r}(g + rf) \right) - L_n \left(\frac{1}{r}g \right) \right\| \\ &\leq \frac{1}{r} \|L_n(g + rf)\| + \frac{1}{r} \|L_n g\| \\ &\leq \frac{1}{r} \sup_k \|L_k(g + rf)\| + \frac{1}{r} \sup_k \|L_k g\| \\ &\leq \frac{2m}{r}. \end{aligned}$$

Da f mit $\|f\| = 1$ beliebig war, folgt $\|L_n\| \leq \frac{2m}{r}$. Da n beliebig war, folgt $\sup_n \|L_n\| \leq \frac{2m}{r}$. \square

SATZ IV.6.7 (Erster Satz von Haršiladze-Lozinski). *Zu jeder Folge von Polynomoperatoren L_n , die $C_{2\pi}$ in \mathbb{T}_n abbilden, gibt es ein $f \in C_{2\pi}$, für das die Folge $(L_n f)_{n \in \mathbb{N}}$ divergiert.*

BEWEIS. Angenommen, die Behauptung sei falsch. Dann ist für jedes $f \in C_{2\pi}$ die Folge $(\|L_n f\|)_{n \in \mathbb{N}}$ konvergent und damit beschränkt. Nach Satz IV.6.6 ist daher $\sup_n \|L_n\| < \infty$. Dies ist ein Widerspruch zu Satz IV.6.3 und IV.6.4. \square

BEMERKUNG IV.6.8. Das Beweisprinzip von Satz IV.6.7 benutzt man auch, um zu zeigen, dass die Newton-Cotes-Formeln zur numerischen Integration nicht für jeden stetigen Integranden konvergieren. Ist nämlich

$$Q_n f = \sum_{k=0}^n a_{n,k} f(t_{n,k})$$

eine Folge von Newton-Cotes-Formeln zur Berechnung von

$$\int_a^b \omega(x) f(x) dx,$$

so zeigt man, dass

$$\| \| Q_n \| \| = \sum_{k=0}^n |a_{n,k}|$$

ist und dass dieser Ausdruck für $n \rightarrow \infty$ unbeschränkt ist. Damit folgt die Behauptung wieder aus Satz IV.6.5.

Wir wenden uns nun Polynomoperatoren für algebraische Polynome zu. Ist $L_n : C([-1, 1]) \rightarrow \mathbb{P}_n$ ein Polynomoperator vom Grad n , so induziert er durch die Transformation $g(\varphi) = f(\cos \varphi)$ einen Polynomoperator \tilde{L}_n von den geraden 2π -periodischen Funktionen in die geraden trigonometrischen Polynome vom Grad $\leq n$, der diese Polynome invariant lässt. Da unter der Transformation $[-1, 1] \ni x \mapsto \cos \varphi$, $\varphi \in [-\pi, \pi]$ die Maximumnorm unverändert bleibt, ist $\| \| \tilde{L}_n \| \| = \| \| L_n \| \|$.

SATZ IV.6.9. Sei L ein Projektionsoperator von den geraden Funktionen in $C_{2\pi}$ auf die geraden trigonometrischen Polynome vom Grad $\leq n$. Dann ist

$$\| \| L \| \| \geq \frac{1}{2}(\lambda_n - 3)$$

mit λ_n aus Satz IV.6.4.

BEWEIS. Wir zeigen

$$\| \| I - L \| \| \geq \frac{1}{2}(\lambda_n - 1).$$

Hieraus folgt die Behauptung mit der Dreiecksungleichung. Zum Nachweis obiger Abschätzung betrachten wir den Operator

$$\Phi f(\varphi) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [T_s(I - L)(T_{-s} + T_s)f](\varphi) ds$$

mit dem Translationsoperator T_s aus dem Beweis von Satz IV.6.4 und zeigen $\Phi = I - S_n$. Zum Nachweis dieser Identität betrachten wir die Funktionen

$$u_k(\varphi) = \cos(k\varphi).$$

Aus den Additionstheoremen folgt

$$\begin{aligned}(T_{-s} + T_s)u_k(\varphi) &= \cos(k(\varphi - s)) + \cos(k(\varphi + s)) \\ &= 2 \cos(k\varphi) \cos(ks) \\ &= 2 \cos(ks)u_k(\varphi).\end{aligned}$$

Für $0 \leq k \leq n$ folgt aus dieser Identität und

$$(I - L)u_k = u_k - Lu_k = u_k - u_k = 0$$

sofort

$$\Phi u_k = 0 = (I - S_n)u_k.$$

Sei nun $k > n$. Da Lu_k ein Polynom p vom Grad $\leq n$ ist, folgt aus den Orthogonalitätsrelationen

$$\begin{aligned}\int_{-\pi}^{\pi} [T_s L(T_{-s} + T_s)u_k](\varphi) ds &= \int_{-\pi}^{\pi} 2 \cos(ks)p(\varphi + s) ds \\ &= \int_{-\pi}^{\pi} 2 \cos(k(t - \varphi))p(t) dt \\ &= 0.\end{aligned}$$

Daher ist

$$\begin{aligned}\Phi u_k(\varphi) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} [T_s(T_{-s} + T_s)u_k](\varphi) ds \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} 2 \cos(ks) \cos(k(\varphi + s)) ds \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} 2 \cos(ks) [\cos(k\varphi) \cos(ks) - \sin(k\varphi) \sin(ks)] ds \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(ks) \cos(ks) \cos(k\varphi) ds \\ &= \cos(k\varphi) \\ &= u_k(\varphi).\end{aligned}$$

Also ist auch in diesem Fall

$$\Phi u_k = (I - S_n)u_k.$$

Daher stimmen Φ und $I - S_n$ auf der Menge der geraden trigonometrischen Polynome beliebigen Grades überein. Da diese in der Menge der geraden stetigen 2π -periodischen Funktionen dicht ist, folgt $\Phi = I - S_n$. Da die Operatornorm von S_n nicht kleiner wird, wenn man sich auf die

geraden Funktionen in $C_{2\pi}$ beschränkt, folgt hieraus

$$\begin{aligned}\lambda_n &= \||S_n\|| \\ &= \||I - \Phi\|| \\ &\leq 1 + \||\Phi\|| \\ &\leq 1 + \||T_s\|| \||I - L\|| \{\||T_s\|| + \||T_{-s}\|\|\} \\ &= 1 + 2 \||I - L\||.\end{aligned}$$

Dies beweist die Behauptung. \square

SATZ IV.6.10 (Zweiter Satz von Haršiladze-Lozinski). *Für jede Folge von Polynomoperatoren $L_n : C([-1, 1]) \rightarrow \mathbb{P}_n$ gibt es ein $f \in C([-1, 1])$, für das die Folge $(L_n f)_{n \in \mathbb{N}}$ nicht gegen f konvergiert.*

BEWEIS. Wegen Satz IV.6.9 ist die Folge $(\||L_n\||)_{n \in \mathbb{N}} = (\||\tilde{L}_n\||)_{n \in \mathbb{N}}$ unbeschränkt. Damit folgt die Behauptung aus dem Prinzip der gleichmäßigen Beschränktheit. \square

Trotz dieser negativen Ergebnisse kann man mit Projektionsoperatoren vernünftige Approximationen erreichen, wenn man zusätzliche Glattheitsbedingungen fordert.

SATZ IV.6.11. *Sei L ein Polynomoperator vom Grad n . Dann ist für jedes f*

$$\|f - Lf\| \leq \||I - L\|| E_n(f).$$

BEWEIS. Sei p die b.A. an f in \mathbb{P}_n bzw. \mathbb{T}_n . Dann folgt

$$\begin{aligned}\|f - Lf\| &= \left\| \underbrace{f - p + Lp}_{=0} - Lf \right\| \\ &= \|(I - L)(f - p)\| \\ &\leq \||I - L\|| \|f - p\| \\ &\leq \||I - L\|| E_n(f).\end{aligned}$$
 \square

Aus diesem Satz folgt das berühmte Dini-Lipschitz-Theorem.

SATZ IV.6.12 (Dini-Lipschitz-Theorem). *Für $f \in C_{2\pi}$ gelte*

$$\omega(f; h) \ln h^{-1} = o(1) \quad \text{für } h \rightarrow 0.$$

Dann konvergiert die Folge $(S_n f)_{n \in \mathbb{N}}$ der Partialsummen der Fourier-Reihe von f gleichmäßig gegen f .

BEWEIS. Die Voraussetzungen garantieren nach den Jackson-Sätzen und Satz IV.6.4 gerade

$$\lim_{n \rightarrow \infty} \||I - S_n\|| E_n(f) = 0.$$

Damit folgt die Behauptung aus Satz IV.6.11. \square

BEMERKUNG IV.6.13. Betrachtet man die Lagrange-Interpolation in den Extremalpunkten der Čebyšev-Polynome, kann man zeigen, dass die Norm der Interpolationsoperatoren $\frac{2}{\pi} \ln n + O(1)$ ist. Damit überträgt sich Satz [IV.6.12](#) direkt auf diese Operatoren.

KAPITEL V

Numerische Behandlung der Čebyšev-Approximation

Zur numerischen Behandlung der Čebyšev-Approximation stehen im wesentlichen zwei Verfahrensklassen zur Auswahl: *Descent-Verfahren* und *Ascent-Verfahren*.

Bei den Descent-Verfahren wird das Approximationsproblem wie in der Einleitung zu Kapitel II als semi-infinites Optimierungsproblem aufgefasst und durch Verfahren, die dem Simplexalgorithmus ähneln, eine Folge von Näherungen erzeugt, deren Fehlernormen monoton fallend gegen das Optimum konvergieren.

Bei den Ascent-Verfahren macht man sich das Alternantenkriterium und den Satz von de la Vallée-Poussin zunutze und konstruiert Folgen von Referenzen, so dass die zugehörigen Interpolationen eine monoton wachsende Folge von Fehlernormen erzeugen, die gegen das Optimum konvergieren. Diese Verfahren sind den Descent-Verfahren weit überlegen. Ihr wichtigster Vertreter ist der Remez-Algorithmus, den wir in § V.2 kennenlernen. Da dieser das Approximationsproblem auf einem Intervall auf eine Folge von Approximationsproblemen auf diskreten Punktmengen reduziert, betrachten wir zunächst derartige diskrete Approximationsprobleme.

V.1. Satz von Stiefel

SATZ V.1.1. Sei V ein n -dimensionaler Haarscher Unterraum von $C([a, b])$. Zu $n + 1$ Punkten $a \leq t_0 < \dots < t_n \leq b$ sei $T = \{t_0, \dots, t_n\}$ und

$$\|f\|_T = \max_{0 \leq i \leq n} |f(t_i)|.$$

Dann ist die b.A. an f in V bzgl. $\|\cdot\|_T$ die Lösung u der Interpolationsaufgabe

$$u(t_i) = f(t_i) - (-1)^i \eta \quad 0 \leq i \leq n$$

mit passendem $\eta \in \mathbb{R}$. Außerdem ist

$$\|f - u\|_T = |\eta|.$$

BEWEIS. Gemäß Satz II.2.11 (S. 26) enthält für die b.A. u an f in V die Extremalmenge von $f - u$ mindestens $n + 1$ Punkte. Wie im Beweis des Alternantensatzes II.3.8 (S. 32) folgt, dass die Vorzeichen der Fehlerfunktion alternieren. \square

Sei v_1, \dots, v_n eine Basis von V . Dann kann man die b.A. an f in V auf T durch Lösen des linearen Gleichungssystems

$$\sum_{k=1}^n \alpha_k v_k(t_i) + (-1)^i \eta = f(t_i) \quad 0 \leq i \leq n$$

bestimmen. Zumindest bei der Approximation durch Polynome ist dieser Ansatz allenfalls für kleine n geeignet, da das lineare Gleichungssystem schlecht konditioniert ist. Stattdessen bestimmt man die Lösung u des Approximationsproblems als Lösung der Interpolationsaufgabe

$$u(t_i) = y(t_i) \quad 0 \leq i \leq n$$

mit

$$y_i = f(t_i) - (-1)^i \eta.$$

Offensichtlich stellt sich bei diesem Ansatz das Problem, die Größe η zu bestimmen. Einen Ansatz liefert das folgende Lemma.

LEMMA V.1.2. *Es gibt genau ein Funktional der Form*

$$\ell(f) = \sum_{i=0}^n \beta_i f(t_i)$$

mit den folgenden Eigenschaften

$$\operatorname{sgn} \beta_i = (-1)^i \quad 0 \leq i \leq n, \quad \sum_{i=0}^n |\beta_i| = 1, \quad \ell(u) = 0 \quad \forall u \in V.$$

Für dieses Funktional gilt

$$\eta = \sum_{i=0}^n \beta_i f(t_i).$$

BEWEIS. Wir betrachten das lineare Gleichungssystem

$$\sum_{i=0}^n (-1)^i \beta_i = 1, \quad \sum_{i=0}^n \beta_i v_k(t_i) = 0 \quad 1 \leq k \leq n$$

mit den $n+1$ Unbekannten β_0, \dots, β_n .

Seien $(\beta_0, \dots, \beta_n)$ und $(\widehat{\beta}_0, \dots, \widehat{\beta}_n)$ zwei Lösungen dieses Gleichungssystems. Wegen der Haarschen Bedingung sind alle β_i und alle $\widehat{\beta}_i$ von Null verschieden. Für das Funktional

$$\mu(g) = \sum_{i=1}^n \left(\widehat{\beta}_0 \beta_i - \beta_0 \widehat{\beta}_i \right) g(t_i)$$

folgt für alle k

$$\begin{aligned} \mu(v_k) &= \widehat{\beta}_0 \sum_{i=1}^n \beta_i v_k(t_i) - \beta_0 \sum_{i=1}^n \widehat{\beta}_i v_k(t_i) \\ &= \widehat{\beta}_0 \sum_{i=0}^n \beta_i v_k(t_i) - \beta_0 \sum_{i=0}^n \widehat{\beta}_i v_k(t_i) = 0. \end{aligned}$$

Wegen der Haarschen Bedingung ergibt sich

$$\widehat{\beta}_0 \beta_i - \beta_0 \widehat{\beta}_i = 0 \quad \forall 1 \leq i \leq n$$

d.h.

$$\widehat{\beta}_i = \frac{\widehat{\beta}_0}{\beta_0} \beta_i \quad \forall 1 \leq i \leq n.$$

Daher ist

$$\begin{aligned} 1 &= \sum_{i=0}^n (-1)^i \widehat{\beta}_i = \widehat{\beta}_0 + \sum_{i=1}^n (-1)^i \frac{\widehat{\beta}_0}{\beta_0} \beta_i \\ &= \frac{\widehat{\beta}_0}{\beta_0} \left\{ \underbrace{\beta_0 + \sum_{i=1}^n (-1)^i \beta_i}_{=1} \right\} \\ &= \frac{\widehat{\beta}_0}{\beta_0} \end{aligned}$$

und somit

$$\widehat{\beta}_i = \beta_i \quad \forall 0 \leq i \leq n.$$

Also hat das lineare Gleichungssystem höchstens eine Lösung und damit genau eine Lösung. Sei ℓ das zu dieser Lösung gehörige Funktional. Dann ist automatisch die dritte Bedingung des Lemmas erfüllt. Wie im Beweis des Alternantensatzes II.3.8 (S. 32) folgt, dass die Vorzeichen der β_i alternieren. Aus der Bedingung $\sum_i (-1)^i \beta_i = 1$ folgen dann die ersten beiden Bedingungen des Lemmas.

Sei nun $u \in V$ die b.A. an f auf T . Aus Satz V.1.1 folgt

$$\begin{aligned} 0 &= \ell(u) = \sum_{i=0}^n \beta_i u(t_i) \\ &= \sum_{i=0}^n \beta_i f(t_i) - \sum_{i=0}^n \beta_i (-1)^i \eta \\ &= \sum_{i=0}^n \beta_i f(t_i) - \eta. \end{aligned}$$

Dies beweist die Behauptung. □

Für den folgenden Satz beachte man, dass $\dim \mathbb{P}_n = n + 1$ ist.

SATZ V.1.3 (Stiefel). Sei $t_0 < \dots < t_{n+1}$ und

$$\|f\|_T = \max_{0 \leq i \leq n+1} |f(t_i)|.$$

Dann ist die b.A. an f in \mathbb{P}_n bzgl. $\|\cdot\|_T$ die Lösung p der Interpolationsaufgabe

$$p(t_i) = f(t_i) - (-1)^i \eta \quad 0 \leq i \leq n$$

mit

$$\eta = \frac{\sum_{i=0}^{n+1} \lambda_i f(t_i)}{\sum_{i=0}^{n+1} (-1)^i \lambda_i}$$

und

$$\lambda_i = \prod_{\substack{j=0 \\ j \neq i}}^{n+1} \frac{1}{t_i - t_j}.$$

BEWEIS. Wegen

$$\operatorname{sgn} \lambda_i = (-1)^{n+1-i}$$

und Satz V.1.1 und Lemma V.1.2 müssen wir nur noch zeigen, dass gilt

$$\sum_{i=0}^{n+1} \lambda_i q(t_i) = 0 \quad \forall q \in \mathbb{P}_n.$$

Sei dazu $q \in \mathbb{P}_n$ beliebig. Da q mit seinem Lagrangeschen Interpolationspolynom zu den Knoten t_0, \dots, t_{n+1} übereinstimmt, gilt

$$\begin{aligned} q(x) &= \sum_{i=0}^{n+1} q(t_i) \prod_{\substack{j=0 \\ j \neq i}}^{n+1} \frac{x - t_j}{t_i - t_j} \\ &= x^{n+1} \sum_{i=0}^{n+1} \lambda_i q(t_i) + r(x) \end{aligned}$$

mit $r \in \mathbb{P}_n$. Hieraus folgt die Behauptung. \square

BEMERKUNG V.1.4. In der Praxis koppelt man die Berechnung von p und η . Seien dazu q und q_0 aus \mathbb{P}_{n+1} die Lösungen der Interpolationsaufgaben

$$\begin{aligned} q(t_i) &= f(t_i) \quad 0 \leq i \leq n+1, \\ q_0(t_i) &= (-1)^i \quad 0 \leq i \leq n+1. \end{aligned}$$

Die Polynome q und q_0 schreibt man dann in der Form

$$\begin{aligned} q(x) &= \alpha x^{n+1} + r(x) \\ q_0(x) &= \alpha_0 x^{n+1} + r_0(x) \end{aligned}$$

mit $r, r_0 \in \mathbb{P}_n$. Wegen des Satzes von Rolle ist $\alpha_0 \neq 0$. Dann leisten

$$\eta = \frac{\alpha}{\alpha_0} \quad \text{und} \quad p = q - \eta q_0$$

das Gewünschte.

Für den Remez-Algorithmus benötigen wir noch eine untere Schranke für den Abstand der Referenzpunkte.

LEMMA V.1.5. Seien $f \in C([a, b])$ und $a \leq t_0 < \dots < t_n \leq b$. Für die Lösung des diskreten Approximationsproblems aus Satz V.1.1 sei $|\eta| \geq c > 0$. Dann gibt es ein $\delta > 0$, das nur von f , c und V abhängt, mit

$$t_{i+1} - t_i \geq \delta$$

für alle $0 \leq i \leq n - 1$.

BEWEIS. Wir nehmen an, die Behauptung sei falsch. Dann gibt es Folgen $(t_{0,m})_{m \in \mathbb{N}}, \dots, (t_{n,m})_{m \in \mathbb{N}}$ mit

$$a \leq t_{0,m} < \dots < t_{n,m} \leq b$$

und

$$\lim_{m \rightarrow \infty} \min_{0 \leq i \leq n-1} (t_{i+1,m} - t_{i,m}) = 0.$$

Für $m \in \mathbb{N}$ sei

$$\ell_m(g) = \sum_{i=0}^n \beta_{i,m} g(t_{i,m})$$

wie in Lemma V.1.2. Dann ist

$$\|\ell_m\| = \sum_{i=0}^n |\beta_{i,m}| = 1$$

und

$$c \leq \ell_m(f)$$

sowie

$$\ell_m(u) = 0 \quad \forall u \in V.$$

Indem wir gegebenenfalls zu einer Teilfolge übergehen, können wir

$$\lim_{m \rightarrow \infty} (t_{j+1,m} - t_{j,m}) = 0$$

für einen Index j voraussetzen. Da die Folgen $(t_{i,m})_{m \in \mathbb{N}}$ alle beschränkt sind, können wir gegebenenfalls durch Übergang zu weiteren Teilfolgen annehmen, dass

$$\lim_{m \rightarrow \infty} t_{i,m} = z_i$$

ist für alle i . Insbesondere ist $z_{j+1} = z_j$. Also gibt es ein $u \in V$ mit

$$u(z_i) = f(z_i) \quad \forall 0 \leq i \leq n.$$

Wegen der Stetigkeit von $f - u$ gibt es ein $m_0 \in \mathbb{N}$ mit

$$\max_{0 \leq i \leq n} |f(t_{i,m}) - u(t_{i,m})| < \frac{c}{2} \quad \forall m \geq m_0.$$

Damit folgt für $m \geq m_0$

$$\begin{aligned} |\ell_m(f)| &\leq \underbrace{\left| \ell_m(u) \right|}_{=0} + |\ell_m(f - u)| \\ &\leq \underbrace{\|\ell_m\|}_{=1} \max_{0 \leq i \leq n} |f(t_{i,m}) - u(t_{i,m})| \\ &\leq \frac{c}{2}. \end{aligned}$$

Wegen des Satzes von de la Vallée-Poussin ist dies ein Widerspruch zu $|\eta| \geq c$. \square

V.2. Der Remez-Algorithmus

Im Folgenden ist stets V ein n -dimensionaler Haarscher Unterraum von $C([a, b])$ und $f \in C([a, b])$. Zu $u_i \in V$ sei $\varepsilon_i = f - u_i$ die zugehörige Fehlerfunktion. Seien $a \leq t_{0,0} < \dots < t_{0,n} \leq b$ und $u_0 \in V$ die beste diskrete Approximation zu f auf diesen Punkten. Wenn diese Punkte eine Alternante von ε_0 bilden, ist u_0 die b.A. an f in $C([a, b])$. Im allgemeinen wird dies nicht der Fall sein und die Punkte bilden eine Referenz im Sinne des Satzes von de la Vallée-Poussin. Die Fehlerfunktion wird dann etwa wie in Abbildung V.2.1 skizziert aussehen. Wenn man die Punkte $t_{0,i}$ durch Punkte $t_{1,i}$ ersetzt, in denen die Fehlerfunktion ε_0 einen größeren Betrag und alternierendes Vorzeichen hat, erhält man eine Referenz mit größerer de la Vallée-Poussin Schranke. Für die neue Punktmenge bestimmt man wieder die beste diskrete Approximation u_1 und erhält so eine bessere Näherung an die b.A. an f in $C([a, b])$. Diese Idee wurde 1934 von Remez zu einem Iterationsverfahren ausgebaut.

ALGORITHMUS V.2.1 (Remez-Algorithmus). *Gesucht die b.A. in V an $f \in C([a, b])$.*

- (1) *Wähle $n + 1$ Punkte $a \leq t_{0,0} < \dots < t_{0,n} \leq b$ und bestimme zu diesen die beste diskrete Approximation $u_0 \in V$ an f und die zugehörige Abweichung η_0 . Setze $m = 1$.*
- (2) *Suche $n + 1$ Referenzpunkte $a \leq t_{m,0} < \dots < t_{m,n} \leq b$, so dass gilt*

$$\begin{aligned} |\varepsilon_{m-1}(t_{m,i})| &\geq \eta_{m-1} && \forall 0 \leq i \leq n, \\ \operatorname{sgn} \varepsilon_{m-1}(t_{m,i}) &= -\operatorname{sgn} \varepsilon_{m-1}(t_{m,i-1}) && \forall 1 \leq i \leq n, \\ |\varepsilon_{m-1}(t_{m,j})| &= \|\varepsilon_{m-1}\| && \text{für ein } j. \end{aligned}$$

- (3) *Bestimme zu der Punktmenge $t_{m,i}$ die beste diskrete Approximation $u_m \in V$ an f und die zugehörige Abweichung η_m . Falls $\|\varepsilon_m\| - |\eta_m|$ klein ist im Vergleich zu $\|\varepsilon_m\|$, beende das Verfahren. Andernfalls ersetze m durch $m + 1$ und gehe zu Schritt (2) zurück.*

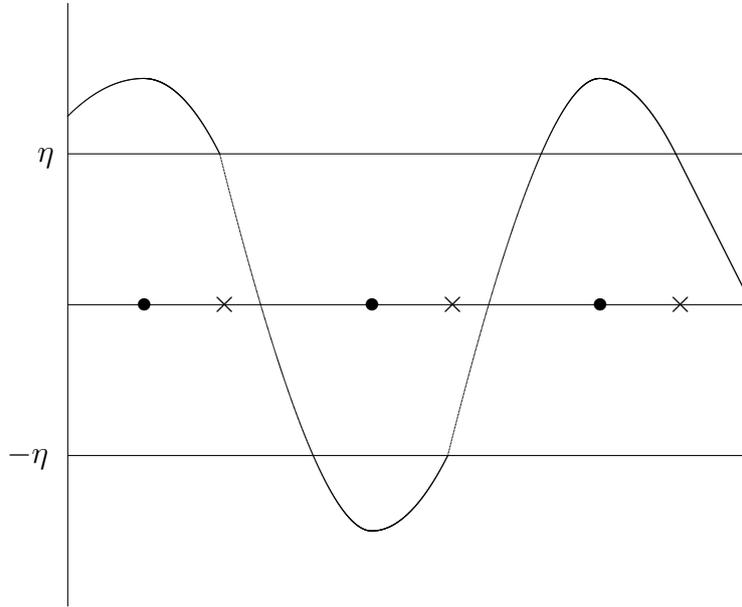


ABBILDUNG V.2.1. Fehlerfunktion ε_0 und Punkte $t_{0,i}$ (\times) und $t_{1,i}$ (\bullet)

Wesentlicher Punkt bei dem Remez-Algorithmus ist der Austausch der Referenz. Für die Konvergenz ist es erforderlich, einen Punkt τ aufzunehmen, an dem die Fehlernorm angenommen wird. Der Austausch weiterer Punkte ist zwar nicht erforderlich, verbessert aber i.a. die Konvergenz des Verfahrens.

Wir beschreiben nun, wie ein Punkt τ mit $|\varepsilon_{m-1}(\tau)| = \|\varepsilon_{m-1}\|$ in die Referenz aufgenommen wird. Der Einfachheit halber unterdrücken wir dabei den Index $m-1$. Setze $t_{-1} = a$ und $t_{n+1} = b$. Dann sind drei Fälle zu unterscheiden:

Normalfall: Es gibt einen Index $j \in \{0, \dots, n\}$ mit $\operatorname{sgn} \varepsilon(\tau) = \operatorname{sgn} \varepsilon(t_j)$ und $t_{j-1} \leq \tau \leq t_j$. Setze $t_{m,j} = \tau$ und $t_{m,i} = t_{m-1,i}$ für $i \neq j$.

Erster Ausnahmefall, neuer Referenzpunkt am linken Rand: Es ist $\tau < t_0$ und $\operatorname{sgn} \varepsilon(\tau) = -\operatorname{sgn} \varepsilon(t_0)$. Setze $t_{m,0} = \tau$ und $t_{m,i} = t_{m-1,i-1}$ für $1 \leq i \leq n$.

Zweiter Ausnahmefall, neuer Referenzpunkt am rechten Rand: Es ist $\tau > t_n$ und $\operatorname{sgn} \varepsilon(\tau) = -\operatorname{sgn} \varepsilon(t_n)$. Setze $t_{m,n} = \tau$ und $t_{m,i} = t_{m-1,i+1}$ für $0 \leq i \leq n-1$.

SATZ V.2.2. Sei V ein n -dimensionaler Haarscher Unterraum von $C([a, b])$. Dann konvergiert die durch den Remez-Algorithmus erzeugte Folge $(u_m)_{m \in \mathbb{N}}$ gegen die b.A. an f in V .

BEWEIS. 1. *Schritt:* Bezeichne mit

$$\ell_m(g) = \sum_{i=0}^n \beta_{m,i} g(t_{m,i})$$

das zu der Referenz $t_{m,0} < \dots < t_{m,n}$ gehörige Funktional aus Lemma V.1.2 (S. 88). Dann ist

$$\begin{aligned}\eta_m &= \ell_m(f) = \ell_m(f - u_{m-1}) \\ &= \sum_{i=0}^n \beta_{m,i} (f(t_{m,i}) - u_{m-1}(t_{m,i})).\end{aligned}$$

Sei j der Index, für den der Wert $\|\varepsilon_{m-1}\|$ angenommen wird. Für die anderen Referenzpunkte ist nach Konstruktion

$$|\varepsilon_{m-1}(t_{m,i})| \geq |\eta_{m-1}|.$$

Da die Vorzeichen der $\beta_{m,i}$ und der Fehlerfunktion alternieren, erhalten wir

$$\begin{aligned}|\eta_m| &= \sum_{i=0}^n |\beta_{m,i}| |\varepsilon_{m-1}(t_{m,i})| \\ &\geq \sum_{\substack{i=0 \\ i \neq j}}^n |\beta_{m,i}| |\eta_{m-1}| + |\beta_{m,j}| \|\varepsilon_{m-1}\| \\ &= \sum_{i=0}^n |\beta_{m,i}| |\eta_{m-1}| + |\beta_{m,j}| \{\|\varepsilon_{m-1}\| - |\eta_{m-1}|\} \\ &= |\eta_{m-1}| + |\beta_{m,j}| \{\|\varepsilon_{m-1}\| - |\eta_{m-1}|\}.\end{aligned}$$

2. *Schritt:* Wegen $|\eta_m| \geq |\eta_1| > 0$ gibt es gemäß Lemma V.1.5 (S. 91) ein $\delta > 0$ mit $t_{m,i} - t_{m,i-1} \geq \delta$ für alle m und i . Die Menge

$$S = \{(t_0, \dots, t_n) \in \mathbb{R}^{n+1} : a \leq t_0 < \dots < t_n \leq b, \\ t_i - t_{i-1} \geq \delta \text{ für } 1 \leq i \leq n\}$$

ist kompakt. Die Abbildung, die jedem Tupel (t_0, \dots, t_n) die Koeffizienten $(\beta_0, \dots, \beta_n)$ des Funktionals aus Lemma V.1.2 (S. 88) zuordnet, ist stetig und bildet S in $(\mathbb{R}^*)^{n+1}$ ab. Daher gibt es ein $\gamma > 0$ mit $|\beta_i| \geq \gamma$ für alle Koeffizienten β_i dieser Funktionale.

3. *Schritt:* Aus der Monotonie

$$|\eta_0| \leq |\eta_1| \leq \dots \leq |\eta_{m-1}| \leq |\eta_m| \leq \dots \leq \rho_V(f)$$

folgt

$$\lim_{m \rightarrow \infty} (|\eta_m| - |\eta_{m-1}|) = 0.$$

Zusammen mit den Abschätzungen der ersten beiden Schritte liefert dies

$$\begin{aligned}\lim_{m \rightarrow \infty} (\|\varepsilon_m\| - \rho_V(f)) &\leq \lim_{m \rightarrow \infty} (\|\varepsilon_m\| - |\eta_m|) \\ &\leq \lim_{m \rightarrow \infty} (|\eta_{m+1}| - |\eta_m|) \\ &= 0.\end{aligned}$$

Also ist $(u_m)_{m \in \mathbb{N}}$ eine Minimalfolge. Da die b.A. u^* gemäß Satz II.4.3 (S. 34) streng eindeutig ist, gibt es ein $c > 0$ mit

$$\|f - u\| \geq \|f - u^*\| + c \|u - u^*\|$$

für alle $u \in V$. Also gilt

$$\begin{aligned} \|u_m - u^*\| &\leq c^{-1} \{ \|f - u_m\| - \|f - u^*\| \} \\ &= c^{-1} \{ \|f - u_m\| - \rho_V(f) \} \\ &\xrightarrow{m \rightarrow \infty} 0. \end{aligned}$$

Dies beweist die Behauptung. \square

BEMERKUNG V.2.3. (1) Aus dem Beweis von Satz V.2.2 folgt, dass es beim Austausch der Referenz reicht, einen Punkt τ mit

$$|\varepsilon_m(\tau)| \geq \frac{1}{2} |\eta_{m-1}| + \frac{1}{2} \|\varepsilon_{m-1}\|$$

zu bestimmen. Denn in diesem Fall liefert der erste Schritt des Beweises die Abschätzung

$$|\eta_m| \geq |\eta_{m-1}| + \frac{1}{2} |\beta_{m,j}| \{ \|\varepsilon_{m-1}\| - |\eta_{m-1}| \}.$$

(2) Im Prinzip ist die erste Referenz $t_{0,i}$ beim Remez-Algorithmus beliebig. Bei der Polynomapproximation ist es jedoch günstig, hierfür die Extremalpunkte der Čebyšev-Polynome zu wählen.

V.3. Diskretisierungstheorie

Das Auffinden eines globalen Extremums einer stetigen Funktion bzw. einer Näherung gemäß Bemerkung V.2.3 (1) ist extrem aufwändig. Daher ersetzt man in der Praxis das Intervall $I = [a, b]$ durch eine diskrete Punktmenge Y , deren Kardinalität groß ist im Vergleich zur Dimension des approximierenden Raumes. Dann stellt sich die Frage, wie gut approximiert die beste diskrete Approximation auf Y die b.A. auf I . Zur Beantwortung dieser Frage muss man die Dichte von Y in I geeignet messen. O.E. ist dabei $I = [-1, 1]$. Dann bieten sich zwei Dichtemaße an

$$\begin{aligned} \delta(Y) &= \sup_{x \in I} \inf_{y \in Y} |x - y|, \\ \delta_T(Y) &= \sup_{x \in I} \inf_{y \in Y} |\arccos x - \arccos y|. \end{aligned}$$

Das zweite Maß entspricht der Dichte der Punktmenge, die man bei der Transformation $x = \cos \varphi$ in $[0, \pi]$ erhält.

SATZ V.3.1. Sei $Y \subset [-1, 1]$. Bezeichne mit $\|\cdot\|$ und $\|\cdot\|_Y$ die Maximumnormen auf I und Y . Dann gilt für jedes $p \in \mathbb{P}_n$ und jedes $n \in \mathbb{N}$

$$\begin{aligned} \|p\|_Y &\geq \{1 - n^2 \delta(Y)\} \|p\|, \\ \|p\|_Y &\geq \{1 - n \delta_T(Y)\} \|p\|. \end{aligned}$$

BEWEIS. Seien $x \in I$ und $y \in Y$ so, dass $|p(x)| = \|p\|$ und $|x - y| \leq \delta(Y)$ ist. Mit dem Mittelwertsatz und der Markovschen Ungleichung, Satz IV.4.5 (S. 70) folgt

$$\begin{aligned} \|p\|_Y &\geq |p(y)| \\ &= |p(x) + p'(\theta x + (1 - \theta)y)(x - y)| \quad \text{für ein } \theta \in [0, 1] \\ &\geq |p(x)| - |p'(\theta x + (1 - \theta)y)(x - y)| \\ &\geq \|p\| - \delta(Y) \|p'\| \\ &\geq \|p\| - \delta(Y)n^2 \|p\|. \end{aligned}$$

Dies beweist die erste Ungleichung.

Zum Nachweis der zweiten Ungleichung ersetze man p durch das trigonometrische Polynom $S(\varphi) = p(\cos \varphi)$ und wende die Bernsteinsche Ungleichung, Satz IV.4.4 (S. 69) an. \square

SATZ V.3.2. Seien $p \in \mathbb{P}_n$ und $q \in \mathbb{P}_n$ die b.A. an f auf dem Intervall $I = [-1, 1]$ bzw. auf der diskreten Punktmenge $Y \subset [-1, 1]$. Weiter sei $\delta(Y) < n^{-2}$ und $\delta_T(Y) < n^{-1}$. Dann ist

$$\begin{aligned} \|p - q\| &\leq c^{-1} \left\{ \omega(f; \delta(Y)) + 2n^2 \delta(Y) (1 - n^2 \delta(Y))^{-1} \|f\| \right\}, \\ \|p - q\| &\leq c^{-1} \left\{ \omega(f; \delta_T(Y)) + 2n \delta_T(Y) (1 - n \delta_T(Y))^{-1} \|f\| \right\}. \end{aligned}$$

Dabei ist c die Konstante aus der strengen Eindeutigkeit, Satz II.4.3 (S. 34).

BEWEIS. Seien $x \in I$ und $y \in Y$ so gewählt, dass $|f(x) - q(x)| = \|f - q\|$ und $|x - y| \leq \delta(Y)$ ist. Mit dem Mittelwertsatz folgt

$$\begin{aligned} \|f - q\| &= |f(x) - q(x)| \\ &\leq |f(x) - f(y)| + |f(y) - q(y)| + |q(y) - q(x)| \\ &\leq \omega(f; \delta(Y)) + \|f - q\|_Y + |x - y| \|q'\| \\ &\leq \omega(f; \delta(Y)) + \|f - q\|_Y + \delta(Y)n^2 \|q\|. \end{aligned}$$

Aus der Optimalität von q über Y ergibt sich

$$\|f - q\|_Y \leq \|f - p\|_Y \leq \|f - p\| = E_n(f)$$

und

$$\|q\|_Y \leq \|f\|_Y + \|f - q\|_Y \leq 2 \|f\|_Y \leq 2 \|f\|.$$

Aus Satz V.3.1 folgt

$$\|q\| \leq (1 - n^2 \delta(Y))^{-1} \|q\|_Y.$$

Insgesamt ergibt sich somit

$$\|f - q\| \leq \omega(f; \delta(Y)) + E_n(f) + 2\delta(Y)n^2 (1 - n^2 \delta(Y))^{-1} \|f\|.$$

Gemäß Satz II.4.3 (S. 34) ist andererseits

$$\|f - q\| \geq E_n(f) + c \|p - q\|.$$

Aus diesen beiden Abschätzungen folgt die erste Behauptung. Die zweite Behauptung wird analog bewiesen. \square

KAPITEL VI

Nichtlineare Approximation

VI.1. Einführung

Zur Motivation betrachten wir zwei Beispiele.

BEISPIEL VI.1.1. Betrachte die Funktion

$$\begin{aligned} f(x) &= \frac{1}{2} \sum_{k \in \mathbb{Z}} 2^{-|k|} e^{i(\operatorname{sgn} k) 2^{|k|} x} \\ &= \frac{1}{2} + \sum_{k \in \mathbb{N}^*} 2^{-k} \cos(2^k x). \end{aligned}$$

Die beiden Reihen werden durch die geometrische Reihe $\sum_{k \geq 0} 2^{-k}$ majorisiert und konvergieren daher gleichmäßig und absolut. Daher ist f wohl definiert und in $C_{2\pi}$. (Es ist sogar $f \in C_{2\pi}^\infty$!) Die b.A. an f in \mathbb{T}_n bezüglich der L^2 -Norm ist bekanntlich durch die n -te Partialsumme der Fourier-Reihe von f gegeben. Definieren wir daher für $n \in \mathbb{N}^*$ die Zahl k_n durch die Bedingung $2^{k_n} \leq n < 2^{k_n+1}$, so ist die b.A. an f in \mathbb{T}_n gegeben durch

$$p_n(x) = \frac{1}{2} + \sum_{1 \leq k \leq k_n} 2^{-k} \cos(2^k x),$$

und für den Fehler der b.A. gilt

$$E_n(f) = \sum_{k > k_n} 2^{-k} \leq 2^{-k_n} \sum_{\ell=0}^{\infty} 2^{-\ell} = 2^{-k_n} \approx \frac{1}{n}.$$

Wollen wir dagegen f durch eine trigonometrische Summe mit n Termen, deren Frequenzen aber nicht durch n beschränkt sind, approximieren, gibt es offensichtlich eine bessere Approximation nämlich

$$p_n^*(x) = \frac{1}{2} + \sum_{k=1}^n 2^{-k} \cos(2^k x).$$

Für den Fehler dieser Approximation gilt

$$\|f - p_n^*\|_{L^2([-\pi, \pi])} = \sum_{k=n+1}^{\infty} 2^{-k} = 2^{-n}.$$

Im ersten Fall fällt der Fehler linear, im zweiten Fall exponentiell.

Formal haben wir im ersten Fall die b.A. in dem linearen Unterraum \mathbb{T}_n gesucht. Die Theorie hierzu haben wir in den vorigen Kapiteln entwickelt. Im zweiten Fall suchen wir die b.A. in der Menge

$$\tilde{\mathbb{T}}_n = \left\{ \sum_{k \in \Lambda} \alpha_k e^{ikx} : \alpha_k \in \mathbb{C}, \Lambda \subset \mathbb{Z}, \#\Lambda \leq n \right\}.$$

Diese Menge ist offensichtlich *kein* linearer Unterraum von $C_{2\pi}$. Das Approximationsproblem ist also *nichtlinear*.

BEISPIEL VI.1.2. Wähle ein $\alpha \in (0, 1)$ und betrachte die Funktion $f(x) = x^\alpha$ auf $[0, 1]$. Wir wollen diese Funktion durch stückweise konstante Funktionen bezüglich der Maximumnorm best möglich approximieren. Dazu definieren wir für $n \in \mathbb{N}^*$ und $0 \leq k \leq n-1$

$$I_{k,n} = \begin{cases} \left[\frac{k}{n}, \frac{k+1}{n} \right) & \text{falls } k < n-1, \\ \left[\frac{k}{n}, \frac{k+1}{n} \right] & \text{falls } k = n-1 \end{cases}$$

und

$$\begin{aligned} S_n^0 &= \text{span} \{ \chi_{I_{k,n}} : 0 \leq k \leq n-1 \} \\ &= \left\{ u = \sum_{k=0}^{n-1} \alpha_k \chi_{I_{k,n}} : \alpha_k \in \mathbb{R} \right\}. \end{aligned}$$

Setze

$$E_n(f) = \inf_{u \in S_n^0} \|f - u\|.$$

Wir wollen $E_n(f)$ abschätzen. Dazu definieren wir

$$u^* = \sum_{k=0}^{n-1} f\left(\frac{2k+1}{2n}\right) \chi_{I_{k,n}} \in S_n^0.$$

Für $0 \leq k \leq n-1$ und $x \in I_{k,n}$ ist offensichtlich

$$|f(x) - u^*(x)| = \left| f(x) - f\left(\frac{2k+1}{2n}\right) \right| \leq \omega\left(f; \frac{1}{2n}\right).$$

Hieraus folgt

$$E_n(f) \leq \|f - u^*\| \leq \omega\left(f; \frac{1}{2n}\right) \approx n^{-\alpha}.$$

Wie wir später sehen werden, ist diese Abschätzung optimal. Schwächen wir die Bedingung an die approximierende Funktion dahingehend ab, dass sie auf n Intervallen stückweise konstant sein soll, ohne diese Intervalle vorzugeben, können wir die Approximationsgüte wesentlich verbessern. Um dies einzusehen, bezeichne mit

$$J_{k,n} = f^{-1}(I_{k,n})$$

die Urbilder der $I_{k,n}$. Wir unterteilen also die Ordinate statt der Abszisse. Es ist

$$J_{k,n} = \left[\left(\frac{k}{n} \right)^{\frac{1}{\alpha}}, \left(\frac{k+1}{n} \right)^{\frac{1}{\alpha}} \right)$$

mit der offensichtlichen Modifikation für $k = n - 1$. Setze

$$v_n^* = \sum_{k=0}^{n-1} \frac{2k+1}{2n} \chi_{J_{k,n}}.$$

Diese Funktion ist auf n Intervallen stückweise konstant. Für $0 \leq k \leq n - 1$ und $x \in J_{k,n}$ ist $f(x) \in I_{k,n}$ und daher

$$|f(x) - v^*(x)| = \left| \underbrace{f(x)}_{\in [\frac{k}{n}, \frac{k+1}{n}]} - \frac{2k+1}{2n} \right| \leq \frac{1}{2n}.$$

Daher ist

$$\|f - v^*\| \leq \frac{1}{2n} < n^{-\alpha} \approx E_n(f).$$

Im ersten Fall haben wir f in dem linearen Unterraum S_n^0 approximiert. Im zweiten Fall haben wir die b.A. in der Menge

$$\tilde{S}_n^0 = \left\{ \sum_{k=0}^{n-1} \alpha_k \chi_{I_k} : \alpha_k \in \mathbb{R}, I_k \subset [0, 1], \right. \\ \left. I_k \text{ paarweise disjunkt, } \bigcup I_k = [0, 1] \right\}$$

gesucht. Diese Menge ist offensichtlich *kein* linearer Unterraum von $L^\infty([0, 1])$. Das Approximationsproblem ist also *nichtlinear*.

Diese beiden Beispiele beschreiben die Prototypen der nichtlinearen Approximationsprobleme, die wir in diesem Kapitel betrachten werden. Um den funktionalanalytischen Aufwand zu minimieren, beschränken wir uns dabei weitgehend auf die Approximation in $C_{2\pi}$ bzw. $C([0, 1])$ bezüglich der Maximum- oder L^2 -Norm. Die Konzepte übertragen sich aber weitgehend auf mehrdimensionale Probleme und L^p -Normen. Insbesondere geben die Abschnitte über Wavelets nur einen ersten Einblick und Motivation der viel allgemeineren Theorie. Im Mittelpunkt steht stets das Problem, die asymptotische Konvergenzordnung der b.A. und die Menge der Funktionen, deren b.A. eine vorgegebene Konvergenzordnung hat, zu charakterisieren.

VI.2. Approximation im Hilbert-Raum

Sei X ein Hilbert-Raum mit Norm $\|\cdot\|$ und Skalarprodukt (\cdot, \cdot) . Wir nehmen an, dass X unendlich dimensional und separabel ist, d.h. dass

es eine Folge $(u_n)_{n \in \mathbb{N}} \subset X$ gibt mit

$$(VI.2.1) \quad \dim \operatorname{span} \{u_1, \dots, u_n\} = n,$$

$$(VI.2.2) \quad \bigcup_n \overline{\operatorname{span} \{u_1, \dots, u_n\}} = X.$$

Wegen des Schmidtschen Orthogonalisierungsverfahrens können wir o.E. annehmen, dass die u_n ein Orthonormalsystem bilden, d.h.

$$(VI.2.3) \quad (u_i, u_j) = \delta_{ij} \quad \forall i, j.$$

Setze

$$X_n = \operatorname{span} \{u_1, \dots, u_n\}.$$

Für $f \in X$ und $n \in \mathbb{N}^*$ sei

$$f_n = (f, u_n)$$

und

$$\rho_n(f) = \inf_{u \in X_n} \|f - u\|.$$

SATZ VI.2.1. *Für jedes $n \in \mathbb{N}^*$ ist X_n eine Čebyšev-Menge. Die b.A. an f in X ist gegeben durch*

$$P_n f = \sum_{k=1}^n f_k u_k.$$

Es ist

$$\rho_n(f) = \left\{ \sum_{k=n+1}^{\infty} |f_k|^2 \right\}^{\frac{1}{2}}.$$

BEWEIS. Gemäß Satz I.2.4 (S. 8) ist X_n als linearer Raum eine Existenzmenge. Gemäß Beispiel I.3.8 ist X strikt konvex und gemäß Satz I.3.5 (S. 12) X_n damit eine Eindeutigkeitsmenge. Sei

$$p_n = \sum_{k=1}^n \alpha_k u_k$$

die b.A. an f in X_n . Dann ist p_n das eindeutige Minimum der Funktion

$$X_n \ni u \mapsto J(u) = \|f - u\|^2 \in \mathbb{R}$$

und damit auch der einzige kritische Punkt von J . Wegen

$$(VI.2.4) \quad DJ(u)v = 2(f - u, v)$$

folgt

$$(f - p_n, u_k) = 0 \quad \forall 1 \leq k \leq n.$$

Zusammen mit der Orthogonalitätsrelation (VI.2.3) folgt hieraus die Darstellung von p_n .

Wegen der Dichterrelation (VI.2.2) ist

$$\lim_{n \rightarrow \infty} \|f - p_n\|^2 = 0.$$

Wegen der Orthogonalitätsbeziehungen (VI.2.4) und (VI.2.3) ist

$$\begin{aligned}\|f - p_n\|^2 &= (f - p_n, f - p_n) \\ &= (f - p_n, f + p_n) - 2 \underbrace{(f - p_n, p_n)}_{=0} \\ &= \|f\|^2 - \|p_n\|^2 \\ &= \|f\|^2 - \sum_{k=1}^n |f_k|^2.\end{aligned}$$

Also ist

$$\|f\|^2 = \sum_{k=1}^{\infty} |f_k|^2.$$

Hieraus folgt die Aussage über $\rho_n(f)$. \square

Gegeben sei eine Zahl $\alpha > 0$. Wie sehen die Elemente $f \in X$ aus, für die es ein $M \geq 0$ gibt mit $\rho_n(f) \leq Mn^{-\alpha}$ für alle $n \in \mathbb{N}^*$?

DEFINITION VI.2.2. Für $\alpha > 0$ ist der *Approximationsraum* $\mathcal{A}^\alpha((X_n))$ definiert durch

$$\mathcal{A}^\alpha((X_n)) = \{f \in X : \exists M \geq 0 \forall n \in \mathbb{N}^* \rho_n(f) \leq Mn^{-\alpha}\}.$$

Für $f \in \mathcal{A}^\alpha((X_n))$ ist $|f|_{\mathcal{A}^\alpha((X_n))}$ definiert als das Infimum über alle Konstanten M , für die $\rho_n(f) \leq Mn^{-\alpha}$ gilt.

SATZ VI.2.3. (1) $\mathcal{A}^\alpha((X_n))$ ist ein linearer Raum und $|\cdot|_{\mathcal{A}^\alpha((X_n))}$ ist eine Semi-Norm.

(2) Für $m \in \mathbb{N}^*$ sei

$$F_m = \left\{ \sum_{k=2^{m-1}+1}^{2^m} |f_k|^2 \right\}^{\frac{1}{2}}.$$

Dann ist $f \in \mathcal{A}^\alpha((X_n))$ genau dann, wenn es ein $M' \geq 0$ gibt mit $F_m \leq M'2^{-m\alpha}$ für alle $m \in \mathbb{N}^*$. Das Infimum über alle derartigen Konstanten M' ist eine zu $|\cdot|_{\mathcal{A}^\alpha((X_n))}$ äquivalente Semi-Norm.

BEWEIS. ad (1): Da X_n ein linearer Raum ist, ist $\rho_n(\lambda f) = |\lambda| \rho_n(f)$ für alle $\lambda \in \mathbb{R}$ und alle $f \in X$. Daher ist $|\cdot|_{\mathcal{A}^\alpha((X_n))}$ homogen und mit f auch λf in $\mathcal{A}^\alpha((X_n))$. Wegen

$$\begin{aligned}\rho_n(f + g) &\leq \|f + g - P_n f - P_n g\| \\ &\leq \|f - P_n f\| + \|g - P_n g\| \\ &= \rho_n(f) + \rho_n(g)\end{aligned}$$

ist mit f, g auch $f + g$ in $\mathcal{A}^\alpha((X_n))$ und $|\cdot|_{\mathcal{A}^\alpha((X_n))}$ erfüllt die Dreiecksungleichung. Wegen $\rho_n(f) = 0$ für $f \in X_n$ ist $|\cdot|_{\mathcal{A}^\alpha((X_n))}$ keine Norm.

ad (2): Wegen Satz VI.2.1 ist

$$\begin{aligned}
F_m^2 &= \sum_{k=2^{m-1}+1}^{2^m} |f_k|^2 \\
&= \sum_{k=2^{m-1}+1}^{\infty} |f_k|^2 - \sum_{k=2^m+1}^{\infty} |f_k|^2 \\
&= \rho_{2^{m-1}}(f)^2 - \rho_{2^m}(f)^2 \\
&\leq \rho_{2^{m-1}}(f)^2.
\end{aligned}$$

Also folgt aus $f \in \mathcal{A}^\alpha((X_n))$

$$F_m \leq M2^{-\alpha m + \alpha} = M'2^{-\alpha m}$$

mit $M' = M2^\alpha$. Dies ist die erste Implikation und die Abschätzung der Semi-Norm in eine Richtung.

Für die andere Implikation und die umgekehrte Abschätzung der Semi-Norm definieren wir für $n \in \mathbb{N}^*$ die Zahl $m \in \mathbb{N}^*$ durch $2^{m-1} \leq n < 2^m$. Wiederum mit Satz VI.2.1 folgt dann

$$\begin{aligned}
\rho_n(f)^2 &= \sum_{k=n+1}^{\infty} |f_k|^2 \\
&\leq \sum_{k=2^{m-1}+1}^{\infty} |f_k|^2 \\
&= \sum_{\ell=m}^{\infty} F_\ell^2 \\
&\leq \sum_{\ell=m}^{\infty} (M')^2 2^{-2\ell\alpha} \\
&= (M')^2 2^{-2m\alpha} \frac{1}{1 - 2^{-2\alpha}} \\
&\leq (M')^2 n^{-2\alpha} \frac{1}{1 - 2^{-2\alpha}}.
\end{aligned}$$

Dies beweist die andere Implikation mit $M = \frac{M'}{\sqrt{1-2^{-2\alpha}}}$. \square

Für viele Anwendungen ist folgende Variante von $\mathcal{A}^\alpha((X_n))$ günstiger.

DEFINITION VI.2.4. Für $\alpha > 0$ bezeichnet $\mathcal{A}_2^\alpha((X_n))$ den *Approximationsraum* aller $f \in X$, für die

$$|f|_{\mathcal{A}_2^\alpha((X_n))} = \left\{ \sum_{n=1}^{\infty} [n^\alpha \rho_n(f)]^2 \frac{1}{n} \right\}^{\frac{1}{2}}$$

endlich ist.

SATZ VI.2.5. (1) $\mathcal{A}_2^\alpha((X_n))$ ist ein linearer Raum; $|\cdot|_{\mathcal{A}_2^\alpha((X_n))}$ ist eine Semi-Norm.

(2) Es ist $f \in \mathcal{A}_2^\alpha((X_n))$ genau dann, wenn $\left\{ \sum_{k=0}^{\infty} 2^{2k\alpha} \rho_{2^k}(f)^2 \right\}^{\frac{1}{2}}$ endlich ist. In diesem Fall definiert dieser Ausdruck eine zu $|\cdot|_{\mathcal{A}_2^\alpha((X_n))}$ äquivalente Semi-Norm.

(3) Es ist $f \in \mathcal{A}_2^\alpha((X_n))$ genau dann, wenn $\left\{ \sum_{k=1}^{\infty} k^{2\alpha} |f_k|^2 \right\}^{\frac{1}{2}}$ endlich ist. In diesem Fall definiert dieser Ausdruck eine zu $|\cdot|_{\mathcal{A}_2^\alpha((X_n))}$ äquivalente Semi-Norm.

(4) Es ist $\mathcal{A}_2^\alpha((X_n)) \subsetneq \mathcal{A}^\alpha((X_n))$.

BEWEIS. *ad* (1): Folgt wie Satz VI.2.3 (1).

ad (2): Es ist

$$\sum_{n=1}^{\infty} [n^\alpha \rho_n(f)]^2 \frac{1}{n} = \sum_{k=0}^{\infty} \left\{ \sum_{n=2^k}^{2^{k+1}-1} n^{2\alpha-1} \rho_n(f)^2 \right\}.$$

Wegen der Schachtelung $X_{n-1} \subset X_n \subset X_{n+1}$ ist $\rho_{2^{k+1}}(f) \leq \rho_n(f) \leq \rho_{2^k}(f)$ für $2^k \leq n \leq 2^{k+1}$. Ebenso kann $n^{2\alpha-1}$ durch $2^{k(2\alpha-1)}$ und $2^{(k+1)(2\alpha-1)}$ eingeschachtelt werden, wobei die Fälle $\alpha \leq \frac{1}{2}$ und $\alpha > \frac{1}{2}$ unerschieden werden müssen. Da jeder $\{\cdot\}$ -Ausdruck in obiger Gleichheit 2^k Summanden hat, folgt, dass $\sum [n^\alpha \rho_n(f)]^2 \frac{1}{n}$ bis auf multiplikative Konstanten in beide Richtungen durch $\sum 2^{2k\alpha} \rho_{2^k}(f)^2$ abgeschätzt werden kann.

ad (3): Wegen Satz VI.2.1 ist

$$\begin{aligned} \sum_{k=0}^{\infty} 2^{2k\alpha} \rho_{2^k}(f)^2 &= \sum_{k=0}^{\infty} 2^{2k\alpha} \left\{ \sum_{m=2^{k+1}}^{\infty} |f_m|^2 \right\} \\ &= \sum_{m=2}^{\infty} |f_m|^2 \left\{ \sum_{k=0}^{[\ln_2 m]} 2^{2k\alpha} \right\}. \end{aligned}$$

Wegen

$$\sum_{k=0}^{[\ln_2 m]} 2^{2k\alpha} \approx m^{2\alpha}$$

folgt die Behauptung.

ad (4): Aus dem Beweis von Satz VI.2.3 (2) wissen wir, dass $F_m^2 \leq \rho_{2^{m-1}}(f)^2$ ist. Ist also $f \in \mathcal{A}_2^\alpha((X_n))$, so folgt

$$\sum_{m=1}^{\infty} 2^{2m\alpha} F_m^2 \leq \sum_{m=1}^{\infty} 2^{2m\alpha} \rho_{2^{m-1}}(f)^2 < \infty.$$

Also ist $(2^{2m\alpha} F_m^2)_{m \in \mathbb{N}^*}$ eine Nullfolge und damit beschränkt. Wegen Satz VI.2.3 ist daher $f \in \mathcal{A}^\alpha((X_n))$. Für den Nachweis von $\mathcal{A}_2^\alpha((X_n)) \neq \mathcal{A}^\alpha((X_n))$ betrachte ein $f \in X$ mit $\rho_n(f) = n^{-\alpha}$ für alle n . (Wegen Satz

VI.2.1 kann man ein solches f konstruieren.) Dann ist $f \in \mathcal{A}^\alpha((X_n))$ aber $\sum [n^\alpha \rho_n(f)]^2 \frac{1}{n} = \sum \frac{1}{n}$ unbeschränkt und damit $f \notin \mathcal{A}_2^\alpha((X_n))$. \square

BEISPIEL VI.2.6. Betrachte den Fall $X = L^2(-\pi, \pi)$, $u_k(x) = e^{ikx}$. Dann sind die f_k die Fourier-Koeffizienten von f . Daher stimmen wegen Satz VI.2.5 (4) die Räume $\mathcal{A}_2^\alpha((X_n))$ mit den klassischen Sobolev-Räumen $W^{\alpha,2}$ überein. Man beachte, dass diese äquivalent definiert sind über das Abklingverhalten der Fourier-Koeffizienten und über Glattheitseigenschaften der Funktionen. Daher kann man hoffen, auch in anderen Fällen die Räume $\mathcal{A}_2^\alpha((X_n))$ mit Räumen identifizieren zu können, die über die Glattheit von Funktionen definiert sind. Man beachte, dass man aus dieser Charakterisierung von $W^{\alpha,2}$ die Fehlerabschätzung

$$\left\{ \sum_{n=1}^{\infty} [n^\alpha E_n(f)]^2 \frac{1}{n} \right\}^{\frac{1}{2}} \leq c \|f\|_{W^{\alpha,2}}$$

erhält. Diese ist schärfer als die Abschätzung

$$E_n(f) \leq cn^{-\alpha} \|f\|_{W^{\alpha,2}},$$

die man üblicherweise in der Literatur z.B. über Finite Elemente findet.

Wir kommen nun zur nichtlinearen Approximation. Dazu definieren wir für $n \in \mathbb{N}^*$

$$\begin{aligned} \Sigma_n &= \left\{ u = \sum_{k \in \Lambda} \alpha_k u_k : \alpha_k \in \mathbb{R}, \Lambda \subset \mathbb{N}^*, \#\Lambda = n \right\} \\ &= \left\{ u = \sum_{i=1}^n \alpha_{k_i} u_{k_i} : \alpha_{k_i} \in \mathbb{R}, 1 \leq k_1 < \dots < k_n \right\} \end{aligned}$$

und

$$\sigma_n(f) = \inf_{u \in \Sigma_n} \|f - u\|.$$

Wie das Beispiel

$$u = \sum_{i=1}^n u_i, \quad v = \sum_{i=n+1}^{2n} u_i$$

zeigt, ist die Menge Σ_n weder linear noch konvex. (Sie ist allerdings sternförmig bezüglich des Nullpunktes.) Daher können wir die allgemeinen Sätze aus Kapitel I nicht anwenden.

SATZ VI.2.7. Σ_n ist eine Existenz- aber keine Eindeutigkeitsmenge. Für $f \in X$ bezeichne $(\gamma_n(f))_{n \in \mathbb{N}^*}$ die betragsmäßig absteigend umsortierte Folge $(f_n)_{n \in \mathbb{N}^*}$. Dann ist

$$\sigma_n(f) = \left\{ \sum_{k=n+1}^{\infty} |\gamma_k(f)|^2 \right\}^{\frac{1}{2}}.$$

BEWEIS. Offensichtlich ist

$$\sigma_n(f) = \inf_{\substack{V \subset X \\ \dim V = n}} \inf_{u \in V} \|f - u\|.$$

Sei $\Lambda \subset \mathbb{N}^*$ mit $\#\Lambda = n$ und $V = \text{span}\{u_k : k \in \Lambda\}$. Aus Satz VI.2.1 ergibt sich

$$\inf_{u \in V} \|f - u\| = \left\{ \sum_{k \in \mathbb{N}^* \setminus \Lambda} |f_k|^2 \right\}^{\frac{1}{2}}.$$

Daher ist

$$\sigma_n(f) = \inf_{\substack{\Lambda \subset \mathbb{N}^* \\ \#\Lambda = n}} \left\{ \sum_{k \in \mathbb{N}^* \setminus \Lambda} |f_k|^2 \right\}^{\frac{1}{2}}.$$

Wegen $|\gamma_1(f)| \geq |\gamma_2(f)| \geq \dots$ und $\{f_k : k \in \mathbb{N}^*\} = \{\gamma_k(f) : k \in \mathbb{N}^*\}$ folgt hieraus die behauptete Formel für $\sigma_n(f)$. Offensichtlich ist $\sum_{k=1}^n \gamma_k(f)u_k$ ein Minimalelement, so dass Σ_n eine Existenzmenge ist. Das Beispiel

$$f = \sum_{k=1}^{n+1} u_k + \sum_{k=n+2}^{\infty} 2^{-k} u_k$$

zeigt, dass Σ_n keine Eindeutigkeitsmenge sein kann. \square

Wegen Satz VI.2.7 kann man analog zum linearen Fall auch Approximationsräume für Σ_n definieren.

DEFINITION VI.2.8. Für $\alpha > 0$ sei

$$\mathcal{A}^\alpha((\Sigma_n)) = \{f \in X : \sigma_n(f) = O(n^{-\alpha})\}$$

und

$$\mathcal{A}_2^\alpha((\Sigma_n)) = \left\{ f \in X : \sum_{n=1}^{\infty} [n^\alpha \sigma_n(f)]^2 \frac{1}{n} < \infty \right\}.$$

Es gilt folgende zum linearen Fall analoge Charakterisierung.

SATZ VI.2.9. (1) Es ist $f \in \mathcal{A}^\alpha((\Sigma_n))$ genau dann, wenn $\gamma_n(f) = O(n^{-\alpha-\frac{1}{2}})$ ist.

(2) Es ist $f \in \mathcal{A}_2^\alpha((\Sigma_n))$ genau dann, wenn $\sum_{n=1}^{\infty} n^{2\alpha} \gamma_n(f)^2$ endlich ist.

BEWEIS. ad (1): Wegen Satz VI.2.7 und Satz VI.2.3 (2) mit $(\gamma_n(f))$ an Stelle von (f_n) gilt

$$f \in \mathcal{A}^\alpha((\Sigma_n)) \iff 2^{2m\alpha} \sum_{k=2^{m-1}+1}^{2^m} |\gamma_k(f)|^2 = O(1).$$

Da die Folge $(\gamma_n(f))$ monoton fallend ist, ist

$$\begin{aligned} 2^{2m\alpha+m-1} |\gamma_{2^m}(f)|^2 &\leq 2^{2m\alpha} \sum_{k=2^{m-1}+1}^{2^m} |\gamma_{2^m}(f)|^2 \\ &\leq 2^{2m\alpha} \sum_{k=2^{m-1}+1}^{2^m} |\gamma_k(f)|^2 \\ &\leq 2^{2m\alpha} \sum_{k=2^{m-1}+1}^{2^m} |\gamma_{2^{m-1}}(f)|^2 \\ &\leq 2^{2m\alpha+m-1} |\gamma_{2^{m-1}}(f)|^2. \end{aligned}$$

Also ist

$$f \in \mathcal{A}^\alpha((\Sigma_n)) \iff 2^{m(\alpha+\frac{1}{2})} |\gamma_{2^m}(f)| = O(1).$$

Für $n \in \mathbb{N}^*$ sei $m \in \mathbb{N}$ so, dass $2^m \leq n < 2^{m+1}$ ist. Dann ist wegen der Monotonie von $(\gamma_n(f))$

$$\begin{aligned} n^{\alpha+\frac{1}{2}} |\gamma_n(f)| &\leq n^{\alpha+\frac{1}{2}} |\gamma_{2^m}(f)| \\ &= (n2^{-m})^{\alpha+\frac{1}{2}} \left| 2^{m(\alpha+\frac{1}{2})} \gamma_{2^m}(f) \right| \\ &\leq 2^{\alpha+\frac{1}{2}} \left| 2^{m(\alpha+\frac{1}{2})} \gamma_{2^m}(f) \right| \end{aligned}$$

und

$$\begin{aligned} n^{\alpha+\frac{1}{2}} |\gamma_n(f)| &\geq n^{\alpha+\frac{1}{2}} |\gamma_{2^{m+1}}(f)| \\ &= (n2^{-m-1})^{\alpha+\frac{1}{2}} \left| 2^{(m+1)(\alpha+\frac{1}{2})} \gamma_{2^{m+1}}(f) \right| \\ &\geq 2^{-\alpha-\frac{1}{2}} \left| 2^{(m+1)(\alpha+\frac{1}{2})} \gamma_{2^{m+1}}(f) \right|. \end{aligned}$$

Hieraus folgt die Behauptung.

ad (2): Folgt aus Satz VI.2.7 und Satz VI.2.5 (3) mit $(\gamma_n(f))$ an Stelle von (f_n) . \square

BEMERKUNG VI.2.10. Wie in Satz VI.2.5 (4) folgt aus Satz VI.2.9 $\mathcal{A}_2^\alpha((\Sigma_n)) \subsetneq \mathcal{A}^\alpha((\Sigma_n))$.

VI.3. Lineare Approximation durch stückweise konstante Funktionen

Sei $\tau = \{0 = t_0 < t_1 < \dots < t_N = 1\}$ eine geordnete Punktmenge in $[0, 1]$. Die Zahl N ist dabei beliebig. τ induziert eine Zerlegung von $[0, 1]$ in N disjunkte Intervalle I_1, \dots, I_N mit

$$I_k = \begin{cases} [t_{k-1}, t_k) & \text{für } 1 \leq k \leq N-1, \\ [t_{k-1}, t_k] & \text{für } k = N. \end{cases}$$

Wir nennen τ daher auch Synonym eine *Zerlegung* von $[0, 1]$ und schreiben $I \in \tau$, wenn wir ein zu dieser Zerlegung gehörendes Intervall kennzeichnen wollen. Die Menge aller Zerlegungen τ bezeichnen wir mit \mathcal{T} . Für festes $\tau \in \mathcal{T}$ ist

$$\delta_\tau = \max_{1 \leq k \leq N} |t_k - t_{k-1}| = \max_{I \in \tau} |I|$$

die zugehörige *Gitterweite* oder *Feinheit* von τ . Für $\tau \in \mathcal{T}$ ist

$$S^0(\tau) = \text{span} \{\chi_I : I \in \tau\} = \text{span} \{\chi_{I_k} : 1 \leq k \leq N\}$$

der zugehörige Raum der stückweise konstanten Funktionen.

Diese Notation orientiert sich an der Theorie der Finiten Elemente. In der Approximationstheorie findet man auch häufig die Bezeichnung $S^1(\tau)$.

Da $S^0(\tau)$ für jedes $p \in [1, \infty]$ ein linearer Unterraum von $L^p([0, 1])$ ist, ist $S^0(\tau)$ eine Existenzmenge. Daher ist die Definition

$$\rho_\tau(f)_p = \inf_{u \in S^0(\tau)} \|f - u\|_p$$

sinnvoll. Für $p = \infty$ betrachten wir dabei nur stetige Funktionen, so dass $\|\cdot\|_\infty$ die klassische Maximumnorm ist. Wenn keine Verwechslung möglich ist, lassen wir häufig den Index ∞ weg.

SATZ VI.3.1. *Für $0 < \alpha \leq 1$ sind folgende Aussagen äquivalent:*

- (1) $f \in \text{Lip } \alpha$.
- (2) $\rho_\tau(f)_\infty = O(\delta_\tau^\alpha)$ für alle $\tau \in \mathcal{T}$.

BEWEIS. „(1) \implies (2)“: Sei $f \in \text{Lip } \alpha$ und $\tau \in \mathcal{T}$. Für $I \in \tau$ sei η_I der Mittelpunkt von I . Setze

$$u = \sum_{I \in \tau} f(\eta_I) \chi_I \in S^0(\tau).$$

Für jedes $x \in [0, 1]$ gibt es ein $I_x \in \tau$ mit $x \in I_x$. Dann ist

$$|x - \eta_{I_x}| \leq \frac{1}{2} \delta_\tau$$

und wegen $f \in \text{Lip } \alpha$

$$\begin{aligned} |f(x) - u(x)| &= |f(x) - f(\eta_{I_x})| \leq M |x - \eta_{I_x}|^\alpha \\ &\leq 2^{-\alpha} M \delta_\tau^\alpha. \end{aligned}$$

Hieraus folgt

$$\rho_\tau(f)_\infty \leq \|f - u\|_\infty \leq 2^{-\alpha} M \delta_\tau^\alpha.$$

„(2) \implies (1)“: Seien $x, y \in [0, 1]$ mit $x \neq y$ beliebig und $\varepsilon > 0$. Dann gibt es ein $\tau \in \mathcal{T}$ mit $\delta_\tau \leq |x - y| + \varepsilon$, so dass x und y in dem gleichen

Intervall I^* von τ liegen. Sei $u \in S^0(\tau)$ eine b.A. an f . Da u auf I^* konstant ist, folgt

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - u(x)| + \underbrace{|u(x) - u(y)|}_{=0} + |u(y) - f(y)| \\ &\leq 2 \|f - u\|_\infty \\ &\leq 2M\delta_\tau^\alpha \\ &= 2M [|x - y| + \varepsilon]^\alpha. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt $f \in \text{Lip } \alpha$. \square

Satz VI.3.1 ist für die Praxis nicht befriedigend. Dort betrachtet man in der Regel Folgen von Unterteilungen. Als wichtigstes Beispiel betrachten wir äquidistante Unterteilungen und definieren für $n \in \mathbb{N}^*$

$$\Delta_n = \left\{ \frac{k}{n} : 0 \leq k \leq n \right\}$$

und setzen

$$\rho_n(f) = \rho_{\Delta_n}(f)_\infty.$$

SATZ VI.3.2. Sei $0 < \alpha \leq 1$. Dann ist $f \in \text{Lip } \alpha$ genau dann, wenn $\rho_n(f) = O(n^{-\alpha})$ ist.

BEWEIS. „ \implies “: Folgt wegen $\delta_{\Delta_n} = \frac{1}{n}$ aus Satz VI.3.1.

„ \impliedby “: Diese Richtung ist komplizierter. Ihr Beweis nutzt aus, dass jedes $x \in [0, 1)$ in das Innere von hinreichend vielen Intervallen aus den Δ_n fällt. Genauer gilt:

Lemma: Sei $n \geq 64$ und $x \in [0, 1 - \frac{1}{16n}]$. Dann gibt es ein m mit $n \leq m \leq 2n$ und ein j mit $0 \leq j \leq m - 1$ und $[x, x + \frac{1}{16n}] \subset [\frac{j}{m}, \frac{j+1}{m})$ (siehe [5, Lemmas XII.2.2, XII.2.3, S. 356–357]).

Sei nun $n \geq 64$, $0 < h \leq \frac{1}{16n}$, $x \in [0, 1 - h)$ und m, j wie im Lemma. Sei $u_m \in S^0(\Delta_m)$ eine b.A. an f . Da u_m auf $[\frac{j}{m}, \frac{j+1}{m})$ konstant ist, folgt

$$\begin{aligned} |f(x+h) - f(x)| &= \left| f(x+h) - \underbrace{u_m(x+h) + u_m(x)}_{=0} - f(x) \right| \\ &\leq 2 \|f - u_m\|_\infty \\ &\leq 2\rho_m(f) \\ &\leq 2 \max_{n \leq k \leq 2n} \rho_k(f) \\ &\leq Cn^{-\alpha}. \end{aligned}$$

Da die rechte Seite dieser Abschätzung nicht von x und h abhängt, folgt $\omega(f; h) = O(h^\alpha)$ und damit $f \in \text{Lip } \alpha$. \square

BEMERKUNG VI.3.3. (1) Mit den gleichen Techniken kann man die Sättigungseigenschaft „ $\rho_n(f) = o(n^{-1}) \iff f$ konstant“ beweisen.

(2) Für die Richtung „ \impliedby “ in Satz VI.3.2 ist es wesentlich, dass die

Intervalle in $\Delta_n, \Delta_{n+1}, \dots, \Delta_{2n}$ hinreichend überlappen. Betrachte z.B. die dyadischen Unterteilungen Δ_{2^k} . Jedes $I \in \Delta_{2^\ell}$ mit $\ell > k$ wird durch mehrfaches Halbieren eines $J \in \Delta_{2^k}$ gewonnen. Daher ist $S^0(\Delta_{2^k}) \subset S^0(\Delta_{2^\ell})$ für jedes $\ell \geq k$ und $\rho_{2^\ell}(f) = 0$ für jedes $f \in S^0(\Delta_{2^k})$ und jedes $\ell \geq k$. Aber es ist $f \notin C([0, 1])$ und natürlich erst recht nicht $f \in \text{Lip } \alpha$ für irgendein $\alpha > 0$.

Wir wollen die Jackson-Teile der Sätze VI.3.1 und VI.3.2 nun auf den L^p -Fall übertragen. Dazu müssen wir zunächst den Begriff des Stetigkeitsmoduls und der Lipschitz-Räume übertragen.

Im Folgenden ist stets $1 \leq p < \infty$. Funktionen auf $[0, 1]$ stellen wir uns durch Null auf ganz \mathbb{R} fortgesetzt vor.

DEFINITION VI.3.4. (1) Für $h \in \mathbb{R}$ ist der *Translationsoperator* T_h definiert durch

$$\begin{aligned} T_h : L^p([0, 1]) &\rightarrow L^p([0, 1]) \\ T_h f(x) &= f(x + h). \end{aligned}$$

(2) Der *Glattheitsmodul* oder *Stetigkeitsmodul* $\omega_p(f; h)$ ist definiert durch

$$\omega_p(f; h) = \sup_{0 \leq t \leq h} \|T_t f - f\|_p.$$

(3) Für $\alpha > 0$ ist

$$\text{Lip}_p \alpha = \{f \in L^p([0, 1]) : \omega_p(f; h) = O(h^\alpha)\}.$$

BEMERKUNG VI.3.5. $\omega_p(f; \cdot)$ ist monoton wachsend.

DEFINITION VI.3.6. Für gegebenes $\tau \in \mathcal{T}$ definiere den *Quasi-Interpolationsoperator* $I_\tau : L^1([0, 1]) \rightarrow S^0(\tau)$ durch

$$I_\tau f = \sum_{I \in \tau} \left\{ \frac{1}{|I|} \int_I f \right\} \chi_I,$$

wobei $|I|$ die Länge von I ist.

SATZ VI.3.7. Für alle $1 \leq p < \infty$, alle $\tau \in \mathcal{T}$, alle $f \in L^p([0, 1])$ und alle $g \in L^p([0, 1])$ mit $g' \in L^p([0, 1])$ gilt:

- (1) $\|I_\tau f\|_p \leq \|f\|_p,$
- (2) $\|g - I_\tau g\|_p \leq \delta_\tau \|g'\|_p.$

BEWEIS. ad (1): Definiere den dualen Lebesgue-Exponenten $p' \in (1, \infty]$ zu p durch $\frac{1}{p} + \frac{1}{p'} = 1$. Dann ist insbesondere $\frac{p}{p'} = p - 1$. Damit folgt aus der Hölderschen Ungleichung

$$\frac{1}{|I|} \left| \int_I f \right| \leq \frac{1}{|I|} \int_I |f| \leq \frac{1}{|I|} |I|^{\frac{1}{p'}} \left\{ \int_I |f|^p \right\}^{\frac{1}{p}}.$$

Also ist

$$\left| \frac{1}{|I|} \int_I f \right|^p \leq \frac{|I|^{p-1}}{|I|^p} \int_I |f|^p$$

und

$$\int_I \left| \frac{1}{|I|} \int_I f \right|^p \leq \frac{|I|^p}{|I|^p} \int_I |f|^p = \int_I |f|^p.$$

Wegen

$$\|I_\tau f\|_p = \left\{ \sum_{I \in \tau} \int_I \left| \frac{1}{|I|} \int_I f \right|^p \right\}^{\frac{1}{p}}$$

folgt hieraus die Behauptung.

ad (2): Sei $g \in L^p$ differenzierbar mit $g' \in L^p$ und $I \in \tau$. Für $t \in I$ folgt

$$\begin{aligned} \left| g(t) - \frac{1}{|I|} \int_I g \right| &= \frac{1}{|I|} \left| \int_I (g(t) - g(s)) ds \right| \\ &= \frac{1}{|I|} \left| \int_I \left\{ \int_I \chi_{[\min\{t,s\}, \max\{t,s\}]}(u) g'(u) du \right\} ds \right| \\ &\leq \frac{1}{|I|} \int_I \left\{ \int_I |g'(u)| du \right\} ds \\ &= \int_I |g'(u)| du. \end{aligned}$$

Für $p = 1$ folgt hieraus

$$\begin{aligned} \int_I |g(t) - I_\tau g(t)| dt &= \int_I \left| g(t) - \frac{1}{|I|} \int_I g \right| dt \leq \int_I \left\{ \int_I |g'(u)| du \right\} dt \\ &\leq |I| \int_I |g'(u)| du \\ &\leq \delta_\tau \int_I |g'(u)| du \end{aligned}$$

und daher

$$\|g - I_\tau g\|_1 \leq \delta_\tau \|g'\|_1.$$

Für $p > 1$ dagegen folgt mit der Hölderschen Ungleichung

$$\begin{aligned} \int_I |g(t) - I_\tau g(t)|^p dt &= \int_I \left| g(t) - \frac{1}{|I|} \int_I g \right|^p dt \leq \int_I \left\{ \int_I |g'(u)| du \right\}^p dt \\ &= |I| \left\{ \int_I |g'(u)| du \right\}^p \\ &\leq |I| |I|^{\frac{p}{p'}} \int_I |g'(u)|^p du \\ &= |I|^p \int_I |g'(u)|^p du \\ &\leq \delta_\tau^p \int_I |g'(u)|^p du \end{aligned}$$

und daher

$$\|g - I_\tau g\|_p^p \leq \delta_\tau^p \|g'\|_p^p.$$

Hieraus folgt die Behauptung für $p > 1$. □

SATZ VI.3.8. *Zu jedem $f \in L^p([0, 1])$ und jedem $\delta > 0$ gibt es ein $g = g(f, \delta) \in L^p([0, 1])$, das differenzierbar ist mit $g' \in L^p([0, 1])$ und das die Abschätzung*

$$\|f - g\|_p + \delta \|g'\|_p \leq 2\omega_p(f; \delta)$$

erfüllt.

BEWEIS. Wir definieren die Funktion g durch

$$(VI.3.1) \quad g = f + \int_0^1 (T_{\delta u} - I)f du$$

d.h.

$$(VI.3.2) \quad \begin{aligned} g(x) &= f(x) + \int_0^1 [f(x + \delta u) - f(x)] du \\ &= \int_0^1 f(x + \delta u) du. \end{aligned}$$

Aus (VI.3.1) folgt

$$\begin{aligned} \|f - g\|_p &= \left\| \int_0^1 (T_{\delta u} - I)f du \right\|_p \\ &\leq \int_0^1 \|(T_{\delta u} - I)f\|_p du \\ &\leq \int_0^1 \omega_p(f; \delta u) du \\ &\leq \int_0^1 \omega_p(f; \delta) du \\ &= \omega_p(f; \delta). \end{aligned}$$

Aus (VI.3.2) folgt mit der Transformation $v = x + \delta u$ die Darstellung

$$g(x) = \int_x^{x+\delta} \frac{1}{\delta} f(v) dv.$$

Also ist g differenzierbar mit

$$g'(x) = \frac{1}{\delta} [f(x + \delta) - f(x)]$$

bzw.

$$g' = \frac{1}{\delta} (T_\delta - I)f.$$

Also ist $g' \in L^p([0, 1])$ und

$$\|g'\|_p = \frac{1}{\delta} \|(T_\delta - I)f\|_p \leq \frac{1}{\delta} \omega_p(f; \delta).$$

Hieraus folgt die Behauptung. □

SATZ VI.3.9. Für jedes $f \in L^p([0, 1])$, jedes $p \in [1, \infty)$ und jedes $\tau \in \mathcal{T}$ ist $\rho_\tau(f)_p \leq 4\omega_p(f; \delta_\tau)$. Insbesondere ist $\rho_\tau(f) = O(\delta_\tau^\alpha)$ für $f \in \text{Lip}_p \alpha$ mit $0 < \alpha \leq 1$.

BEWEIS. Offensichtlich müssen wir nur die erste Abschätzung beweisen. Sei hierzu zu f und δ_τ die Funktion g wie in Satz VI.3.8. Dann folgt

$$\begin{aligned} \rho_\tau(f)_p &\leq \|f - I_\tau f\|_p \\ &\leq \|f - g\|_p + \underbrace{\|g - I_\tau g\|_p}_{\leq \delta_\tau \|g'\|_p} + \underbrace{\|I_\tau f - I_\tau g\|_p}_{\leq \|f - g\|_p} \\ &\leq 2\|f - g\|_p + \delta_\tau \|g'\|_p \\ &\leq 2\left\{\|f - g\|_p + \delta_\tau \|g'\|_p\right\} \\ &\leq 4\omega_p(f; \delta_\tau). \quad \square \end{aligned}$$

BEMERKUNG VI.3.10. Mit technischem Mehraufwand kann man auch die Umkehrung von Satz VI.3.9 und das entsprechende Resultat für die Unterteilungen Δ_n beweisen. Ebenso kann man stückweise Polynome höheren Grades betrachten. In diesem Fall muss man bei der Definition des Stetigkeitsmoduls ω_p den Ausdruck $T_h - I = \sum_{k=0}^1 (-1)^k \binom{1}{k} T_h^k$ durch $\sum_{k=0}^r (-1)^k \binom{r}{k} T_h^k$ ersetzen, wobei $r - 1$ der Polynomgrad ist.

VI.4. Nichtlineare Approximation durch stückweise konstante Funktionen

Sei $\tau \in \mathcal{T}$ eine Zerlegung von $[0, 1]$. Wir sagen $\#\tau = n$, wenn τ aus n Intervallen besteht. Die Punktmenge τ besteht dann aus $n + 1$ Punkten. Setze

$$\Sigma_n = \bigcup_{\substack{\tau \in \mathcal{T} \\ \#\tau = n}} S^0(\tau).$$

Die Approximation in Σ_n ist das nichtlineare Analogon zum vorigen Abschnitt. Der Einfachheit halber beschränken wir uns im Folgenden auf die Approximation stetiger Funktionen in der Supremumsnorm $\|\cdot\|_\infty$.

Die Menge Σ_n ist weder konvex noch linear. Daher können wir die Existenzaussagen aus § 1.2 nicht heranziehen.

SATZ VI.4.1. Σ_n ist eine Existenzmenge.

BEWEIS. Sei $\tau = \{0 = t_0 < t_1 < \dots < t_n = 1\}$ eine beliebige Zerlegung von $[0, 1]$. Dann ist

$$\rho_\tau(f) = \inf_{u \in S^0(\tau)} \|f - u\| = \max_{1 \leq k \leq n} \min_{c \in \mathbb{R}} \max_{t_{k-1} \leq x \leq t_k} |f(x) - c|.$$

Offensichtlich nimmt die Funktion

$$c \mapsto \max_{t_{k-1} \leq x \leq t_k} |f(x) - c|$$

ihr Minimum genau an der Stelle

$$m_k(f) = \frac{1}{2} \max_{t_{k-1} \leq x \leq t_k} f(x) + \frac{1}{2} \min_{t_{k-1} \leq x \leq t_k} f(x)$$

an. Daher ist

$$u_\tau = \sum_{k=1}^n m_k(f) \chi_{[t_{k-1}, t_k)}$$

die eindeutige b.A. an f in $S^0(\tau)$.

Seien nun $(u_k)_{k \in \mathbb{N}} \subset \Sigma_n$ eine Minimalfolge und $\tau_k = \{0 = t_{0,k} < t_{1,k} < \dots < t_{n,k} = 1\}$ die zugehörigen Zerlegungen. Da die Folgen $(t_{i,k})_{k \in \mathbb{N}}$, $0 \leq i \leq n$, beschränkt sind, können wir nach Übergang zu einer Teilfolge annehmen, dass sie konvergieren, d.h. $t_{i,k} \rightarrow t_i^*$ für $0 \leq i \leq n$. Sei τ^* die zugehörige Zerlegung. Wegen der expliziten Darstellung der b.A. $u_\tau \in S^0(\tau)$ an f in $S^0(\tau)$ und der Stetigkeit von f konvergieren die u_{τ_k} gleichmäßig gegen ein $u^* \in S^0(\tau^*)$. Da $(u_{\tau_k})_{k \in \mathbb{N}}$ auch eine Minimalfolge ist, ist u^* eine b.A. an f in Σ_n . \square

Wegen Satz VI.4.1 ist

$$\sigma_n(f) = \inf_{u \in \Sigma_n} \|f - u\|$$

wohl definiert. Wir wollen nun die Klasse aller f mit $\sigma_n(f) = O(n^{-1})$ charakterisieren. Dazu benötigen wir den Begriff der totalen Variation.

DEFINITION VI.4.2. Sei $[a, b]$ ein Intervall und $f : [a, b] \rightarrow \mathbb{R}$ eine Funktion. Dann heißt

$$\text{Var}_{[a,b]}(f) = \sup_{n \in \mathbb{N}^*} \sup_{a \leq x_0 < \dots < x_n \leq b} \sum_{i=1}^n |f(x_i) - f(x_{i-1})|$$

die *totale Variation* von f . Die Funktion f heißt von *beschränkter Variation*, wenn $\text{Var}_{[a,b]}(f)$ endlich ist. Die Menge aller Funktionen mit beschränkter Variation wird mit $\text{BV}(a, b)$ bezeichnet.

BEMERKUNG VI.4.3. (1) Es ist $\text{BV}(0, 1) \not\subset C([0, 1])$ und $C([0, 1]) \not\subset \text{BV}(0, 1)$.

(2) Ist $f : [0, 1] \rightarrow \mathbb{R}$ monoton, so ist $f \in \text{BV}(0, 1)$.

(3) Ist $f : [0, 1] \rightarrow \mathbb{R}$ differenzierbar und f' integrierbar, so ist $f \in \text{BV}(0, 1)$ und

$$\text{Var}_{[0,1]}(f) \leq \int_0^1 |f'(x)| dx.$$

(4) Ist $f \in \text{BV}(0, 1)$, so ist die Funktion $x \mapsto \text{Var}_{[0,x]}(f)$ monoton wachsend.

(5) Es gilt

$$\text{Var}_{[a,b]}(f) + \text{Var}_{[b,c]}(f) \leq \text{Var}_{[a,c]}(f),$$

wenn jeder dieser Ausdrücke endlich ist.

BEWEIS. *ad (1)*: Die Funktion

$$f(x) = \sum_{k=0}^{\infty} 2^{-k} \chi_{(2^{-k-1}, 2^{-k}]}$$

ist nicht stetig, aber gemäß Teil (2) in $BV(0, 1)$. Die Funktion

$$f(x) = \begin{cases} x \sin\left(\frac{\pi}{2x}\right) & \text{für } 0 < x \leq 1, \\ 0 & \text{für } x = 0 \end{cases}$$

ist stetig. Für $n \in \mathbb{N}^*$ ist aber

$$\begin{aligned} \text{Var}_{[0,1]}(f) &\geq \sum_{i=1}^n \left| f\left(\frac{1}{2i+1}\right) - f\left(\frac{1}{2i-1}\right) \right| \\ &= \sum_{i=1}^n \left| \frac{1}{2i+1} \underbrace{\sin\left(\frac{2i+1}{2}\pi\right)}_{=(-1)^i} - \frac{1}{2i-1} \underbrace{\sin\left(\frac{2i-1}{2}\pi\right)}_{=(-1)^{i-1}} \right| \\ &= \sum_{i=1}^n \left\{ \frac{1}{2i+1} + \frac{1}{2i-1} \right\} \\ &\geq \sum_{i=1}^n \left\{ \frac{1}{2i+1} + \frac{1}{2i} \right\} \\ &= \sum_{j=2}^{2n+1} \frac{1}{j} \\ &\xrightarrow{n \rightarrow \infty} \infty. \end{aligned}$$

Also ist $f \notin BV(0, 1)$.

ad (2): O.E. ist f monoton wachsend. Für $0 \leq x_0 \leq \dots \leq x_n \leq 1$ folgt dann

$$\begin{aligned} \sum_{i=1}^n |f(x_i) - f(x_{i-1})| &= \sum_{i=1}^n (f(x_i) - f(x_{i-1})) = f(x_n) - f(x_0) \\ &\leq f(1) - f(0). \end{aligned}$$

ad (3): Für $0 \leq x_0 \leq \dots \leq x_n \leq 1$ folgt

$$\begin{aligned} \sum_{i=1}^n |f(x_i) - f(x_{i-1})| &= \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f'(t) dt \right| \leq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} |f'(t)| dt \\ &\leq \int_0^1 |f'(t)| dt. \end{aligned}$$

ad (4): Ist offensichtlich.

ad (5): Sei $\varepsilon > 0$. Dann gibt es Zerlegungen $a \leq x_0 \leq \dots \leq x_n \leq b$ und $b \leq y_0 \leq \dots \leq y_m \leq c$ mit

$$\begin{aligned} \sum_{i=1}^n |f(x_i) - f(x_{i-1})| &\geq \text{Var}_{[a,b]}(f) - \varepsilon, \\ \sum_{j=1}^m |f(y_j) - f(y_{j-1})| &\geq \text{Var}_{[b,c]}(f) - \varepsilon. \end{aligned}$$

O.E. ist $x_n = y_0$. Sonst nehme man den Punkt b zu den Zerlegungen hinzu. Dann folgt

$$\begin{aligned} \text{Var}_{[a,c]}(f) &\geq \sum_{i=1}^n |f(x_i) - f(x_{i-1})| + \sum_{j=1}^m |f(y_j) - f(y_{j-1})| \\ &\geq \text{Var}_{[a,b]}(f) + \text{Var}_{[b,c]}(f) - 2\varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig war, folgt hieraus die Behauptung. \square

SATZ VI.4.4. Für $f \in C([0, 1])$ ist $\sigma_n(f) = O(n^{-1})$ genau dann, wenn $f \in \text{BV}(0, 1)$ ist.

BEWEIS. „ \Leftarrow “: Sei $M = \text{Var}_{[0,1]}(f)$. Definiere die Funktion $\varphi : [0, 1] \rightarrow \mathbb{R}$ durch $\varphi(x) = \text{Var}_{[0,x]}(f)$ und setze $t_i = \varphi^{-1}(\frac{i}{n}M)$ für $0 \leq i \leq n$. Da gemäß Bemerkung VI.4.3 (4) φ monoton wachsend ist, gilt $0 = t_0 \leq \dots \leq t_n = 1$. Wegen Bemerkung VI.4.3 (5) ist weiter für jedes $1 \leq i \leq n$

$$\begin{aligned} \text{Var}_{[t_{i-1}, t_i]}(f) &= \text{Var}_{[t_{i-1}, t_i]}(f) + \text{Var}_{[0, t_{i-1}]}(f) - \text{Var}_{[0, t_{i-1}]}(f) \\ &\leq \text{Var}_{[0, t_i]}(f) - \text{Var}_{[0, t_{i-1}]}(f) \\ &= \varphi(t_i) - \varphi(t_{i-1}) \\ &= \frac{M}{n}. \end{aligned}$$

Setze

$$m_i(f) = \frac{1}{2} \max_{t_{i-1} \leq x \leq t_i} f(x) + \frac{1}{2} \min_{t_{i-1} \leq x \leq t_i} f(x)$$

und

$$u = \sum_{i=1}^n m_i(f) \chi_{[t_{i-1}, t_i]} \in \Sigma_n.$$

Für jedes i ist dann

$$\begin{aligned} \max_{t_{i-1} \leq x \leq t_i} |f(x) - u(x)| &\leq \frac{1}{2} \left[\max_{t_{i-1} \leq x \leq t_i} f(x) - \min_{t_{i-1} \leq x \leq t_i} f(x) \right] \\ &\leq \frac{1}{2} \text{Var}_{[t_{i-1}, t_i]}(f) \\ &\leq \frac{M}{2n}. \end{aligned}$$

Also ist

$$\sigma_n(f) \leq \|f - u\| \leq \frac{M}{2n}.$$

„ \implies “: Es gibt ein $M > 0$ und eine Folge $(u_n)_{n \in \mathbb{N}}$ mit $u_n \in \Sigma_n$ und $\|f - u_n\| \leq \frac{M}{n}$ für jedes n . Sei $0 \leq x_0 \leq \dots \leq x_m \leq 1$ beliebig und $\nu_{k,n}$ die Zahl der verschiedenen Werte, die u_n auf $[x_{k-1}, x_k)$ annimmt. Falls $\nu_{k,n} = 1$ ist, d.h., falls u_n auf $[x_{k-1}, x_k)$ konstant ist, folgt

$$\begin{aligned} |f(x_k) - f(x_{k-1})| &\leq |f(x_k) - u_n(x_k)| + \underbrace{|u_n(x_k) - u_n(x_{k-1})|}_{=0} \\ &\quad + |f(x_{k-1}) - u_n(x_{k-1})| \\ &\leq 2 \|f - u_n\|. \end{aligned}$$

Sei nun $\nu_{k,n} > 1$ und $x_{k-1} < t_1 < \dots < t_{\nu_{k,n}-1} < x_k$ die zugehörigen Sprungstellen von u_n . Dann folgt wie oben

$$|f(x_k) - f(x_{k-1})| \leq 2 \|f - u_n\| + |u_n(x_k) - u_n(x_{k-1})|.$$

Da f stetig ist, gilt

$$\begin{aligned} |u_n(x_k) - u_n(x_{k-1})| &= |u_n(t_{\nu_{k,n}-1} + 0) - u_n(t_1 - 0)| \\ &= |u_n(t_{\nu_{k,n}-1} + 0) - u_n(t_{\nu_{k,n}-1} - 0) \\ &\quad + \underbrace{u_n(t_{\nu_{k,n}-1} - 0) - u_n(t_{\nu_{k,n}-2} + 0)}_{=0} \\ &\quad + \dots + \\ &\quad + u_n(t_1 + 0) - u_n(t_1 - 0)| \\ &\leq \sum_{i=1}^{\nu_{k,n}-1} |u_n(t_i + 0) - u_n(t_i - 0)| \\ &\leq \sum_{i=1}^{\nu_{k,n}-1} \{|u_n(t_i + 0) - f(t_i + 0)| \\ &\quad + \underbrace{|f(t_i + 0) - f(t_i - 0)|}_{=0} \\ &\quad + |f(t_i - 0) - u_n(t_i - 0)|\} \\ &\leq (\nu_{k,n} - 1) 2 \|f - u_n\|. \end{aligned}$$

Insgesamt erhalten wir also

$$|f(x_k) - f(x_{k-1})| \leq 2\nu_{k,n} \|f - u_n\|$$

und

$$\sum_{k=1}^m |f(x_k) - f(x_{k-1})| \leq 2 \|f - u_n\| \sum_{k=1}^m \nu_{k,n}.$$

Die Zerlegung $0 \leq x_0 \leq \dots \leq x_m \leq 1$ und die zu u_n gehörige Zerlegung ergeben zusammen eine verfeinerte Zerlegung von $[0, 1]$ aus höchstens $m + n$ Intervallen, auf denen u_n stückweise konstant ist. Daher ist

$$\sum_{k=1}^m \nu_{k,n} \leq m + n$$

und

$$\sum_{k=1}^m |f(x_k) - f(x_{k-1})| \leq 2 \frac{M}{n} (m + n) = 2M \left(1 + \frac{m}{n}\right).$$

Da die linke Seite dieser Ungleichung nicht von n abhängt, folgt

$$\sum_{k=1}^m |f(x_k) - f(x_{k-1})| \leq 2M.$$

Da die Zerlegung $0 \leq x_0 \leq \dots \leq x_m \leq 1$ beliebig war, folgt $f \in \text{BV}(0, 1)$. \square

BEMERKUNG VI.4.5. Wir vergleichen die Sätze [VI.3.2](#) (S. 110) und [VI.4.4](#). Für eine feste Unterteilung verhält sich der Approximationsfehler wie $O(n^{-1})$ genau dann, wenn $f \in \text{Lip } 1$ ist. Für eine beliebige, variable Unterteilung ist diese Konvergenzrate äquivalent zu $f \in \text{BV}(0, 1)$. Das Beispiel $f(x) = x^\alpha$ mit $0 < \alpha < 1$ zeigt, dass die zweite Bedingung wesentlich schwächer ist. Grob gesagt, bedeutet $f \in \text{Lip } 1$, dass $f' \in L^\infty$ ist, und $f \in \text{BV}$, dass $f' \in L^1$ ist.

VI.5. Adaptive Approximation

Das nichtlineare Approximationsproblem des vorigen Abschnittes ist für die Praxis nur bedingt geeignet, da es i.a. zu aufwändig ist, die optimale Unterteilung zu bestimmen. Stattdessen bestimmt man in der Regel eine nahezu optimale Unterteilung adaptiv. Für die Beschreibung dieses Prozesses bezeichnen wir mit \mathcal{I} die Menge aller abgeschlossenen, perfekten Teilintervalle in $[0, 1]$ und mit \mathcal{D} die Menge aller dyadischen Intervalle. Jedes $I \in \mathcal{I}$ ist also von der Form $I = [a, b]$ mit $0 \leq a < b \leq 1$, und jedes $D \in \mathcal{D}$ ist von der Form $D = [(k-1)2^{-n}, k2^{-n}]$ mit $n \in \mathbb{N}^*$ und $1 \leq k \leq 2^n$. Wir nehmen an, dass wir für eine feste, gegebene, differenzierbare Funktion f über eine Funktion $\mathcal{E} : \mathcal{I} \rightarrow \mathbb{R}$ verfügen, die folgende Eigenschaften hat:

- $\inf_{c \in \mathbb{R}} \|f - c\|_{L^\infty(I)} \leq \mathcal{E}(I)$ für jedes $I \in \mathcal{I}$,
- $\mathcal{E}(I_1) + \mathcal{E}(I_2) \leq \mathcal{E}(I_1 \cup I_2)$ für alle disjunkten $I_1, I_2 \in \mathcal{I}$,
- $\mathcal{E}(I) \leq \int_I |f'(y)| dy$ für jedes $I \in \mathcal{I}$.

Beispiele für derartige Funktionen sind

- $\frac{1}{2} \left[\max_{x \in I} f(x) - \min_{x \in I} f(x) \right],$
- $\text{Var}_I(f),$
- $\int_I |f'(y)| dy.$

Mit Hilfe einer solchen Funktion \mathcal{E} können wir folgenden Algorithmus durchführen.

ALGORITHMUS VI.5.1 (Adaptive Approximation). Gegeben: *Eine Funktion f und eine Toleranz $\varepsilon > 0$.*

Gesucht: *Eine stückweise konstante Funktion u_ε mit $\|f - u_\varepsilon\| \leq \varepsilon$.*

- (1) Vorbereitung: *Falls $\mathcal{E}([0, 1]) \leq \varepsilon$ ist, setze $\mathcal{G} = \{[0, 1]\}$, $\mathcal{B} = \emptyset$ und gehe zu (3). Andernfalls setze $\mathcal{G} = \emptyset$ und $\mathcal{B} = \{[0, 1]\}$ und gehe zu (2).*
- (2) Allgemeiner Schritt: *Betrachte ein Intervall $I \in \mathcal{B}$ und bezeichne mit J_1 und J_2 die beiden Kinder von I , die durch Halbieren von I entstehen. Falls $\mathcal{E}(J_i) \leq \varepsilon$ ist, füge J_i zu \mathcal{G} hinzu. Andernfalls füge J_i zu \mathcal{B} hinzu. Entferne I aus \mathcal{B} .*
- (3) Ende: *Falls $\mathcal{B} = \emptyset$ ist, beende das Verfahren. Setze $\mathcal{G}_\varepsilon = \mathcal{G}_\varepsilon(f) = \mathcal{G}$. Bestimme zu jedem $I \in \mathcal{G}_\varepsilon$ ein $c_I \in \mathbb{R}$ mit*

$$\|f - c_I\|_{L^\infty(I)} \leq \mathcal{E}(I)$$

(z.B. $c_I = \frac{1}{2} \max_{x \in I} f(x) + \frac{1}{2} \min_{x \in I} f(x)$) und setze

$$u_\varepsilon = \sum_{I \in \mathcal{G}_\varepsilon} c_I \chi_I.$$

Da für jedes $I \in \mathcal{G}_\varepsilon$ konstruktionsgemäß $\mathcal{E}(I) \leq \varepsilon$ ist, gilt

$$\|f - u_\varepsilon\|_\infty \leq \varepsilon.$$

Die Effizienz von Algorithmus VI.5.1 hängt offensichtlich von $N_\varepsilon = N_\varepsilon(f) = \#\mathcal{G}_\varepsilon$ ab.

SATZ VI.5.2. *Sei f differenzierbar und $f' \in L^\infty([0, 1])$. Dann ist*

$$N_\varepsilon \leq \frac{2}{\varepsilon} \|f'\|_\infty.$$

BEWEIS. Sei $I \in \mathcal{G}_\varepsilon$ und \bar{I} ein Vorfahr von I , d.h., I entsteht durch Halbieren aus \bar{I} . Dann ist wegen der dritten Eigenschaft von \mathcal{E} und der Konstruktion von \mathcal{G}_ε

$$\varepsilon < \mathcal{E}(I) \leq \int_I |f'| \leq |\bar{I}| \|f'\|_\infty = 2|I| \|f'\|_\infty.$$

Da die Intervalle in \mathcal{G}_ε paarweise disjunkt sind, folgt

$$\varepsilon N_\varepsilon = \sum_{I \in \mathcal{G}_\varepsilon} \varepsilon \leq \sum_{I \in \mathcal{G}_\varepsilon} 2|I| \|f'\|_\infty \leq 2 \|f'\|_\infty. \quad \square$$

Um die adaptive Approximation mit der nichtlinearen Approximation des vorigen Abschnittes zu vergleichen, setzen wir

$$\alpha_n(f) = \inf\{\varepsilon > 0 : N_\varepsilon(f) \leq n\}.$$

Aus Satz VI.5.2 folgt mit der Wahl $\varepsilon = 2 \|f\|_\infty n^{-1}$ die Abschätzung

$$\alpha_n(f) \leq \|f - u_\varepsilon\|_\infty \leq \frac{2 \|f'\|_\infty}{n}.$$

Diese ist zu vergleichen mit der Abschätzung

$$\sigma_n(f) \leq \frac{\text{Var}_{[0,1]}(f)}{2n} \leq \frac{\|f'\|_1}{2n},$$

die wir im Beweis von Satz VI.4.4 (S. 117) erhalten haben. Man kann die Lücke zwischen diesen beiden Abschätzungen verkleinern, wenn man schärferes mathematisches Geschütz aufführt.

DEFINITION VI.5.3. (1) Für $f \in L^1([0, 1])$ ist die *maximale Funktion* $Mf : [0, 1] \rightarrow \mathbb{R}$ definiert durch

$$Mf(x) = \sup_{I; x \in I} \frac{1}{|I|} \int_I |f(y)| dy.$$

(2) $LlogL$ ist der Raum aller messbaren Funktionen f auf $[0, 1]$, für die $f[1 + \ln(|f|)]$ integrierbar ist. Für $f \in LlogL$ setze

$$\|f\|_{LlogL} = \int_0^1 |f(x)| |1 + \ln(|f(x)|)| dx.$$

BEMERKUNG VI.5.4. (1) Ist $f \in L^\infty$, so ist $Mf(x) \leq \|f\|_\infty$ für alle x .

(2) Für $p > 1$ ist $L^p \subset LlogL$.

(3) Es ist $LlogL \not\subset L^1$.

SATZ VI.5.5. *Es gibt eine Konstante $c > 0$, so dass für alle $f \in LlogL$ gilt*

$$\int_0^1 Mf(x) dx \leq c \|f\|_{LlogL}.$$

BEWEIS. [1, S. 243–246] □

SATZ VI.5.6. *Sei f differenzierbar und $f' \in LlogL$. Dann ist*

$$N_\varepsilon \leq \frac{2c}{\varepsilon} \|f'\|_{LlogL}$$

mit der Konstanten c aus Satz VI.5.5.

BEWEIS. Mit den Bezeichnungen aus Satz VI.5.2 haben wir

$$\varepsilon < \mathcal{E}(\bar{I}) \leq \int_{\bar{I}} |f'(y)| dy = |\bar{I}| \frac{1}{|\bar{I}|} \int_{\bar{I}} |f'(y)| dy \leq |\bar{I}| Mf'(x)$$

für alle $x \in \bar{I}$. Also ist

$$\varepsilon \leq \inf_{x \in I} |\bar{I}| M f'(x) \leq |\bar{I}| \frac{1}{|I|} \int_I M f'(y) dy = 2 \int_I M f'(y) dy.$$

Hieraus folgt

$$\varepsilon N_\varepsilon \leq 2 \sum_I \int_I M f'(y) dy \leq 2 \int_0^1 M f'(y) dy \leq 2c \|f'\|_{L \log L}. \quad \square$$

Aus Satz VI.5.6 folgt die verbesserte Abschätzung

$$\alpha_n(f) \leq \|f - u_\varepsilon\|_\infty \leq \frac{2c \|f'\|_{L \log L}}{n}.$$

Grob gesprochen ist diese Abschätzung nur um einen Faktor $\ln(|f'|)$ schlechter als die Abschätzung für $\sigma_n(f)$.

VI.6. Haar-Wavelets

In diesem Abschnitt bezeichnet $\|\cdot\|$ stets die L^2 -Norm auf $[0, 1]$ und (\cdot, \cdot) das zugehörige Skalarprodukt. Wir erinnern an die Zerlegungen Δ_n aus § VI.3 (S. 108) und die zugehörigen Räume $S^0(\Delta_n)$ stückweise konstanter Funktionen. Für $k \in \mathbb{N}$ setze $S_k = S^0(\Delta_{2^k})$ und bezeichne mit $P_k : L^2([0, 1]) \rightarrow S_k$ den zugehörigen orthogonalen Projektor.



ABBILDUNG VI.6.1. Funktionen $\varphi_{0,1}$ links und $\varphi_{1,2}$ rechts

DEFINITION VI.6.1. Bezeichne mit $\varphi = \chi_{[0,1]}$ die charakteristische Funktion des Intervalls $[0, 1]$ und definiere für $k \in \mathbb{N}$ und $0 \leq j \leq 2^k - 1$ die Funktion $\varphi_{j,k}$ durch (vgl. Abbildung VI.6.1)

$$\varphi_{j,k}(x) = 2^{\frac{k}{2}} \varphi(2^k x - j).$$

SATZ VI.6.2. Für alle $k \in \mathbb{N}$ und alle $f \in L^2([0, 1])$ ist

$$P_k f = \sum_{j=0}^{2^k-1} (f, \varphi_{j,k}) \varphi_{j,k}.$$

BEWEIS. Offensichtlich ist $\varphi_{j,k} = 2^{\frac{k}{2}} \chi_{[j2^{-k}, (j+1)2^{-k}]}$. Daher sind $\varphi_{0,k}, \dots, \varphi_{2^k-1,k}$ paarweise orthogonal und haben jeweils die L^2 -Norm 1. Hieraus folgt die Behauptung. \square

Die Räume S_0, S_1, \dots bilden eine Skala, d.h., es ist $S_k \subset S_{k+1}$ für jedes k . Bezeichne mit W_k das orthogonale Komplement von S_k in S_{k+1} und mit Q_k den zugehörigen Projektor.

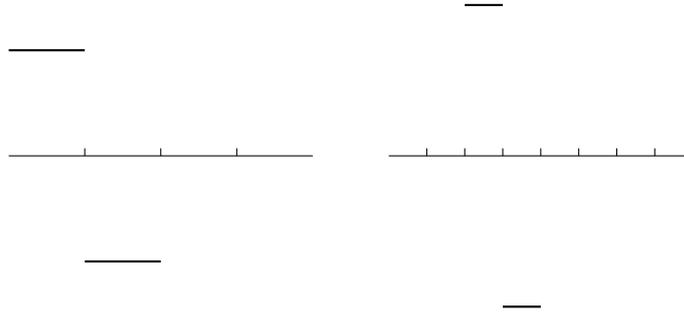


ABBILDUNG VI.6.2. Funktionen $\psi_{0,1}$ links und $\psi_{1,2}$ rechts

DEFINITION VI.6.3. Bezeichne mit

$$\psi = H = \chi_{[0, \frac{1}{2})} - \chi_{[\frac{1}{2}, 1)} = \begin{cases} 1 & \text{für } 0 \leq x < \frac{1}{2}, \\ -1 & \text{für } \frac{1}{2} \leq x < 1 \end{cases}$$

die *Haar-Funktion* des Intervalles $[0, 1)$ und definiere für $k \in \mathbb{N}$ und $0 \leq j \leq 2^k - 1$ die Funktion $\psi_{j,k} = H_{j,k}$ durch (vgl. Abbildung VI.6.2)

$$\psi_{j,k}(x) = H_{j,k}(x) = 2^{\frac{k}{2}} \psi(2^k x - j).$$

Die Funktionen $\psi_{j,k}$ heißen *Haar-Wavelets*.

SATZ VI.6.4. Für jedes $k \in \mathbb{N}$ und jedes $f \in L^2([0, 1])$ ist

$$Q_k f = \sum_{j=0}^{2^k-1} (f, \psi_{j,k}) \psi_{j,k}.$$

BEWEIS. Offensichtlich ist

$$\psi_{j,k} = 2^{\frac{k}{2}} [\chi_{[2^j 2^{-k-1}, (2j+1)2^{-k-1})} - \chi_{[(2j+1)2^{-k-1}, (2j+2)2^{-k-1})}].$$

Daher sind $\psi_{0,k}, \dots, \psi_{2^k-1,k}$ paarweise orthogonal, haben jeweils die L^2 -Norm 1, liegen in S_{k+1} und erfüllen für jedes j und k

$$\int_0^1 \psi_{j,k} = \int_{j2^{-k}}^{(j+1)2^{-k}} \psi_{j,k} = 0.$$

Hieraus folgt $\psi_{j,k} \perp S_k$ für jedes j und k . Dies beweist die Behauptung. \square

SATZ VI.6.5. Für jedes $f \in L^2([0, 1])$ und jedes $m \in \mathbb{N}$ ist

$$P_m f = (f, \varphi) + \sum_{k=0}^{m-1} \sum_{j=0}^{2^k-1} (f, \psi_{j,k}) \psi_{j,k}$$

und

$$f = (f, \varphi) + \sum_{k=0}^{\infty} \sum_{j=0}^{2^k-1} (f, \psi_{j,k}) \psi_{j,k}.$$

Insbesondere ist

$$\|f\|^2 = |(f, \varphi)|^2 + \sum_{k=0}^{\infty} \sum_{j=0}^{2^k-1} |(f, \psi_{j,k})|^2.$$

BEWEIS. Aus der Definition von W_k und Q_k folgt

$$Q_k = P_{k+1} - P_k$$

und damit

$$P_m = P_0 + \sum_{k=0}^{m-1} Q_k.$$

Hieraus folgt mit Satz VI.6.4 die erste Behauptung.

Aus Satz VI.3.9 (S. 114) folgt

$$(VI.6.1) \quad \lim_{m \rightarrow \infty} \|f - P_m f\| = 0.$$

Dies beweist die zweite Behauptung.

Die letzte Behauptung folgt aus (VI.6.1) und dem Beweis von Satz VI.2.1 (S. 102), in dem wir

$$\|f\|^2 = \sum_{k=0}^{\infty} |f_k|^2$$

gezeigt haben, wenn f_k die Fourier-Koeffizienten von f bzgl. einer separablen Basis sind. \square

Satz VI.6.5 hat enorme praktische Bedeutung und zeigt die Überlegenheit der Haar-Wavelets. Angenommen, wir haben bereits $P_m f$ berechnet und wollen nun $P_{m+1} f$ zur Verbesserung der Approximation bestimmen. Dann ist

$$P_{m+1} f = P_m f + \sum_{j=0}^{2^m-1} (f, \psi_{j,m}) \psi_{j,m}.$$

Wir müssen also nur die neuen Koeffizienten $(f, \psi_{j,m})$ mit $0 \leq j \leq 2^m - 1$ berechnen. Dies kann rekursiv geschehen. Um dies einzusehen, beachten wir, dass aus den Definitionen VI.6.1 und VI.6.3 folgt

$$\begin{aligned} \varphi_{j,k} &= \frac{1}{\sqrt{2}} [\varphi_{2j,k+1} + \varphi_{2j+1,k+1}], \\ \psi_{j,k} &= \frac{1}{\sqrt{2}} [\varphi_{2j,k+1} - \varphi_{2j+1,k+1}]. \end{aligned}$$

Daher gelten die Rekursionsformeln

$$\begin{aligned} (f, \varphi_{j,k}) &= \frac{1}{\sqrt{2}} [(f, \varphi_{2j,k+1}) + (f, \varphi_{2j+1,k+1})] \\ (f, \psi_{j,k}) &= \frac{1}{\sqrt{2}} [(f, \varphi_{2j,k+1}) - (f, \varphi_{2j+1,k+1})] \end{aligned}$$

und

$$\begin{aligned}(f, \varphi_{2j,k+1}) &= \frac{1}{\sqrt{2}} [(f, \varphi_{j,k}) + (f, \psi_{j,k})] \\ (f, \varphi_{2j+1,k+1}) &= \frac{1}{\sqrt{2}} [(f, \varphi_{j,k}) - (f, \psi_{j,k})].\end{aligned}$$

Sind also die Koeffizienten $(f, \varphi_{j,m})$ für $0 \leq j \leq 2^m - 1$ bekannt, können alle Koeffizienten der Darstellung

$$P_m f = (f, \varphi) + \sum_{k=0}^{m-1} \sum_{j=0}^{2^k-1} (f, \psi_{j,k}) \psi_{j,k}$$

rekursiv berechnet werden. Sind umgekehrt alle Koeffizienten dieser Darstellung bekannt, können die Koeffizienten $(f, \varphi_{j,k})$ für $0 \leq k \leq m$ und $0 \leq j \leq 2^k - 1$ rekursiv berechnet werden¹.

Wegen Satz VI.6.5 erfüllen die Funktionen $\psi_{j,k}$ die Voraussetzungen an die Basis u_n aus § VI.2 (S. 101). Daher können wir die dortigen Ergebnisse über Approximationsräume verwenden. Für die Praxis wesentlich ist die Erkenntnis, dass man in diesem Fall die abstrakten Räume durch Glattheitseigenschaften der Funktion f charakterisieren kann. So erhält man folgendes Ergebnis.

SATZ VI.6.6. Für $p > 0$ sei

$$\omega_p(f; h) = \sup_{0 < t \leq h} \|f(\cdot + t) - f(\cdot)\|_p.$$

Dann gilt für den Fehler $\rho_n(f)$ der besten linearen Approximation in S_n

$$\begin{aligned}\rho_n(f) = O(n^{-\alpha}) &\iff \sup_{t>0} t^{-\alpha} \omega_2(f; t) < \infty, \\ \sum_{n \in \mathbb{N}} [n^\alpha \rho_n(f)]^2 \frac{1}{n} < \infty &\iff \int_0^\infty [t^{-\alpha} \omega_2(f; t)]^2 \frac{1}{t} dt < \infty.\end{aligned}$$

Für den Fehler $\sigma_n(f)$ der besten nichtlinearen Approximation mit n Termen gilt

$$\sigma_n(f) = O(n^{-\alpha}) \iff \int_0^\infty [t^{-\alpha} \omega_\tau(f; t)]^\tau \frac{1}{t} dt < \infty$$

mit $\frac{1}{\tau} = \alpha + \frac{1}{2}$.

¹Vergleiche die Algorithmen I.5.1 und I.5.2 von Goertzel zur trigonometrischen Interpolation und die schnelle Fourier-Transformation aus §I.5 der Vorlesung "Einführung in die Numerik".

VI.7. Haar-Wavelets im \mathbb{R}^d und glatte Wavelets

Die wesentlichen Strukturen, die dem vorigen Abschnitt zugrunde liegen, sind:

- Die Konstruktion einer Basis von L^2 aus einer Funktion mittels Translation und Dilatation.
- Eine hierarchische oder Multi-Skalen Zerlegung von L^2 in orthogonale Räume.

Wir wollen in diesem Abschnitt skizzieren, wie man diese Konzepte auf eine allgemeine Basis stellen kann und wie man dadurch die Ergebnisse des vorigen Abschnittes auf mehrere Raumdimensionen und glatte Basisfunktionen ausdehnen kann.

Die Grundoperationen sind *Translation* und *Dilatation*. Sind $f : \mathbb{R}^d \rightarrow \mathbb{R}$ eine Funktion, $j \in \mathbb{Z}^d$ und $a \in \mathbb{R}^*$, so sind das *Translat* von f und sein *Dilatat* definiert durch

$$\begin{aligned} f(\cdot - j) & \text{ d.h. } x \mapsto f(x - j) \\ f(a \cdot) & \text{ d.h. } x \mapsto f(ax). \end{aligned}$$

Sei nun $\varphi \in L^2(\mathbb{R}^d)$ eine Funktion mit kompaktem Träger. In § VI.6 hatten wir den Spezialfall $\varphi = \chi_{[0,1]}$. Bezeichne mit $\tilde{S}(\varphi)$ den Raum aller endlichen Linearkombinationen von Translaten von φ und mit $S = S(\varphi)$ den Abschluss von $\tilde{S}(\varphi)$ in $L^2(\mathbb{R}^d)$. S heißt der *shift invariante Raum* generiert von φ . Weiter setzen wir voraus, dass die Translate von φ eine orthonormale Basis von S bilden. In § VI.6 ist diese Bedingung erfüllt. Diese Voraussetzung ist für die allgemeine Theorie nicht notwendig, erleichtert aber die Darstellung enorm.

Für $k \in \mathbb{Z}$ ist $S_k = S_k(\varphi)$ der Raum der Dilatate um 2^k der Funktionen aus S . Das heißt, es ist $f \in S_k$ genau dann, wenn $f = g(2^k \cdot)$ ist für ein $g \in S$.

Die wesentliche Voraussetzung ist nun, dass die Räume S_k eine *Skala* oder *Hierarchie* bilden, d.h.

$$(VI.7.1) \quad S_k \subset S_{k+1} \quad \forall k \in \mathbb{Z}.$$

Wegen der Konstruktion der Räume durch Dilatation ist (VI.7.1) äquivalent zu

$$(VI.7.2) \quad S_0 \subset S_1.$$

Wegen der Konstruktion von S durch Translation ist (VI.7.2) äquivalent zu $\varphi \in S_1$ und somit zu

$$(VI.7.3) \quad \varphi(x) = \sum_{k \in \mathbb{Z}^d} c_k \varphi(2x - k) \quad \forall x \in \mathbb{R}^d.$$

Die Koeffizienten c_k heißen *Verfeinerungsmaske* von φ . Da φ kompakten Träger hat, sind nur endlich viele c_k von Null verschieden. Im Rahmen von § VI.6 ist

$$c_0 = c_1 = 1, \quad c_k = 0 \quad \forall k \in \mathbb{Z} \setminus \{0, 1\}.$$

Sei W das orthogonale Komplement von S_0 in S_1 . W heißt *Wavelet-Raum*. Er stellt die Details dar, die beim Übergang von S_0 nach S_1 gewonnen werden. Das wesentliche Ergebnis der Theorie ist die Tatsache, dass W ebenfalls ein shift invarianter Raum ist, der durch die Funktion

$$(VI.7.4) \quad \psi(x) = \sum_{k \in \mathbb{Z}^d} (-1)^k c_{e-k} \varphi(2x - k)$$

erzeugt wird. Dabei ist $e = (1, \dots, 1) \in \mathbb{Z}^d$ und $(-1)^k = (-1)^{k_1} \dots (-1)^{k_d}$. Im Rahmen von § VI.6 ist ψ genau die Haar-Funktion H .

Bezeichne mit $P : L^2(\mathbb{R}^d) \rightarrow S$ den orthogonalen Projektor. Dann ist

$$Pf = \sum_{j \in \mathbb{Z}^d} (f, \varphi(\cdot - j)) \varphi(\cdot - j),$$

wenn (\cdot, \cdot) das L^2 -Skalarprodukt bezeichnet. Durch Dilatation erhalten wir den Projektor $P_k : L^2(\mathbb{R}^d) \rightarrow S_k$. Aus dem Transformationssatz folgt²

$$P_k f = \sum_{j \in \mathbb{Z}^d} 2^{kd} (f, \varphi(2^k \cdot - j)) \varphi(2^k \cdot - j).$$

Sei $Q = P_1 - P_0$ der orthogonale Projektor auf W . Dann ist

$$Qf = \sum_{j \in \mathbb{Z}^d} (f, \psi(\cdot - j)) \psi(\cdot - j).$$

Sei W_k das orthogonale Komplement von S_k in S_{k+1} und $Q_k = P_{k+1} - P_k$ der zugehörige orthogonale Projektor. Aus der Konstruktion der S_k und von W folgt, dass W_k das Dilatat von W ist und dass entsprechend gilt

$$Q_k f = \sum_{j \in \mathbb{Z}^d} 2^{kd} (f, \psi(2^k \cdot - j)) \psi(2^k \cdot - j).$$

Wir setzen nun voraus, dass $\bigcup_{k \in \mathbb{N}} S_k$ dicht ist in $L^2(\mathbb{R}^d)$. Dies ist äquivalent zu

$$\lim_{k \rightarrow \infty} \|f - P_k f\|_2 = 0 \quad \forall f \in L^2(\mathbb{R}^d).$$

Man kann zeigen, dass hieraus für alle $f \in L^2$ folgt

$$\lim_{k \rightarrow -\infty} P_k f = 0.$$

Daher gilt für alle $f \in L^2$

$$\begin{aligned} f &= \sum_{k \in \mathbb{Z}} (P_{k+1} f - P_k f) = \sum_{k \in \mathbb{Z}} Q_k f \\ &= \sum_{k \in \mathbb{Z}} \sum_{j \in \mathbb{Z}^d} 2^{kd} (f, \psi(2^k \cdot - j)) \psi(2^k \cdot - j). \end{aligned}$$

Dies ist die *Wavelet-Darstellung* von f .

²Wegen $\int_{\mathbb{R}^d} \varphi(x)^2 dx = 1$ ist $\int_{\mathbb{R}^d} \varphi(2^k x)^2 dx = 2^{-kd} \int_{\mathbb{R}^d} \varphi(y)^2 dy = 2^{-kd}$.

Wegen der Beziehungen (VI.7.3) und (VI.7.4) und der Tatsache, dass nur endlich viele Koeffizienten c_k von Null verschieden sind, gilt eine zu § VI.6 analoge Rekursionsformel für die Koeffizienten $(f, \varphi(2^k \cdot -j))$ und $(f, \psi(2^k \cdot -j))$. Dies ist die so genannte *schnelle Wavelet-Transformation*.

Die bisherigen Ausführungen sind insofern von akademischer Natur, als wir bisher nur das Beispiel aus § VI.6 für ein eindimensionales, nicht glattes Wavelet kennen. Kann man Wavelets in \mathbb{R}^d konstruieren, die glatt, d.h. in C^r mit beliebigem vorgegebenem r sind? Dies ist aufgrund einer berühmten Konstruktion von (Baronesse³) I. Daubechies tatsächlich möglich. Die Argumente sind zu technisch, um sie im Detail auszuführen, aber wir wollen die Grundidee doch skizzieren.

Wegen der Tensorproduktstruktur von $L^2(\mathbb{R}^d)$ (Satz von Fubini!) kann man Funktionen $\varphi \in L^2(\mathbb{R}^d)$ mit den gewünschten Eigenschaften aus entsprechenden Funktionen $\tilde{\varphi} \in L^2(\mathbb{R})$ durch den Ansatz

$$\varphi(x_1, \dots, x_d) = \tilde{\varphi}(x_1) \cdot \dots \cdot \tilde{\varphi}(x_d)$$

konstruieren. (Die ist nicht die einzige Möglichkeit!) Wir beschränken uns daher auf den Fall $d = 1$.

Die Grundidee ist, die Koeffizienten c_k in (VI.7.3) vorzugeben und aus ihren Eigenschaften auf Eigenschaften der Funktion φ zu schließen.

Sei also $(c_k)_{k \in \mathbb{Z}}$ mit $c_k = 0$ für $|k| > m$ und $c_m \neq 0$ gegeben. Setze

$$\mathcal{A}(y) = \frac{1}{2} \sum_{k=-m}^m c_k e^{-iky}.$$

Dann erfüllt die Fourier-Transformierte $\hat{\varphi}$ der Funktion φ aus (VI.7.3) die Funktionalgleichung

$$\hat{\varphi}(y) = \mathcal{A}\left(\frac{y}{2}\right) \hat{\varphi}\left(\frac{y}{2}\right).$$

Die Translate von φ sind genau dann orthonormal, wenn gilt

$$(VI.7.5) \quad |\mathcal{A}(y)|^2 + |\mathcal{A}(y + \pi)|^2 = 1 \quad \forall y \in \mathbb{R}.$$

Aus der Annahme $\int_{\mathbb{R}} \varphi(x) dx \neq 0$ und der Beziehung (VI.7.3) folgt durch Integration wegen des Transformationssatzes

$$\sum_{k \in \mathbb{Z}} c_k = 2.$$

Dies impliziert

$$\mathcal{A}(0) = 1 \quad \text{und} \quad \mathcal{A}(\pi) = 0.$$

³Frau Daubechies wurde 2012 wegen ihrer mathematischen(!) Verdienste vom belgischen König in den Adelsstand erhoben.

Daher gibt es ein $N \in \mathbb{N}^*$, eine Funktion α und eine Zahl θ mit

$$(VI.7.6) \quad \begin{aligned} \mathcal{A}(y) &= (1 + e^{iy})^N \alpha(y), \\ \|\alpha\|_\infty &= 2^{-\theta}, \\ \alpha(0) &= 2^{-N}. \end{aligned}$$

Wir suchen α in der Klasse der trigonometrischen Polynome mit reellen Koeffizienten. Dann ist

$$|\alpha(y)|^2 = \alpha(y)\overline{\alpha(y)} = \alpha(y)\alpha(-y)$$

ein gerades trigonometrisches Polynom und daher von der Form

$$|\alpha(y)|^2 = T(\cos y) = T\left(1 - 2\sin^2 \frac{y}{2}\right) = R\left(\sin^2 \frac{y}{2}\right)$$

mit einem algebraischen Polynom R . Die Bedingung (VI.7.5) reduziert sich nun auf

$$\left(\cos^2 \frac{y}{2}\right)^N R\left(\sin^2 \frac{y}{2}\right) + \left(\sin^2 \frac{y}{2}\right)^N R\left(\cos^2 \frac{y}{2}\right) = 2^{-2N}.$$

Mit $t = \sin^2 \frac{y}{2}$ gibt dies

$$(VI.7.7) \quad (1-t)^N R(t) + t^N R(1-t) = 2^{-2N}.$$

Wir schreiben R nun in der Bernstein-Form

$$R(t) = \sum_{k=0}^{N-1} \lambda_k \binom{N-1}{k} t^k (1-t)^{N-1-k}.$$

Dann wird aus (VI.7.7)

$$\begin{aligned} 2^{-2N} &= (1-t)^N \sum_{k=0}^{N-1} \lambda_k \binom{N-1}{k} t^k (1-t)^{N-1-k} \\ &\quad + t^N \sum_{k=0}^{N-1} \lambda_k \binom{N-1}{k} (1-t)^k t^{N-1-k} \\ &= \sum_{k=0}^{N-1} \lambda_k \binom{N-1}{k} t^k (1-t)^{2N-1-k} \\ &\quad + \sum_{k=0}^{N-1} \lambda_k \binom{N-1}{k} t^{2N-1-k} (1-t)^k \\ &= \sum_{k=0}^{N-1} \lambda_k \binom{N-1}{k} t^k (1-t)^{2N-1-k} \\ &\quad + \sum_{\ell=N}^{2N-1} \lambda_{2N-1-\ell} \binom{N-1}{2N-1-\ell} t^\ell (1-t)^{2N-1-\ell}. \end{aligned}$$

Wegen

$$2^{-2N} = 2^{-2N} \sum_{m=0}^{2N-1} \binom{2N-1}{m} t^m (1-t)^{2N-1-m}$$

ist

$$\lambda_k = 2^{-2N} \frac{\binom{2N-1}{k}}{\binom{N-1}{k}} \quad k = 0, 1, \dots, N-1$$

eine Lösung. Das zugehörige Polynom nennen wir R_N . Da die λ_k nicht negativ sind, ist R_N auf $[0, 1]$ nicht negativ. Wegen eines klassischen Satzes von Féjer-Riesz gibt es daher ein trigonometrisches Polynom α_N mit

$$(VI.7.8) \quad R_N \left(\sin^2 \frac{y}{2} \right) = |\alpha_N(y)|^2.$$

Die Funktion

$$\mathcal{A}_N(y) = (1 + e^{iy})^N \alpha_N(y)$$

erfüllt die Beziehung (VI.7.5), da R_N die Beziehung (VI.7.7) erfüllt. Also hat die zugehörige Funktion φ einen kompakten Träger. Man kann zeigen, dass das zugehörige Wavelet ψ einen Träger in $[-(N-1), N]$ hat. Die Frage ist, welche Glattheit ψ hat.

Wegen

$$\alpha_N(0) = \sqrt{R_N(0)} = \sqrt{\lambda_0} = 2^{-N}$$

ist die dritte Bedingung von (VI.7.6) erfüllt. Da die λ_k monoton sind, ist R_N monoton wachsend. Daher ist

$$\max_{0 \leq t \leq 1} R_N(t) = R_N(1) = \lambda_{N-1} = 2^{-2N} \binom{2N-1}{N}.$$

Wegen der Stirlingschen Formel ist deshalb $\|\alpha_N\|_\infty$ beschränkt durch $cN^{-\frac{1}{4}}$. Zu gegebenem $\theta > 0$ können wir nun N so groß wählen, dass die zweite Bedingung von (VI.7.6) ebenfalls erfüllt ist. Daher folgt

$$|\widehat{\varphi}(x)| \leq c(1 + |x|)^{-\theta} \quad \text{für } |x| \rightarrow \infty$$

und damit $\varphi, \psi \in C^r$ für $0 \leq r < \theta - 1$.

Für $N = 1$ liefert diese Konstruktion genau die Haar-Funktion H aus § VI.6. Für $N = 2$ erhält man

$$\begin{aligned} R_2(t) &= \frac{1}{16}(1 + 2t), \\ \mathcal{A}_2(y) &= (1 + e^{iy})^2 \alpha_2(y), \\ \alpha_2(y) &= \frac{\sqrt{3} + 1}{8} - \frac{\sqrt{3} - 1}{8} e^{-iy}. \end{aligned}$$

Die zugehörigen Funktionen φ und ψ sind stetig. Es ist sogar $\psi \in \text{Lip } 0.55$.

Eine Tabelle von Verfeinerungskoeffizienten und Glattheitseigenschaften von φ und ψ in Abhängigkeit von N findet man in [3].

VI.8. Anwendungsbeispiele

Als ein Anwendungsbeispiel betrachten wir die Datenkompression und Rauschunterdrückung bei der Bildverarbeitung.

Ein Bild können wir uns vereinfacht als stückweise konstante Approximation einer Dichtefunktion f vorstellen. Der Bildschirm $\Omega = [0, 1]^2$ wird in 2^N (z.B. 1024 oder 1024×1024) dyadische Quadrate unterteilt und jedem Quadrat wird ein Grauwert zugeordnet. Dies ist die uns bekannte Approximation an f . Im Rahmen der Abschnitte VI.6 und VI.7 bedeutet dies, dass wir die Koeffizienten (f, φ_i) kennen.

Für eine Übertragung eines Bildes möchten wir diese Datenmenge reduzieren und die relevanten Informationen herausfiltern. Dazu sind wir bereit, einen gewissen Informationsverlust zu akzeptieren, d.h., wir suchen eine Approximation \tilde{f} an f mit weniger Koeffizienten, so dass $\|f - \tilde{f}\| \leq \varepsilon$ ist für ein gegebenes ε . Die Lösung mit Hilfe der Wavelets sieht dann so aus:

- (1) Berechne mit Hilfe der schnellen Wavelet-Transformation aus § VI.6 die Wavelet-Koeffizienten (f, ψ_i) aus den bekannten Koeffizienten (f, φ_i) .
- (2) Ordne die Wavelet-Koeffizienten der Größe nach und bestimme so $\gamma_i(f)$.
- (3) Versende nur die ersten $k \ll 2^N$ Koeffizienten $\gamma_i(f)$. Dabei ist k entweder vorgegeben oder wird durch die Forderung

$$\sum_{i>k} |\gamma_i(f)|^2 < \varepsilon^2$$

bestimmt.

- (4) Berechne mit Hilfe der schnellen Wavelet-Transformation aus § VI.6 die Approximation

$$\tilde{f} = \sum_{i \leq k} \gamma_i(f) \psi_{\beta_i}$$

von f . Dabei sind die Indizes β_i so bestimmt, dass $\gamma_i(f) = (f, \psi_{\beta_i})$ ist.

Die Rauschunterdrückung ist analog zur Datenkompression. Jetzt stellt man sich vor, das das Verrauschte Bild einer Darstellung

$$f = \sum_{i \leq 2^N} (f, \varphi_i) \varphi_i$$

mit $N \gg 1$ entspricht, dass man aber eigentlich eine unverrauschte Information

$$\tilde{f} = \sum_{i \leq 2^n} (f, \varphi_i) \varphi_i$$

mit $n \ll N$ hat. Die Berechnung von \tilde{f} erfolgt dann mit $k = 2^n$.

Index

- $\|\cdot\|_\infty$ Maximum-Norm auf \mathbb{R}^n , 8
- $\|\cdot\|_p$ ℓ^p -Norm auf \mathbb{R}^n , 8
- $|\cdot|_{\mathcal{A}^\alpha((X_n))}$ Norm von $\mathcal{A}^\alpha((X_n))$, 103
- $|\mathcal{A}_2^\alpha((X_n))| \cdot$ Norm von $\mathcal{A}_2^\alpha((X_n))$, 104
- $\|\cdot\|_{L\log L}$ Norm von $L\log L$, 121
- $\|\cdot\|$ Norm von X , 7
- $\|\cdot\|_{X'}$ Norm von X' , 23
- $\|\cdot\|_\infty$ Supremum-Norm, 7
- $\|\cdot\|_p$ L^p -Norm, 8
- \mathcal{A}^α Approximationsraum, 103, 107
- \mathcal{A}_2^α Approximationsraum, 104, 107
- $B_n f$ Bernstein-Polynome, 53
- $BV(a, b)$ Raum der Funktionen mit beschränkter Variation auf $[a, b]$, 115
- $C(T)$ Raum der stetigen Funktionen auf einer kompakten Menge T , 7
- $C^{k+\alpha}(I)$ Raum aller k -mal stetig differenzierbaren Funktionen auf I mit $f^{(k)} \in \text{Lip } \alpha$, 58
- $C_{2\pi}$ Raum der 2π -periodischen Funktionen, 7, 58
- Δ_n äquidistante Unterteilung von $[0, 1]$ in n Intervalle, 110
- $E_{n,[a,b]}(f)$ Abstand der b.A. an f in \mathbb{P}_n auf dem Intervall $[a, b]$, 39
- $E_n(f)$ Abstand der b.A. an f in \mathbb{P}_n auf dem Intervall $[-1, 1]$, 39
- $H_{j,k}$ Haar-Wavelet, 123
- I Intervall, 8
- I_τ Quasi-Interpolationsoperator, 111
- J_n , 60
- $K_m(t)$ Jackson-Kern, 60
- $\text{Lip } \alpha$ Lipschitz-Klasse der Ordnung α , 57
- $\text{Lip}_p \alpha$ Lipschitz-Raum in L^p , 111
- $L\log L$, 121
- $L^p(I)$ Raum der zur p -ten Potenz Lebesgue integrierbaren messbaren Funktionen auf einem Intervall I , 8
- $M[f]$ Menge der Extremalpunkte von f , 19
- Mf maximale Funktion zu f , 121
- Φ Abstandsfunktion, 8
- \mathbb{P}_n Raum der algebraischen Polynome vom Grad n , 7, 39
- Σ_n nichtlineare Approximationsmenge, 106, 114
- $S^0(\tau)$ Raum der stückweise konstanten Funktionen zur Zerlegung τ , 109
- T kompakte Menge, 7
- \mathcal{T} Menge aller Zerlegungen, 109
- T_h Translationsoperator, 111
- T_n Čebyšev-Polynome (erster Art), 40
- \mathbb{T}_n Raum der trigonometrischen Polynome vom Grad n , 7, 58
- V Teilmenge von X , 7
- $\text{Var}_{[a,b]}(f)$ totale Variation von f , 115
- X normierter Raum, 7
- X' Raum der beschränkten linearen Funktionale auf X , 23
- $\alpha_n(f)$ adaptive Approximationsgüte, 121
- $\delta_T(Y)$ Dichtheitsmaß, 95
- $\delta(Y)$ Dichtheitsmaß, 95
- δ_τ Gitterweite von τ , 109
- ∂p Grad eines algebraischen Polynomes p , 39
- $\kappa(M)$ konvexe Hülle von M , 24
- $\omega(f; \cdot)$ Stetigkeitsmodul der Funktion f , 56
- $\omega_2(f; \cdot)$ zweite Stetigkeitsmodul der Funktion f , 58
- $\omega_p(f; \cdot)$ Stetigkeitsmodul in L^p , 111
- $\psi_{j,k}$ Haar-Wavelet, 123
- $\rho_V(f)$ Abstand von f zu V , 7

- $\rho_n(f)$ Approximationsgüte einer äquidistanten Unterteilung von $[0, 1]$ in n Intervalle, 110
- $\rho_\tau(f)_p$ Approximationsgüte von $S^0(\tau)$ in $L^p([0, 1])$, 109
- $\sigma_n(f)$ nichtlineare Approximationsgüte, 106, 115
- τ Zerlegung, 108

- Abstand, 7
- Abstandsfunktion, 8
- adaptive Approximation, 120
- Additionstheorem, 5
- Alternante, 32
- Alternantenkriterium, 32
- Anfangsbedingung, 6
- Angleichung von Messdaten, 5
- Approximationsproblem, 5
- Approximationsraum, 103, 104
- approximativ kompakt, 10
- Ascent-Verfahren, 87

- b.A. beste Approximation, 7
- Bernstein-Polynome, 53
- Bernstein-Ungleichung, 69
- beschränkt kompakt, 10
- beschränkte Variation, 115
- beschränktes lineares Funktional, 23
- beste Approximation, 5, 7
- beste Approximation durch Polynome, 5
- Browerscher Fixpunktsatz, 17

- Čebyšev-Approximation, 19
- Čebyšev-Polynome, 40
- Čebyšev-Polynome zweiter Art, 41
- Čebyšev-Menge, 8

- Descent-Verfahren, 87
- Dichtheitsmaß, 95
- Dilatation, 126
- Dilatation, 126
- Dini-Lipschitz-Theorem, 84

- einbettbar, 36
- Eindeutigkeitsmenge, 8
- erster Satz von Haršiladze-Lozinski, 81
- Existenzmenge, 8
- Existenzsatz, 8
- Exponentialfunktion, 5

- Feinheit, 109
- Fourier-Projektor, 76

- Funktionsanpassung, 5

- Gitterweite, 109
- Glattheitsmodul, 111

- Haar-Funktion, 123
- Haar-Wavelets, 123
- Haarsche Bedingung, 28
- Haarscher Raum, 28
- HB Haarsche Bedingung, 28
- Hierarchie, 126
- Hilbert-Raum, 12
- Höldersche Ungleichung, 13
- HR Haarscher Raum, 28

- Interpolationsaufgabe, 28
- Invarianzprinzip, 40

- Jackson-Kern, 60
- Jensensche Ungleichung, 13

- kleinste Fehlerquadrate, 6
- konvexe Funktion, 11
- konvexe Hülle, 24
- konvexe Menge, 11

- Lagrange-Interpolation, 66
- Lemma von Baire, 80
- Lemma von Schur, 67
- Lethargietheorem, 55
- Lipschitz-Bedingung, 57
- Lipschitz-Klasse, 58
- Lösen einer parabolischen Differentialgleichung, 6

- Markov-Ungleichung, 70
- maximale Funktion, 121
- Menge der Extrempunkte, 19
- metrische Projektion, 15
- Minimalfolge, 10

- normale Menge, 30

- Orthonormalsystem, 102

- parabolische Differentialgleichung, 6
- Parallelogrammgleichung, 12
- Petree-Funktional, 75
- Polynomoperator, 76
- positiver linearer Operator, 51
- Prinzip der gleichmäßigen Beschränktheit, 81
- Projektion, 76
- Projektionsoperator, 76

- Quasi-Interpolationsoperator, 111

- Randbedingung, 6
- Referenz, 20
- Regularitätssatz, 6
- Remez-Algorithmus, 92

- Saturation, 54
- Saturationskern, 54
- Saturationsklasse, 54
- Satz von Bernstein, 71
- Satz von Bohmann-Korovkin, 51
- Satz von Carathéodory, 24
- Satz von de la Vallée-Poussin, 32
- Satz von Faber, 76
- Satz von Mairhuber-Curtis, 38
- Satz von Stiefel, 89
- Satz von Weierstraß, 54
- Satz von Zygmund, 73
- schnelle Wavelet-Transformation,
128
- semi-infinites Programm, 19
- separabel, 101
- shift invarianter Raum, 126
- Skala, 122, 126
- Skalarprodukt, 12
- Stetigkeitsmodul, 56, 111
- streng eindeutige beste
Approximation, 34
- strikt konvexer Raum, 11

- Taylor-Polynom, 5
- totale Variation, 115
- Translat, 126
- Translation, 126
- Translationsoperator, 111
- transzendente Funktion, 5

- uniform konvexer Raum, 11

- Verfeinerungsmaske, 126
- Vergleichssatz der Interpolation, 46
- Vergleichssatz von Bernstein, 47

- Wärmeleitungsgleichung, 6
- Wavelet-Darstellung, 127
- Wavelet-Raum, 127

- Zerlegung, 109
- zweite Stetigkeitsmodul, 58
- zweiter Satz von Haršiladze-Lozinski,
84

Literaturverzeichnis

- [1] Colin Bennett and Robert Sharpley, *Interpolation of operators*, Pure and Applied Mathematics, vol. 129, Academic Press Inc., Boston, MA, 1988.
- [2] E. W. Cheney, *Introduction to approximation theory*, McGraw-Hill Book Co., New York, 1966.
- [3] Ingrid Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. Pure Appl. Math. **41** (1988), no. 7, 909–996.
- [4] Ronald A. DeVore, *Nonlinear approximation*, Acta numerica, 1998, Acta Numer., vol. 7, Cambridge Univ. Press, Cambridge, 1998, pp. 51–150.
- [5] Ronald A. DeVore and George G. Lorentz, *Constructive approximation*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 303, Springer-Verlag, Berlin, 1993.
- [6] George G. Lorentz, *Approximation of functions*, second ed., Chelsea Publishing Co., New York, 1986.
- [7] Müller, Manfred W., *Approximationstheorie. Studienbuch für Studierende der Mathematik, Physik und Informatik ab 3. Semester.*, Studien-Texte Mathematik. Wiesbaden: Akademische Verlagsgesellschaft. 247 S., 1978.
- [8] von Querenburg, Boto, *Mengentheoretische Topologie. 2., neubearb. u. erw. Aufl.*, Hochschultext. Berlin, Heidelberg, New York: Springer-Verlag. IX, 208 S., 1979.