Skriptum Approximationstheorie

Prof. Dr. René Grothmann

Stand: 27. September 2006

Inhaltsverzeichnis

Kapitel 1

Existenz und Charakterisierung

1.1 Approximierbarkeit

Sei X eine kompakte Topologie (oder ein kompakter metrischer Raum), $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Sei

$$C_{\mathbb{K}}(X) = \{f : X \to \mathbb{K} : f \text{ stetig}\}.$$

versehen mit der **Supremums-Norm**

$$\|f\| = \|f\|_X = \|f\|_{X,\infty} := \max_{x \in X} |f(x)|.$$

Wir schreiben zur Abkürzung $C(X)$, wenn der Körper keine Rolle spielt oder fest liegt. Seien

$$V_1 \subseteq V_2 \subseteq V_3 \subseteq \ldots \subseteq C(X)$$

Teilmengen, oder, spezieller, Unterräume. f heißt **gleichmäßig approximierbar**, wenn es eine Folge von Funktionen

$$v_n \in V_n \qquad \text{für alle } n \in \mathbb{N}$$

gibt, so dass

$$\lim_{n \to \infty} \|v_n - f\| = 0$$

gilt.

1.1. Bemerkung: Äquivalent dazu ist, dass

$$V = \bigcup_{n \in \mathbb{N}} V_n$$

dicht in $C(X)$ ist, dass also $\overline{V} = C(X)$ gilt, wobei $\overline{V}$ den Abschluss von V im metrischen Raum $C(X)$ bedeutet.

1.2. Beispiel: $X = [a, b]$, $\mathbb{K} = \mathbb{R}$, $V_n = \mathcal{P}_n$, der Raum der algebraischen Polynome vom Grad kleiner oder gleich n. Die Frage ist, ob es zu jedem $\epsilon > 0$ ein und $f \in C[a, b]$ ein Polynom p gibt mit

$$\|f - p\| < \epsilon.$$

Diese Frage wird durch den folgenden Satz positiv beantwortet. Man beachte, dass man die Approximation nur dann mit beschränktem Grad n bewerkstelligen kann, wenn f selbst ein Polynom ist.

1.3 Satz: **(Weierstraß)** *Jede Funktion $f \in C[a, b]$ lässt sich beliebig genau gleichmäßig durch* ***algebraische Polynome*** *approximieren.*

Beweis: In der Tat konvergieren die Polynome

$$B_n f(t) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} t^k (1-t)^{n-k}$$

für $f \in C[0, 1]$ gleichmäßig gegen f auf $[0, 1]$. B_n heißt der **Bernstein-Operator**. Details in der Vorlesung. □

1.4. Bemerkung: Den **Raum der trigonometrischen Polynome** vom Grad n bezeichnen wir als $\mathcal{T}_n$. Jede Funktion in diesem Raum hat eine eindeutige Darstellung der Form

$$f(t) = \sum_{k=-n}^{n} a_k z^k, \qquad z = e^{it}, \quad a_{-k} = \overline{a_k} \text{ für alle } k.$$

Es folgt, dass das Produkt zweier Funktionen aus T_n und T_m in T_{m+n} liegt.

1.5. Bemerkung: Die Transformation $p \mapsto \hat{p}$ bildet

$$p(x) = \sum_{k=0}^{n} a_k T_k(x)$$

(T_k das k-te Chebyshev-Polyom $T_k(x) = \cos(n \arccos x)$) auf

$$\hat{p}(t) = \sum_{k=}^{n} a_k \cos kt$$

ab.

1.6 Satz: **(Weierstraß)** *Jede 2π-periodische stetige Funktion lässt sich beliebig genau gleichmäßig auf $\mathbb{R}$ durch* ***trigonometrische Polynome***

$$g(t) = a_0 + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt)$$

approximieren.

Beweis: In der Vorlesung. □

1.7 Satz: **(Stone)** *Sei $\mathcal{A} \subseteq C(X)$ eine* ***Algebra****. Das heißt, A ist ein Unterraum, der zu je zwei Funktionen das Produkt enthält. Weiter gelte, dass $\mathcal{A}$ jedes $f \in C(X)$ in zwei beliebigen Punkten aus X interpolieren kann. Dann ist $\mathcal{A}$ dicht in $C(X)$.*

Beweis: In der Vorlesung. □

1.8. Bemerkung: Die Interpolationsbedingung ist zum Beispiel erfüllt, wenn f die Funktion 1 enthält und zu je zwei Punkten eine Funktion, die dort verschiedene Werte hat. (Man sagt, $\mathcal{A}$ trenne die Punkte von X.)

1.9. Beispiel: Polynome in n Variablen sind dicht in $C(X)$ für $X \subseteq \mathbb{R}^n$.

1.10 Satz: *Sei $\mathcal{A} \subseteq C_{\mathbb{C}}(X)$ eine Algebra komplexwertiger Funktionen, in der sich jede Funktionen an zwei verschiedenen Punkten interpolieren lässt, und die mit f auch $\overline{f}$ enthält. Dann ist $\mathcal{A}$ dicht in $C_{\mathbb{C}}(X)$.*

Beweis: In der Vorlesung. $\square$

1.11. Bemerkung: Dieser Satz taugt im allgemeinen nicht zur Approximation mit komplexen Polynomen, da $\overline{p(z)}$ kein Polynom ist, wenn $p(z)$ ein Polynom ist. Der Satz von **Mergelyan** sagt aber aus, dass sich jede Funktion, die auf einem Kompaktum $K \subseteq \mathbb{C}$ stetig ist und in K° analytisch, gleichmäßig durch Polynome approximieren lässt.

1.12 Aufgabe: Folgern Sie aus dem Satz von Weierstraß, dass eine stetige Funktion $f : [a,b] \to \mathbb{R}$ mit

$$\int_a^b f(t)t^n = 0, \qquad \text{für alle } n = 0, 1, 2, \ldots$$

identisch 0 sein muss.

1.13 Aufgabe: Sei $f : [a,b] \to \mathbb{R}$ stetig und in (a,b) stetig differenzierbar. Sei außerdem f' auf $[a,b]$ stetig fortsetzbar. Zeigen Sie, dass es dann zu jedem $\epsilon > 0$ ein Polynom p gibt mit

$$\|f - p\| < \epsilon, \quad \|f' - p'\| < \epsilon.$$

1.14 Aufgabe: Sei X ein Kompaktum und $f : X \to \mathbb{R}$ stetig, sowie V der von

$$1, \quad f, \quad f^2, \quad \ldots$$

aufgespannte Unterraum. Geben Sie eine notwendige und hinreichende Bedingung an f an, dass V dicht in $C(X)$ ist.

1.15 Aufgabe: Zeigen Sie, dass die Menge aller geraden Polynome dicht in $C[a,b]$ ist, falls

$$0 \le a < b$$

gilt, nicht aber, falls $a < 0$ und $b > 0$ ist. Was gilt für ungerade Polynome?

1.16 Aufgabe: Man zeige, dass sich eine stetige Funktion $f : [0,1] \to \mathbb{R}$ genau dann durch eine Folge von Polynomen mit ganzzahligen Koeffizienten gleichmäßig approximieren lässt, wenn $f(0)$ und $f(1)$ ganzzahlig sind.

Hinweis: Für die eine Richtung betrachte man die modifizierten Bernsteinpolynome

$$\tilde{B}_n f(x) = \sum_{k=0}^{n} \left[\binom{n}{k} f(\frac{k}{n}) \right] x^k (1-x)^{n-k}.$$

Dabei bezeichne $[\ldots]$ die Gaußsche Klammer, also

$$[x] = \max\{n \in \mathbb{N} : n \le x\}.$$

Berechnen Sie dann $B_n f - \tilde{B}_n f$ und zeigen Sie, dass

$$\sum_{k=1}^{n-1} x^k (1-x)^{n-k}$$

gleichmäßig mit $n \to \infty$ gegen 0 konvergiert. Schätzen Sie diesen Ausdruck dazu für $x \in [0, 1/4]$ und $x \in [3/4, 1]$ mit der geometrischen Reihe ab.

1.2 Beste Approximationen

Für einen metrischen Raum M und $V \subseteq V$ kann man

$$d(f, V) = \inf_{v \in V} d(f, v)$$

definieren. Dies ist der **Minimalabstand** (oder **Approximationsfehler**) von f zu V. Eine Punkt v_0 in V mit

$$d(f, v_0) = d(f, V)$$

heißt **beste Approximation** an f bezüglich V.

Wir haben hier meist $M = C(X)$. Wenn die Metrik auf $C(X)$ zur Supremumsnorm gehört, so nennt man eine beste Approximation **gleichmäßige beste Approximation**. Wenn V ein Unterraum von $C(X)$ ist, so nennt man die Approximationsaufgabe **lineare Approximationsaufgabe**.

1.17 Satz: *Sei M ein normierter Raum und $V \subseteq M$ abgeschlossen und nicht leer, so dass V in einem endlich dimensionalen Unterraum von M enthalten ist. Dann existiert zu jedem $f \in M$ eine beste Approximation bezüglich V.*

Beweis: Man kann zeigen, dass endlich dimensionale Unterräume abgeschlossen sind und beschränkte abgeschlossene Teilmengen kompakt. Deswegen ist für fest gewähltes $v_0 \in M$ die Menge

$$\tilde{V} = \{v \in V : d(v, v_0) \leq d(f, v_0)\}$$

kompakt. Die stetige Funktion $v \mapsto d(f, v)$ nimmt auf $\tilde{V}$ ein Minimum an, das dann beste Approximation an f bezüglich V ist. □

1.18 Satz: *Sei M ein normierter Raum und $V \subseteq M$ konvex (zum Beispiel ein Unterraum). Dann ist die Menge der besten Approximationen*

$$E_f = \{v \in V : \|f - v\| = d(f, V)\}$$

konvex.

Beweis: In der Vorlesung. □

1.19 Satz: *Die Norm auf M sei strikt konvex, d.h. für alle $v_1, v_2 \in M$ mit $\|v_1\| = \|v_2\| = 1$ gelte*

$$\|\frac{1}{2}(v_1 + v_2)\| < 1.$$

Dann ist die beste Approximation bezüglich einer konvexen Menge eindeutig bestimmt.

Beweis: In der Vorlesung. □

1.20. Bemerkung: Eine lineare Abbildung $\phi : V \to W$ eines normierten Vektorraums in einen anderen ist genau dann stetig, wenn

$$\|\phi\| := \sup_{x \neq 0} \frac{\|\phi(x)\|}{\|x\|} < \infty$$

gilt. $\|\phi\|$ ist eine Norm auf dem Raum der stetigen Abbildungen von V nach W, die man **Funktionalnorm** nennt. Es gilt

$$\|\phi\| := \sup_{\|x\|=1} \|\phi(x)\|.$$

Außerdem für die Hintereinanderausführung zweier linearer Abbildungen

$$\|\phi \circ \psi\| \leq \|\phi\| \, \|\psi\|.$$

Folglich

$$\|\phi^{-1}\| \geq \frac{1}{\|\phi\|}$$

für eine invertierbare Abbildung $\phi : V \to V$.

1.21 Aufgabe: Sei V der Raum der Polynome vom Grad 1 auf $[0,1]$ versehen mit der L_1-Norm und $\phi : V \to \mathbb{R}$ definiert durch $\phi(f) = f(0)$. Berechnen Sie die zugehörige Norm $\|\phi\|$.

1.22 Satz: *Sei $V \subseteq M$ ein Unterraum und $\phi : M \to \mathbb{K}$ ein lineares stetiges Funktional ungleich dem Nullfunktional mit*

$$\phi(M) = \{0\}.$$

Dann gilt für alle $f \in M$

$$d(f, V) \geq \frac{|L(f)|}{\|L\|}.$$

Beweis: Der Einfachheit halber beweisen wir den Satz nur für endlich dimensionale V und überlassen den allgemeinen Fall dem Leser. Dann existiert eine beste Approximation $v \in V$ an f, und es gilt

$$|L(f)| = |L(f) - L(v)| \leq \|L\| \, \|f - v\| = \|L\| d(f, V).$$

Es folgt die Behauptung. □

1.23. Beispiel: Sei $p_1, p_2, \ldots$ die Folge der orthogonalen Polynome auf $[a, b]$ zu einer Gewichtsfunktion $w > 0$, also

$$\int_a^b p_i(t) p_j(t) w(t) \, dt = 0$$

für $i \neq j$. Dann ist

$$L(f) = \int_a^b f(t)p_{n+1}(t)w(t)\,dt$$

ein geeignetes Funktional

$$L = V \to \mathbb{R}$$

für $M = \mathcal{P}_n$ und $V = C[a,b]$. Es gilt

$$\|L\| = \int_a^b |p_{n+1}(t)|w(t)\,dt.$$

wobei V mit der Supremumsnorm versehen sei.

1.24. Beispiel: Seien

$$a \leq x_0 < \ldots < x_n \leq b$$

und

$$L(f) = [x_0, \ldots, x_n]_f.$$

Dann ist L ein geeignetes Funktional für den Raum $M = \mathcal{P}_n$ und $V = C[a,b]$. Es gilt

$$L(f) = \sum_{k=0}^{n} \frac{1}{\omega'(x_k)} f(x_k)$$

mit

$$\omega(x) = (x - x_0) \cdot \ldots \cdot (x - x_n).$$

Für jedes Funktional der Form

$$L(f) = \sum_{k=0}^{n} a_k f(x_k)$$

gilt

$$\|L\| = \sum_{k=0}^{n} |a_k|,$$

wenn V mit der Supremumsnorm versehen ist.

1.25 Aufgabe: Berechnen Sie $\|L\|$ mit dem Operator des obigen Beispiels, für den Fall $n = 2$ und

$$a = x_0 < x_1 = \frac{a+b}{2} < x_2 = b.$$

Folgern Sie für die Minimalabweichung $e_1(f)$ einer Funktion $f \in C[a,b]$ bezüglich Π_1

$$e_1(f) \geq \frac{|f(x_0) - 2f(x_1) + f(x_2)|}{4}.$$

1.26 Satz: *Sei $V \subseteq M$ ein Unterraum und $\Phi : M \to M$ ein linearer stetiger Operator mit*

$$\Phi(v) = v \qquad \text{für alle } v \in V.$$

Dann gilt

$$\|f - \Phi(f)\| \leq (\|L\| + 1)\, d(f, V).$$

Beweis: Sei $v \in V$ mit

$$\|f - v\| \leq d(f, V) + \epsilon.$$

Dann gilt

$$\begin{aligned} \|f - \Phi f\| \leq \|f - v\| + \|v - \Phi f\| \leq d(f, V) + \epsilon + \|\Phi\| \|v - f\| \\ = (1 + \|\Phi\|) d(f, V) + \epsilon. \end{aligned}$$

$\square$

1.27. Beispiel: Sei $M = C[a, b]$, $V = \mathcal{P}_n$ und

$$a \leq x_0 < \ldots < x_n \leq b.$$

Dann sei $P_n : C[a, b] \to \mathcal{P}_n$ der Operator, der f auf das Interpolationspolynom $P_n f$ in diesen Punkten abbildet. Es gilt

$$P_n f(x) = \sum_{k=0}^{n} \frac{\omega(x)}{(x - x_k)\omega'(x_k)} f(x_k).$$

Dann ist P_n solch ein Projektionsoperator. Für jeden Operator der Form

$$\Phi f(x) = \sum_{k=0}^{n} f(x_k) v_k(x)$$

mit $v_0, \ldots, v_n \in V$ und paarweise verschiedenen $x_0, \ldots, x_k$ gilt

$$\|\Phi\| = \left\| \sum_{k=0}^{n} |v_k| \right\|,$$

wenn $M = C(X)$ mit der Supremumsnorm versehen wird. Um $=$ zu zeigen, wählt man eine Funktion $f \in C(X)$ mit

$$f(x_k) = \text{sign } v_k(x), \qquad k = 0, \ldots, n.$$

Dies ist in einem kompakten metrischen Raum X immer möglich. Im Fall des Interpolationsoperators nennt man die Norm die **Interpolationsnorm**.

1.3 Trennungssätze

Zum Beweis des Satzes von Kolmogoroff zeigen wir einige Hilfssätze über konvexe Mengen, sowie andere in diesem Zusammenhang interessante Sätze. Die konvexe Hülle einer Menge $K \subseteq V$ ist

$$\text{co}\, K = \{\sum_{k=1}^{n} \lambda_k v_k : 0 \leq \lambda_k \text{ für alle } k,\ \textstyle\sum_{k=1}^{n} \lambda_k = 1,\ n \in \mathbb{N}\}.$$

Dies ist die kleinste konvexe Menge, die K umfasst.

1.28 Satz: **(Carathéodory)** *Sei* $\dim V = n$, $K \subseteq V$. *Im Fall* $\mathbb{K} = \mathbb{R}$ *gilt dann*

$$\operatorname{co} K = \{\sum_{k=1}^{n+1} \lambda_k v_k : 0 \leq \lambda_k \text{ für alle } k,\ \textstyle\sum_{k=1}^{n+1} \lambda_k = 1\},$$

und im Fall $\mathbb{K} = \mathbb{C}$ *gilt*

$$\operatorname{co} K = \{\sum_{k=1}^{2n+1} \lambda_k v_k : 0 \leq \lambda_k \text{ für alle } k,\ \textstyle\sum_{k=1}^{2n+1} \lambda_k = 1\}.$$

Beweis: Siehe Skript über linerare Algebra. □

1.29 Satz: *Sei* $K \subseteq V$ *kompakt,* V *endlich dimensional. Dann ist* $\operatorname{co} K$ *kompakt.*

Beweis: In der Vorlesung. □

1.30 Satz: *Sei* V *ein Skalarproduktraum, versehen mit der zugehörigen Norm, und* $M \subseteq V$ *konvex. Dann ist* $v \in M$ *genau dann beste Approximation an* $f \in V$, *wenn*

$$\operatorname{Re} \langle f - v, w - v \rangle \leq 0$$

für alle $w \in M$ *gilt.*

Beweis: In der Vorlesung. □

Ein **Hibertraum** ist ein Skalarproduktraum, der **vollständig** ist, in dem also jede Cauchy-Folge konvergiert. Wir wissen bereits, dass die beste Approximation an konvexe Mengen in Skalarprodukträumen eindeutig ist, wenn sie existiert. Für endlich dimensionale Unterräume wissen wir, dass sie existiert. Dies gilt aber auch alle für Hilberträume und abgeschlossene, konvexe Teilmengen.

1.31 Satz: *Sei* $M \subseteq H$ *abgeschlossen und konvex,* H *ein Hilbertraum. Dann existiert für jedes* $f \in H$ *eine eindeutig bestimmte beste Approximation* $v \in H$.

Beweis: In der Vorlesung. □

Wir benötigen den folgenden **Trennungssatz**. Dabei trennen wir zwei Mengen M_1 und M_2 durch ein reellwertiges Funktional ϕ

$$\phi(M_1) \leq \phi(M_2).$$

Von starker Trennung spricht man, wenn

$$\phi(M_1) \leq a < b \leq \phi(M_2)$$

ist. Im folgenden Satz ist $M_1 = \{0\}$ und $M_2 = M$ abgeschlossen und konvex und es wird stark getrennt, also

$$\phi(M) \geq \epsilon > 0 = \phi(0).$$

1.32 Satz: *Sei $M \subseteq H$ abgeschlossen und konvex, H ein Hilbertraum. Dann existiert genau dann ein $v \in M$ und ein $\epsilon > 0$ mit*

$$\phi(w) = \operatorname{Re}\langle w, v\rangle \geq \epsilon > 0 \qquad \text{für alle } w \in M,$$

wenn $0 \notin M$ ist.

Beweis: In der Vorlesung. □

1.33 Satz: *Sei $M \subseteq V$, V ein endlich-dimensionaler Skalarproduktraum. Dann existiert genau dann ein $v \in V$, $v \neq 0$, mit*

$$\operatorname{Re}\langle w, v\rangle \geq 0 \qquad \text{für alle } w \in M,$$

wenn

$$0 \notin (\operatorname{co} M)^{\circ}$$

ist.

Beweis: In der Vorlesung. □

1.34. Bemerkung: Um zwei Mengen M_1 und M_2 zu trennen, kann man einfach 0 von $M_2 - M_1$ trennen. Man beachte, dass $M_2 - M_1$ konvex ist, wenn die beiden Mengen konvex sind. Es ist auch kompakt, wenn die beiden Mengen kompakt sind. Es braucht aber nicht abgeschlossen zu sein, auch wenn es die beiden Mengen sind.

1.4 Charakterisierungssätze

In diesem Abschnitt ist V ein Unterraum des Raums $C(X)$ für ein Kompaktum X. Der Körper $\mathbb{K}$ ist entweder $\mathbb{R}$ oder $\mathbb{C}$. Die Norm ist die Supremumsnorm. Wir betrachten also die gleichmäßige, lineare Approximationsaufgabe.

1.35 Satz: **(Kolmogoroff)** *Die Funktion $v \in V$ ist genau dann beste Approximation an $f \in C(X)$, wenn es keine Funktion $w \in V$ gibt mit*

$$\operatorname{Re}\left(w(x)\overline{f(x) - v(x)}\right) > 0$$

für alle

$$x \in E(f - v) := \{x \in X : |f(x) - v(x)| = \|f - v\|\}.$$

Beweis: In der Vorlesung. □

1.36. Bemerkung: Für $\mathbb{K} = \mathbb{R}$ lautet die Bedingung, dass es kein $w \in V$ gibt mit

$$\operatorname{sign} w(x) = \operatorname{sign}(f(x) - v(x)) \qquad \text{für alle } x \in E(f - v).$$

1.37. Bemerkung: $E(f - v)$ heißt die **Extremalpunktmenge** von $f - v$. Diese Menge ist nach dem obigen Satz **charakteristische Menge** für die beste

Approximation. Das heißt, man kann genau dann f besser approximieren auf der gesamten Menge X, wenn man es auf $E(f-v)$ kann.

Man folgert aus diesem Satz leicht den **Alternantensatz** für die beste Approximation in $C[a,b]$ mit Polynomen $V = \mathcal{P}_n$.

1.38 Satz: *$p \in \mathcal{P}_n$ ist genau dann beste Approximation an f auf einem Kompaktum $K \subseteq \mathbb{R}$, wenn es Extremalpunkte*

$$a \leq x_0 < \ldots < x_{n+1} \leq b.$$

gibt mit

$$f(x_k) - p(x_k) = \sigma(-1)^k \|f-p\| \qquad \textit{für alle } k = 0, \ldots, n+1$$

mit festem $\sigma = \pm 1$.

Beweis: Wir wählen Mengen von Extremalpunkten der Fehlerfunktion $f - p$

$$E_1 < \ldots < E_m,$$

mit

$$f(x) - p(x) = \sigma(-1)^i \|f-p\|_K \qquad \text{für alle } x \in E_i$$

mit einem festen $\sigma = \pm 1$. Falls nun $m < n+2$ wäre, so kann man sehr leicht ein Polynom finden, dass auf den $m \leq n+1$ Komponenten das Vorzeichen $\sigma(-1)^i$ hat. Dazu muss man lediglich m einfache Nullstellen zwischen diese Mengen legen, was wegen $m \leq n$ möglich ist. $\square$

Diesen Satz werden wir für Haarsche Systeme später allgemein formulieren.

1.39 Aufgabe: Zeigen Sie für die beste Approximation p an die Funktion $f(x) = x^{n+1}$ auf $[-1,1]$ bezüglich $\mathcal{P}_n$

$$f(x) - p(x) = \frac{1}{2^n} T_{n+1}(x).$$

Folgern Sie für jedes Polynom

$$\omega(x) = x^{n+1} + a_n x^n + \ldots + a_0$$

dass

$$\|\omega\| \geq \frac{1}{2^n}$$

gilt.

1.40 Aufgabe: Aus dem Alternantensatz folgt, dass die beste Approximation p an $f \in C[-1,1]$ die Funktion in $n+1$ Punkten interpoliert. Folgern Sie aus der Darstellung des Interpolationsfehlers und der vorigen Aufgabe

$$\frac{1}{2^n(n+1)!} \min_{\xi \in [-1,1]} |f^{(n+1)}(\xi)| \leq d(f, \mathcal{P}_n) \leq \frac{1}{2^n(n+1)!} \max_{\xi \in [-1,1]} |f^{(n+1)}(\xi)|,$$

wenn f in $(-1,1)$ mindestens $n+1$-mal differenzierbar ist.

1.41 Aufgabe: Sei $p \in \Pi_n$ die gleichmäßig beste Approximation an f auf $[a,b]$, und es gelte

$$f^{(n+1)}(x) \neq 0 \qquad \text{für alle } x \in (a,b).$$

Zeigen Sie, dass $f - p$ dann eine eindeutig bestimmte Alternante mit Punkten

$$a = x_0 < \ldots < x_{n+1} = b$$

hat.

1.42 Satz: *Sei V ein n-dimensionaler Unterraum von $C(X)$. Dann sind äquivalent*

(1) *$v \in V$ ist beste Approximation an $f \in C(X)$.*

(2)

$$0 \in \operatorname{co}\{\overline{f(x) - v(x)}\hat{x} : x \in E(f - v)\}$$

*wobei $\operatorname{co} K$ die **konvexe Hülle** einer Menge K bezeichne und*

$$\hat{x} = \begin{pmatrix} v_1(x) \\ \vdots \\ v_n(x) \end{pmatrix}$$

für eine fest gewählte Basis $v_1, \ldots, v_n$ von V sei.

(3) *Es gibt **charakteristische Punkte** $x_1, \ldots, x_k \in E(f - v)$, wobei $k \leq n+1$ im Fall $\mathbb{K} = \mathbb{R}$ und $k \leq 2n+1$ im Fall $\mathbb{K} = \mathbb{C}$ gewählt werden kann, also Punkte, so dass v auch beste Approximation an f auf diesen Punkten ist.*

Beweis: In der Vorlesung. □

1.43. Beispiel: Die $n + 2$ Alternantenpunkte im Alternantensatz sind charakteristisch. In diesem Fall gilt

$$\dim V = \dim \mathcal{P}_n = n + 1.$$

In der Tat kann es nicht weniger charakteristische Punkte geben, wenn f nicht selbst ein Polynom ist.

1.44. Bemerkung: Dass es $k \leq n + 1$ charakteristische Punkte gibt, bedeutet nicht, dass jede beste Approximation auf diesen Punkten insgesamt beste Approximation ist. Es bedeutet lediglich, dass man auf diesen Punkten auch nicht besser approximieren kann als auf der gesamten Menge. Wenn allerdings die beste Approximation auf diesen Punkten eindeutig ist, so ist sie die insgesamt beste Approximation. Wir werden zeigen, dass dies für Haarsche Unterräume der Fall ist.

1.5 Haarsche Unterräume

1.45. Definition: Sei X ein beliebiger metrischer Raum. Ein n-dimensionaler Unterraum $V \subseteq C(X)$ (oder $V \subseteq C_{\mathbb{C}}(X)$) heißt **Haarscher Unterraum**, wenn jedes $v \in V \setminus \{0\}$ höchstens $n - 1$ verschiedene Nullstellen hat.

Man nennt Haarsche Unterräume auch **Chebyshev-Räume**.

1.46 Satz: *Sei $V \subseteq C(X)$ ein reeller oder komplexer Unterraum der Dimension n. Dann sind äquivalent.*

(1) *Jedes $v \in V \setminus \{0\}$ hat höchstens $n-1$ verschiedene Nullstellen.*

(2) *Für eine fest gewählte Basis $v_1, \ldots, v_n$ von V und für alle paarweise verschiedenen Punkte $x_1, \ldots, x_n \in X$ ist*

$$\det \begin{pmatrix} v_1(x_1) & \ldots & v_n(x_1) \\ \vdots & & \vdots \\ v_1(x_n) & \ldots & v_n(x_n) \end{pmatrix} \neq 0$$

(3) *Für alle paarweise verschiedenen Punkte $x_1, \ldots, x_n \in X$ und für alle vorgegebenen Werte $y_1, \ldots, y_n \in \mathbb{K}$ existiert ein $v \in V$ mit*

$$v(x_i) = y_i \qquad \text{für alle } i = 1, \ldots, n.$$

*(**Interpolation**).*

Beweis: Man beachte, dass X mindestens n Punkte enthalten muss, weil

$$\dim V \leq \dim C(X) = |X|$$

für endliche Mengen X gilt. Der Beweis verwendet die Abbildung

$$\phi : V \to \mathbb{K}^n$$
$$\phi(v) = (v(x_1), \ldots, v(x_n))$$

oder das Gleichungssystem

$$\begin{pmatrix} v_1(x_1) & \ldots & v_n(x_1) \\ \vdots & & \vdots \\ v_1(x_n) & \ldots & v_n(x_n) \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Aus Prinzipien der linearen Algebra folgt die Äquivalenz der Behauptungen. □

1.47. Bemerkung: Wenn V ein Haarscher Unterraum der Dimension n ist, so gilt (2) für alle Basen.

1.48. Bemerkung: Wenn die Bedingung (1) des obigen Satzes gilt, so folgt $\dim V \leq n$. Im Umkehrschluss muss jeder Unterraum mit $\dim V > n$ eine Funktion $v \neq 0$ mit n verschiedenen Nullstellen enthalten. Denn entweder enthält X weniger als n Punkte, oder der Kern der Abbildung ϕ ist für paarweise verschiedene Punkte $x_1, \ldots, x_n \in X$ der Nullraum, so dass $\phi : V \to \mathbb{K}^n$ also injektiv sein muss.

1.49. Bemerkung: Wenn die Bedingung (3) des obigen Satzes gilt und X mindestens n Punkte enthält, so muss $\dim V \geq n$ sein. Denn in diesem Fall ist die Abbildung $\phi : V \to \mathbb{K}^n$, die zu n paarweise verschiedenen Punkten aus X gehört, surjektiv.

1.50. Bemerkung: Wenn V ein Haarscher Unterraum der Dimension n ist, so ist die Interpolation an n paarweise verschiedenen Punkten eindeutig bestimmt.

1.51 Aufgabe: Sei V ein Haarscher Unterraum der Dimension n von $C(X)$ und $Y \subseteq X$. Bezeichne dann $V|_Y$ den Raum der auf Y eingeschränkten Funktionen von

V. Zeigen Sie, dass $V|_Y$ genau dann ein Haarscher Unterraum der Dimension n von $C(Y)$ ist, wenn Y mindestens n Punkte enthält.

1.52. Beispiel: Der Polynomraum $\mathcal{P}_n$ der reellen algebraischen Polynom vom Grad kleiner oder gleich n ist ein Haarscher Unterraum von $C(\mathbb{R})$ der Dimension $n+1$. Denn einerseits besitzt der die Basis $1, x, \ldots, x^n$ und andererseits ist die Bedingung (3) erfüllt, weil man das Interpolationspolynom in der Lagrangeform

$$p(x) = \sum_{k=0}^{n} \frac{\omega(x)}{(x - x_k)\omega'(x_k)}$$

aufschreiben kann, wobei

$$\omega(x) = (x - x_0) \cdot \ldots \cdot (x - x_n)$$

ist. Aus der obigen Bemerkung folgt also $\dim \mathcal{P}_n = n + 1$. Dasselbe gilt für die komplexen algebraischen Polynome auf $\mathbb{C}$.

1.53. Beispiel: Der Raum der trigonometrischen Polynome $\mathcal{T}_n$ vom Grad kleiner oder gleich n ist ein reeller Haarscher Unterraum von $C[0, 2\pi)$ der Dimension $2n + 1$.

1.54. Bemerkung: Der von den Funktionen $v_1, \ldots, v_n$ aufgespannte Unterraum von $C(X)$ ist genau dann ein Haarscher Unterraum der Dimension n, wenn für alle $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$, die nicht alle gleich 0 sind, die Funktion

$$v(x) = \alpha_1 v_1 + \ldots + \alpha_n v_n$$

höchstens $n - 1$ verschiedene Nullstellen haben kann, und wenn X mindestens n Punkte enthält.

1.55. Beispiel: Sei $\lambda_1 < \ldots < \lambda_n$. Der von den Funktionen

$$e^{\lambda_1 x}, \ldots, e^{\lambda_n x}$$

aufgespannte Raum ist ein reeller Haarscher Unterraum von $C(\mathbb{R})$ der Dimension n. In diesem Fall verwendet man Induktion mit dem Satz von Rolle und die obige Bemerkung.

1.56 Aufgabe: Folgern Sie aus dem obigen Beispiel, dass der von den Funktionen

$$x^{\lambda_1}, \ldots, x^{\lambda_n}$$

aufgespannte Raum ein Haarscher Unterraum von $C(0, \infty)$ ist.

1.57 Aufgabe: Zeigen Sie, dass der von den Funktionen

$$1, x, \ldots, x^n, \sin(x), \cos(x)$$

aufgespannte Raum ein reeller Haarscher Unterraum von $C[0, 2\pi)$ ist.

1.58 Satz: *Sei $V \subseteq C[a, b]$ ein reeller Haarscher Unterraum der Dimension n.*

(1) *Dann hat die Determinante aus Bedingung (2) des vorigen Satzes für je n Punkte*

$$a \leq x_1 < \ldots < x_n \leq b.$$

dasselbe Vorzeichen.

(2) *Wenn dann für $v \in V$ und $n+1$ Punkte*

$$a \leq t_1 < \ldots < t_{n+1} \leq b.$$

gilt

$$v(t_1) \leq 0, \quad v(t_2) \geq 0, \ldots,$$

so folgt $v = 0$.

Beweis: Der Beweis von (2) verwendet das Gleichungssystem

$$\begin{pmatrix} v_1(t_1) & \ldots & v_n(t_1) & v(t_1) \\ \vdots & & \vdots & \vdots \\ v_1(t_{n+1}) & \ldots & v_n(t_{n+1}) & v(t_{n+1}) \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ -1 \end{pmatrix} = 0,$$

wobei $v_1, \ldots, v_n$ eine Basis von V ist. Durch Entwickeln der Determinante nach der letzten Spalte sieht man mit Hilfe von (1), dass dieses Gleichungssystem eine Determinante ungleich 0 hat. Es folgt $v = 0$. □

1.59. Bemerkung: Es kann keinen reellen Haarschen Unterraum der Dimension n auf einem Raum X geben, in dem man zwei der Punkte $x_1, \ldots, x_n$ stetig vertauschen kann, ohne die anderen Punkte zu treffen. Denn sonst müsste während des stetigen Vertauschens die Determinante aus Bedingung (2) ihr Vorzeichen ändern. Ebenso gibt es keinen Haarschen Unterraum gerader Dimension auf dem Rand des Einheitskreises.

1.60 Satz: **(de la Vallée Poussin)** *Sei $V \subseteq C[a,b]$ ein reeller Haarscher Unterraum der Dimension n, $f \in C(X)$, $v \in V$ und*

$$a \leq t_1 < \ldots < t_{n+1} \leq b$$

Punkte, so dass ein festes Vorzeichen $\sigma = \pm 1$ existiert mit

$$f(t_k) - v(t_k) = (-1)^k \sigma h_k \qquad \text{für alle } k = 0, \ldots, n.$$

Dann gilt

$$d(f, V) \geq \min_{k=0,\ldots,n} h_k.$$

Beweis: In der Vorlesung. □

1.61. Bemerkung: Die Extremalpunktmenge der reellen oder komplexen besten Approximation bezüglich einem Haarschen Unterraum enthält mindestens $n+1$ Punkte. Denn man könnte sonst interpolieren.

1.62 Satz: *Sei X ein Kompaktum und $V \subseteq C(X)$ (bzw. $V \subseteq C_{\mathbb{C}}(X)$ ein Haarscher Unterraum. Dann ist die beste Approximation an $f \in C(X)$ bezüglich V eindeutig bestimmt.*

Beweis: Wenn v_1, v_2 beste Approximationen sind, so betrachtet man

$$v = \frac{1}{2}(v_1 + v_2).$$

Dies ist ebenfalls beste Approximation und es gibt $n+1$ charakteristische Punkte in $E(f-v)$. Man zeigt, dass v_1 und v_2 in diesen Punkten übereinstimmen. $\square$

1.63 Satz: (**Alternantensatz von Chebyshev**) *Sei V ein reeller Haarscher Unterraum der Dimension n von $C[a,b]$. Dann ist $v \in V$ genau dann beste Approximation an $f \in C[a,b]$, wenn es $(n+1)$ Punkte*

$$a \le t_1 < \ldots < t_{n+1} \le b$$

gibt und ein festes Vorzeichen $\sigma = \pm 1$, so dass

$$f(t_k) - v(t_k) = (-1)^k \sigma \|f-v\|_{[a,b]} \qquad \text{für alle } k = 0, \ldots, n$$

gilt.

Beweis: Aufgrund des Satzes von de la Vallée Poussin ist $v \in V$ beste Approximation an $f \in C[a,b]$, wenn $f - v$ eine Alternante der Länge $n+1$ hat.

Für die umgekehrte Richtung verwendet man $n+1$ charakteristische Punkte $t_0, \ldots, t_{n+1}$ nach dem Charakterisierungssatz. Das Gleichungssystem

$$\begin{pmatrix} v_1(t_1) & \ldots & v_n(t_1) & (-1)^1 \\ \vdots & & \vdots & \vdots \\ v_1(t_{n+1}) & \ldots & v_n(t_{n+1}) & (-1)^{n+1} \end{pmatrix} \cdot \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ h \end{pmatrix} = \begin{pmatrix} f(t_1) \\ \vdots \\ f(t_{n+1}) \end{pmatrix}$$

ist eindeutig lösbar. Auf diese Weise erhält man eine Funktion $\tilde{v}$, die auf den Punkten alterniert, dort also beste Approximation ist. Wegen der Eindeutigkeit der besten Approximation folgt $v = \tilde{v}$ auf $t_1, \ldots, t_{n+1}$, also auf $[a,b]$. $\square$

1.64 Aufgabe: Folgern Sie aus dem Alternantensatz, dass jeder reelle Haarsche Unterraum von $C[a,b]$ eine Funktion enthält, die nirgends 0 ist.

1.65. Bemerkung: Der Alternantensatz und die Abschätzung von de la Valée Poussin lässt sich auf allgemeine Kompakta X übertragen, indem man $(-1)^k$ durch

$$(-1)^k \det \begin{pmatrix} v_1(x_0) & \ldots & v_1(x_0) \\ \vdots & & \vdots \\ v_1(x_{k-1}) & \ldots & v_1(x_{k-1}) \\ v_1(x_{k+1}) & \ldots & v_1(x_{k+1}) \\ \vdots & & \vdots \\ v_1(x_n) & \ldots & v_1(x_n) \end{pmatrix}$$

ersetzt.

1.66. Bemerkung: Wenn X der Rand des Einheitskreises ist, und V ein Haarscher Unterraum, so hat V ungerade Dimension $2n+1$. Die beste Approximation ist in diesem Fall durch eine Alternante der Länge $2n+2$ charakterisiert.

1.6 Algorithmen

1.67. Bemerkung: Man kann das Approximationsproblem für $V \subseteq C(X)$ mit Basis $v_1, \ldots, v_n$ in ein **semi-infinites** Optimierungsproblem umschreiben. Dabei wird h minimiert unter den unendlich vielen Nebenbedingungen

$$-h \leq \sum_{k=1}^{n} \alpha_k v_k(x) - f(x) \leq h, \qquad \text{für alle } x \in X$$

Wenn $X = \{x_1, \ldots, x_m\}$ eine endliche Menge ist, so entsteht ein gewöhnliches Optimierungsproblem. Man minimiert h unter den Nebenbedingungen

$$\begin{aligned} \alpha_1 v_1(x_i) + \ldots + \alpha_n v_n(x_i) - h &\leq f(x_i), \\ -\alpha_1 v_1(x_i) - \ldots - \alpha_n v_n(x_i) - h &\leq -f(x_i), \end{aligned}$$

für $i = 1, \ldots, m$. Man könnte nun versuchen X durch endlich viele Punkte zu überdecken. Das entstehende Optimierungsproblem ist allerdings nicht gut konditioniert, wenn die Punkte gegeneinander rücken. Es ist günstiger, das duale Problem zu lösen.

1.68. Bemerkung: Das Optimierungsproblem ist eine Anwendung des allgemeinen **Ausgleichsproblems**

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|_\infty$$

für ein gegebene Matrix $A \in \mathbb{R}^{m \times n}$ und $b \in \mathbb{R}^m$. Gewöhnlich wird allerdings hier die Euklidsche Norm verwendet, also

$$\min_{x \in \mathbb{R}^n} \|Ax - b\|.$$

Die Lösung ist dann durch die Normalgleichung $A^T Ax = A^T b$ zu berechnen.

1.69. Definition: Der **Remez-Algorithmus** berechnet die beste Approximation an die Funktion $f \in C[a, b]$ aus einem Haarschen Unterraum $V \subseteq C[a, b]$ der Dimension n. In jedem Schritt werden Punkte

$$a \leq x_1 < \ldots < x_{n+1} \leq b$$

verwendet, die man **Referenz** nennt. Man berechnet die beste Approximation $v \in V$ an f auf der Referenz, die durch eine Alternante gekennzeichnet ist.

$$f(x_i) - v(x_i) = (-1)^i h \qquad \text{für alle } i = 1, \ldots, n + 1.$$

Man nennt h die **Referenzabweichung**. Sei nun $\tilde{x}$ ein Punkt, in dem $|f - v|$ maximal ist. Dann tauscht man $\tilde{x}$ gegen einen Punkt x_{i_0} aus, so dass eine neue Referenz

$$a \leq \tilde{x}_1 < \ldots < \tilde{x}_{n+1} \leq b$$

entsteht, und so dass $f - v$ in dieser Referenz im Vorzeichen alterniert. Dies ist immer möglich. Auf der neuen Referenz berechnet man die beste Approximation $\tilde{v}$ mit Referenzabweichung $\tilde{h}$, und es folgt aus der Abschätzung von de la Vallée Poussin

$$|h| \leq |\tilde{h}|.$$

Deswegen heißt der Remez-Algorithmus ein **aufsteigendes Verfahren**. Der Schritt von h nach $\tilde{h}$ wird nun immer wieder wiederholt. Es gilt in jedem Schritt

$$|h| \leq d(f, V) \leq \|f - v\|.$$

Das Verfahren wird solange fortgesetzt, bis

$$\|f - v\| - |h| < \epsilon$$

für vorgegebenes $\epsilon > 0$ wird.

1.70 Satz: *Die besten Approximationen auf den Referenzabweichungen im oben beschriebenen Remez-Algorithmus konvergieren gegen die beste Approximation auf* $[a, b]$.

Beweis: In der Vorlesung. □

1.71. Bemerkung: Man kann den Algorithmus verbessern, indem man zusätzlich lokale Extrema austauscht. Dieses Verfahren nennt man **Simultan-Austausch**. Mit geeigneten Voraussetzungen an f kann man zeigen, dass das Verfahren lokal quadratisch konvergiert.

1.7 Schwach-Chebyshev-Räume

1.72. Definition: Sei $X \subseteq \mathbb{R}$. Ein Unterraum $V \subseteq C(X)$ heißt **Schwach-Chebychev-Raum** der Dimension n, wenn jede Funktion $v \in V$ höchstens n echte **Vorzeichenwechsel** hat. Das heißt, es gibt keine Funktion $v \in V$ und Punkte

$$x_1 < \ldots < x_{n+1}$$

in X, so dass

$$\operatorname{sign} v(x_{i+1}) = -\operatorname{sign} v(x_i) \qquad \text{für alle } i = 1, \ldots, n$$

gilt (**alternierende Punktfolge** der Länge $n + 1$).

1.73. Bemerkung: Offenbar ist jeder Haarsche Unterraum von $C[a, b]$ ein Schwach-Chebyshev-Raum. Denn im Fall $X = C[a, b]$ bedingen n echte Vorzeichenwechsel auch n verschiedene Nullstellen.

1.74. Definition: Man definiert den Splineraum

$$S_m(x_0, \ldots, x_{k+1}) \subseteq C[x_0, x_{k+1}]$$

$m \in \mathbb{N}$, zu den Knoten

$$x_0 < \ldots x_{k+1}$$

dadurch, dass $s \in S_m(x_0, \ldots, x_{k+1})$ überall $m - 1$-mal stetig differenzierbar ist, bzw. stetig für $m = 1$, und in jedem der Teilintervalle

$$(x_i, x_{i+1}), \qquad i = 0, \ldots, k,$$

ein Polynom m-ten Grades ist.

1.75. Beispiel: Wichtige Spezialfälle sind die stetigen Streckenzüge für $m = 1$, die stetig differenzierbaren quadratischen Splines für $m = 2$, und die zweimal stetig differenzierbaren kubischen Splines für $m = 3$.

1.76. Bemerkung: Splineräume mit mehr als 2 Knoten können keine Haarsche Unterräume sein, denn sie können Nullstellenintervalle haben, wie die Splines

$$(x - x_k)_+^m = \begin{cases} (x - x_k)^m, & x > x_k, \\ 0, & x \leq x_k, \end{cases}$$

zeigen.

1.77 Satz: *Die Splineräume $S_m(x_0, \ldots, x_{k+1}$ sind Schwach-Chebyshev-Räume der Dimension $m + k + 1$.*

Beweis: Die Funktionen

$$1, x, \ldots, x^m, (x - x_1)^m, \ldots, (x - x_k)^m$$

bilden eine Basis von $S_m(x_0, \ldots, x_{k+1})$. Man zeigt den Sonderfall $m = 1$ separat, und den Rest der Behauptung mit dem Zwischenwertsatz durch Induktion nach m. □

1.78 Satz: *Sei $V \subseteq C[a, b]$ der Dimension n ein Schwach-Chebyshev-Raum. Dann gibt es für jede fest gewählte Basis $v_1, \ldots, v_n$ von V ein $\sigma = \pm 1$ gibt mit*

$$\sigma \det \begin{pmatrix} v_1(x_1) & \ldots & v_n(x_1) \\ \vdots & & \vdots \\ v_1(x_n) & \ldots & v_n(x_n) \end{pmatrix} \geq 0$$

für alle

$$a \leq x_1 < \ldots < x_n \leq b.$$

Beweis: Man zeigt den Satz zunächst für n Punkte aus $n + 1$ vorgegebenen Punkten. Danach zeigt, dass das Vorzeichen der Determinante für zwei Punktfolgen von n Punkten nicht verschieden sein kann. Dazu verwendet man Induktion nach der Anzahl der übereinstimmenden Punkte. □

1.79 Satz: *Sei $V \subseteq C[a, b]$ ein Unterraum mit einer Basis $v_1, \ldots, v_n$ und einen $\sigma = \pm 1$, so dass für eine fest gewählte Basis $v_1, \ldots, v_n$ der vorige Satz gilt. Dann gibt es eine Folge von Basen $v_{1,k}, \ldots, v_{n,k}$, die jeweils Haarsche Unterräume von $C[a, b]$ aufspannen mit*

$$v_{i,k} \to v_i \qquad \text{für alle } i = 1, \ldots, n$$

gleichmäßig auf $[a, b]$. Insbesondere ist dann V ein Schwach-Chebyshev-Raum.

Beweis: Man setzt zunächst die Basisfunktionen $v_1, \ldots, v_n$ auf ganz $\mathbb{R}$ konstant fort. Die Fortsetzungen spannen einen Schwach-Chebyshev-Raum von

$C(\mathbb{R})$ auf, so dass wir uns auf den Fall $I = \mathbb{R}$ konzentrieren können. Die Funktionen $v \in V$ werden nun durch Faltung mit einem Gaußkern

$$K_k(t) = \frac{k}{\sqrt{2\pi}} e^{-k^2t^2/2}$$

geglättet, also

$$v_{i,k}(x) = \int_{-\infty}^{\infty} v_i(t) K_k(x-t)\, dt = \int_{-\infty}^{\infty} v_i(x-t) K_k(t)\, dt$$

für alle $i = 1, \ldots, n$. Die Funktionen v_i sind auf ganz $\mathbb{R}$ beschränkt und gleichmäßig stetig, und es gilt.

$$\int_{-\infty}^{\infty} K_k(t)\, dt = 1.$$

Man zeigt die gleichmäßige Konvergenz daher mittels

$$|v_i(x) - v_{i,k}(x)| = \left| \int (v_i(x) - v_i(x-t)) K_k(t)\, dt \right| \leq \omega_v(\delta) + 2\|v_i\| \int_{|t|\geq\delta} K_k(t),$$

wobei

$$\omega_v(\delta) = \max_{|t-s|\leq\delta} |v(t) - v(s)|$$

den Stetigkeitsmodul von v bezeichne. Mit Hilfe der Leibnitzschen Formel zeigt man nun

$$\begin{aligned} \det \begin{pmatrix} v_{1,k}(x_1) & \ldots & v_{n,k}(x_1) \\ \vdots & & \vdots \\ v_{1,k}(x_n) & \ldots & v_{n,k}(x_n) \end{pmatrix} \\ = \int \ldots \int_{s_1 < \ldots < s_n} \det \begin{pmatrix} v_1(s_1) & \ldots & v_n(s_1) \\ \vdots & & \vdots \\ v_1(s_n) & \ldots & v_n(s_n) \end{pmatrix} \\ \cdot \det \begin{pmatrix} h_1(s_1) & \ldots & h_n(s_1) \\ \vdots & & \vdots \\ h_1(s_n) & \ldots & h_n(s_n) \end{pmatrix} ds_1 \ldots ds_n, \end{aligned}$$

mit

$$h_i(s) = K_k(s - x_i) \qquad \text{für alle } i = 1, \ldots, n.$$

Für den Gaußschen Kern bilden diese Funktionen einen Haarschen Unterraum, so dass der zweite Faktor im Integral stets dasselbe Vorzeichen hat und nicht 0 wird. Der erste Faktor wechselt das Vorzeichen auch nicht. Außerdem gibt es n Punkte, wo er nicht 0 ist, also auch nicht in einer Umgebung. Es folgt, dass das Integral nicht Null ist, und damit die Behauptung.

Um zu zeigen, dass V ein Schwach-Chebyshev-Raum ist, nehmen wir an, dass es $n+1$ echte Vorzeichenwechsel gibt. Dann hat auch einer der approximierenden Räume $n + 1$ Vorzeichenwechsel, was ein Widerspruch ist. □

1.80. Bemerkung: Die Abschätzung von de la Vallée Poussin für $d(f, V)$ nach unten gilt in Schwach-Chebyshev-Räumen weiterhin. Allerdings braucht die beste Approximation nicht eindeutig bestimmt zu sein.

1.81 Satz: *Sei $V \subseteq C[a, b]$ ein Schwach-Chebyshev-Unterraum der Dimension n. Dann existiert eine beste Approximation $v \in V$ an $f \in C[a, b]$ mit einer Alternante der Länge $n + 1$ aus $E(f - v)$. Umgekehrt ist jedes $v \in V$ mit einer solchen Alternante beste Approximation.*

Beweis: In der Vorlesung. □

1.82 Satz: *$s \in S_m(x_0, \ldots, x_{k+1})$ ist genau dann beste Approximation an $f \in C[x_0, x_{k+1}]$, wenn es ein Intervall $[x_i, x_j]$, $0 \leq i < j \leq k + 1$, gibt, so dass $f - s$ auf $[x_i, x_j]$ eine Alternante der Länge*

$$\dim S_m(x_i, \ldots, x_j) + 1 = m + j - i + 1$$

aus $E(f - s)$ hat.

Beweis: Sei s beste Approximation an f. Es gibt eine beste Approximation $\tilde{s}$ mit einer Alternanten der Länge $m + k + 2$. Sei $[x_i, x_j]$ ein Intervall, so dass $f - \tilde{s}$ dort eine Alternante der Länge $m + j - i + 1$ hat, so dass $j - i$ minimal ist. Dann zeigt man, dass die Alternantenpunkte die Interpolationsbedingung von Schönberg-Whittney erfüllen. $S_m(x_i, \ldots, x_j)$ ist also auf diesen Punkten ein Haarscher Unterraum, so dass $s = \tilde{s}$ gilt.

Wenn umgekehrt s eine gewünschte Alternante hat, so ist die Funktion s auf $S_m(x_i, \ldots, x_j)$ beste Approximation, also auch auf $[a, b]$. □

Kapitel 2

Geschwindigkeit der Approximation

2.1 Der Fejer-Operator

2.1. Definition: Für eine 2π-periodische Funktion f definieren wir die abgebrochene **Fourier-Reihe** durch

$$S_n f(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos(kx) + b_k \sin(kx))$$

mit

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

S_n heißt der **Fourier-Operator** und wir definieren den **Fejer-Operator** durch

$$F_n f(x) = \frac{1}{n} \sum_{k=0}^{n-1} S_k f(x).$$

Den Raum der 2π-periodischen, stetigen Funktionen bezeichnen wir mit $C_{2\pi}$.

2.2. Bemerkung: $S_n f$ ist die orthogonale Projektion von f auf den Raum der Trigonometrischen Polynome $\mathcal{T}_n$ vom Grad n bezüglich des Skalarprodukts

$$< f, g >= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) g(t)\, dt.$$

$S_n f$ konvergiert nicht für alle stetigen Funktionen gegen f. Man kann jedoch zeigen, dass es für alle stetigen, und stückweise stetig differenzierbaren Funktionen gleichmäßig konvergiert. Im Gegensatz dazu stellt sich heraus, dass der Fejeroperator $F_n f$ Konvergenz für alle stetigen Funktionen liefert.

2.3. Definition: Ein linearer Operator auf einem Funktionenraum heißt **positiver Operator**, wenn er positive Funktionen auf positive Funktionen abbildet.

2.4. Bemerkung: Ein positiver linearer Operator L ist monoton. Das heißt aus $f \leq g$ folgt $Lf \leq Lg$.

2.5 Satz: **(Korovkin)** (1) *Sei $L_n : C[a,b] \to \mathcal{P}_n$ eine Folge von positiven linearen Operatoren. Dann konvergiert die Folge $L_n f$ genau dann für alle $f \in C[a,b]$ gleichmäßig gegen f, wenn diese für die Funktionen 1, x und x^2 gilt.*

(2) *Sei $L_n : C_{2\pi} \to \mathcal{T}_n$ eine Folge von positiven linearen Operatoren. Dann konvergiert die Folge $L_n f$ genau dann für alle $f \in C_{2\pi}$ gleichmäßig gegen f, wenn diese für die Funktionen 1, cos(x) und sin(x) gilt.*

Beweis: (1) Die eine Richtung ist trivial. Die Konvergenz gelte nun für $f_0 = 1$, $f_1(x) = x$ und $f_2(x) = x^2$. Es gilt für alle $t, x \in [a,b]$ und $\delta > 0$

$$|f(t) - f(x)| \leq \omega_f(\delta) + \frac{2\|f\|}{\delta^2}(t-x)^2,$$

also

$$-\omega_f(\delta) - \frac{2\|f\|}{\delta^2}(t-x)^2 \leq f(t) - f(x) \leq \omega_f(\delta) + \frac{2\|f\|}{\delta^2}(t-x)^2.$$

Wir fassen diese Funktionen als Funktionen von t auf, wenden den Operator L_n an und benutzen die Monotonie von L_n. Man erhält

$$|L_n f(x) - f(x) L_n f_0(x)| \leq \omega_f(\delta) L_n f_0(x) + \frac{2\|f\|}{\delta^2} L_n h_x(x)$$

mit $f_0 = 1$ und $h_x(t) = (t-x)^2$. Also

$$|L_n f(x) - f(x)| \leq |f(x)||1 - L_n f_0(x)| + \omega_f(\delta)|L_n f_0(x)| + \frac{2\|f\|}{\delta^2}|L_n h_x(x)|.$$

Wie zeigen nun, dass $L_n h_x(x)$ gleichmäßig für alle $x \in [a,b]$ gegen 0 konvergiert. Wegen $(t-x)^2 = t^2 - 2xt + x^2$ folgt

$$\begin{aligned} L_n h_x(x) &= x^2 L_n f_0(x) - 2x L_n f_1(x) - L_n f_2(x) \\ &= x^2 (L_n f_0(x) - 1) - 2x(L_n f_1(x) - x) + (L_n f_2(x) - x^2). \end{aligned}$$

Also

$$|L_n h_x(x)| \leq x^2 \|L_n f_0 - f_0) + 2|x| \|L_n f_1 - f_1\| + \|L_n f_2 - f_2\| \to 0$$

gleichmäßig für alle x. Wählt man nun zu $\epsilon > 0$ ein genügend kleines $\delta > 0$ und anschießend ein genügend großes $N \in \mathbb{N}$, so folgt

$$\|L_n - f\| < \epsilon \qquad \text{für alle } n \geq N.$$

(2) Wir verwenden hier die Funktion

$$h_x(t) = \sin^2\left(\frac{t-x}{2}\right).$$

Die Argumente sind weitgehend äquivalent. □

2.6 Satz: *Der Fejer-Operator F_n ist positiv, es gilt $\|F_n\| = 1$, und $F_n f$ konvergiert für alle $f \in C_{2\pi}$ gleichmäßig gegen f.*

Beweis: Man berechnet zunächst

$$\begin{aligned} S_n f(x) &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left(\frac{1}{2}+\sum_{k=1}^{n}(\cos(kx)\cos(kt)+\sin(kx)\sin(kt))\right)dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left(\frac{1}{2}+\sum_{k=1}^{n}\cos(k(t-x))\right)dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t+x)\frac{\sin((2n+1)t/2)}{2\sin(t/2)}\,dt \\ &= \frac{1}{\pi}\int_{0}^{\pi/2}(f(x+2t)+f(x-2t))\frac{\sin((2n+1)t/2)}{\sin(t/2)}\,dt. \end{aligned}$$

Dies zeigt, dass S_n kein positiver Operator ist. Man erhält jedoch

$$\begin{aligned} F_n f(x) &= \frac{1}{\pi}\int_{0}^{\pi/2}(f(x+2t)+f(x-2t))\frac{1}{n}\sum_{k=0}^{n-1}\frac{\sin((2k+1)t/2)}{\sin(t/2)}\,dt \\ &= \frac{1}{\pi}\int_{0}^{\pi/2}(f(x+2t)+f(x-2t))\frac{\sin(ns)^2}{n\sin(s)^2}\,dt. \end{aligned}$$

Dies ist ein positiver Operator. Man hat für $n \geq 1$, $f_0 = 1$, $f_1 = \cos(x)$, $f_2 = \sin(x)$

$$F_n f_0 = f_0, \quad F_n f_1 = \frac{n-1}{n} f_1, \quad F_n f_2 = \frac{n-1}{n} f_2.$$

Aus dem Satz von Korovkin folgt die Behauptung. Für alle $n \in \mathbb{N}$ gilt $S_n 1 = 1$, also auch $F_n 1 = 1$. Es folgt

$$\frac{1}{2\pi}\int_{0}^{\pi/2}\frac{\sin(ns)^2}{n\sin(s)^2}\,dt = 1.$$

Daraus folgert man $\|F_n\| = 1$. □

2.7 Aufgabe: Zeigen Sie

$$\frac{1}{2}+\sum_{k=1}^{n}\cos(kx) = \frac{\sin((2n+1)x/2}{2\sin(x/2)}.$$

Dazu verwendet man am einfachsten eine komplexe geometrische Reihe. Zeigen Sie außerdem

$$\sum_{k=0}^{n-1}\sin((2k+1)x) = \frac{\sin(nx)^2}{\sin(x)}.$$

2.8 Satz: **(de la Vallée-Poussin)** *Für den Operator $G_n = 2F_{2n} - F_n$ gilt $G_n(f) = f$ für alle $f \in \mathcal{T}_{n-1}$, $n \geq 1$, und*

$$\|f - G_n f\| \leq 4d(f, \mathcal{T}_n).$$

Beweis: Es gilt
$$G_n f = \frac{1}{n}(S_n + \ldots + S_{2n-1}).$$
Daraus folgt die erste Behauptung. $\|G_n\| \leq 3$ folgt aus dem Satz, und die Behauptung folgt aus Satz 1.26. $\square$

2.9 Aufgabe: Sei $f(t) = |\cos(t)|$. Zeigen Sie
$$S_{2n} f = S_{2n+1} f = \frac{2}{\pi} + \frac{4}{\pi} \sum_{k=1}^{n} \frac{(-1)^{k-1}\cos(2kt)}{2k^2 - 1}.$$
Folgern Sie
$$S_n f\left(\frac{\pi}{2}\right) = \frac{2}{(2n+1)\pi}.$$
Folgern Sie daraus
$$G_n f\left(\frac{\pi}{2}\right) \geq \frac{1}{2n}$$
und
$$d(f, \mathcal{T}_n) \geq \frac{1}{8n}.$$
Folgern Sie daraus für $h(x) = |x|$
$$d(h, \mathcal{P}_n) \geq \frac{1}{8n}.$$

2.10. Definition: Wir sagen, eine Funktion $f : [a,b] \to \mathbb{R}$ sei **Lipschitz-stetig** mit Exponent $0 < \alpha$, wenn für alle $x, y \in \mathbb{R}$
$$|f(x) - f(y)| \leq M|x-y|^\alpha.$$
gilt. Die Klasse dieser Funktionen bezeichnen wir mit $\mathrm{Lip}_\alpha M$.

2.11 Satz: *Die $f \in C_{2\pi}$ und $f \in \mathrm{Lip}_\alpha M$. Dann gilt*
$$\|F_n f - f\| \leq \frac{\pi 2^\alpha}{1-\alpha^2} \frac{M}{n^\alpha}$$

Beweis: Aus der Integraldarstellung des Fejer-Operators folgt
$$F_n f(x) - f(x) = \frac{1}{\pi} \int_0^{\pi/2} (f(x+2t) + f(x-2t) - 2f(x)) \frac{\sin(nt)^2}{\sin(t)^2}\, dt.$$
Wegen der Lipschitzstetigkeit folgt
$$|f(x+2t) + f(x-2t) - 2f(x)| \leq 2^{1+\alpha} M x^\alpha.$$
Verwendet man außerdem
$$\int_0^\infty \frac{\sin(t)^2}{t^{2-\alpha}}\, dt < \int_0^1 t^\alpha\, dt + \int_1^\infty \frac{1}{t^{2-\alpha}}\, dt = \frac{2}{1-\alpha^2}$$

so folgt die Behauptung wegen

$$\begin{aligned}\|F_n f - f\| &\leq \frac{M\pi}{n2^{1-\alpha}} \int_0^{\pi/2} \frac{\sin(nt)^2}{t^{2-\alpha}}\,dt \\ &= \frac{M\pi}{n^\alpha 2^{1-\alpha}} \int_0^{n\pi/2} \frac{\sin(t)^2}{t^{2-\alpha}}\,dt \\ &\leq \frac{\pi 2^\alpha}{1-\alpha^2} \frac{M}{n^\alpha}.\end{aligned}$$

□

2.12 Aufgabe: (1) Sei $f : [a,b] \to \mathbb{R}$ stetig. Zeigen Sie

$$\omega_f(\alpha+\beta) \leq \omega_f(\alpha) + \omega_f(\beta).$$

für $\alpha, \beta > 0$.

(2) Folgern Sie für $\lambda \geq 0$

$$\omega_f(\lambda\alpha) \leq (\lambda+1)\omega_f(\alpha).$$

(2) Folgern Sie, dass $\omega_f(\delta)$ stetig in δ ist.

(3) Bestimmen Sie den Stetigkeitsmodul von $\delta \mapsto \omega_f(\delta)$ auf $[0,\infty)$.

2.2 Korovkin-Operatoren

Wir führen die folgende Klasse von positiven **Korovkin-Operaten** ein und geben eine Abschätzung für den Approximationsfehler.

2.13 Satz: *Sei*

$$k_n(t) = \frac{1}{2} + \sum_{k=1}^{n} \rho_{k,n} \cos(kn) \geq 0$$

für alle $t \in [0, 2\pi)$, und

$$K_n f(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) k_n(x-t)\,dt.$$

Dann gilt

$$\|K_n - f\| \leq \omega_f(\delta) \left(1 + \frac{\pi}{\delta\sqrt{2}} \sqrt{1-\rho_{1,n}}\right).$$

Insbesondere konvergiert $K_n f$ gegen f gleichmäßig, falls nur $\rho_{1,n} \to 1$.

2.14. Beispiel: Der Fejer-Operator ist von diesem Typ mit

$$\rho_{1,n} = \frac{n-1}{n}.$$

Deswegen folgt die Konvergenz auch aus diesem Satz. Jedoch bekommt man selbst für differenzierbare f keine optimalen Abschätzungen.

2.15 Satz: *Jeder Kern mit*

$$k_n(t) = \frac{1}{2} + \sum_{k=1}^{n} \rho_{k,n} \cos(kn) \geq 0 \qquad \textit{für alle } t \in \mathbb{R}$$

lässt sich als

$$k_n(t) = \frac{1}{2} \left| \sum_{k=0}^{n} a_k e^{ikt} \right|^2$$

schreiben mit $a_0, \ldots, a_n \in \mathbb{R}$ *und*

$$\sum_{k=0}^{n} a_k^2 = 1.$$

Beweis: In der Vorlesung. □

2.16. Bemerkung: Insbesondere gilt dann

$$\rho_{1,n} = a_0 a_1 + \ldots + a_{n-1} a_n.$$

Um optimale Operatoren zu erhalten, muss man also nur diesen Ausdruck unter der Nebenbedingung

$$a_0^2 + \ldots + a_n^2 = 1$$

maximieren. Die Lagrange-Bedingung für dieses Maximum lautet

$$\begin{aligned} a_1 &= 2\lambda a_0, \\ a_2 + a_0 &= 2\lambda a_1, \\ &\vdots \\ a_n + a_{n-2} &= 2\lambda a_{n-1}, \\ a_{n-1} &= 2\lambda a_n. \end{aligned}$$

Man bezeichnet mit

$$U_k(x) = \frac{1}{n+1} T'_{k+1}(x)$$

die Chebyshev-Polynome 2. Art. Es gilt dann $U_0(x) = 1$, $U_1(x) = 2x$ und

$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x) \qquad \text{für alle } n \in \mathbb{N},$$

genau wie für die Chebyshev-Polynome. Es sei λ eine Nullstelle von U_{n+1}. Dann löst

$$a_k = a_0 U_k(\lambda), \quad a_o = \frac{1}{\sum_{k=0}^{n} U_k(\lambda)^2}$$

die Lagrange-Gleichungen. Außerdem hat man

$$a_0 a_1 + \ldots + a_{n-1} a_n = \lambda.$$

Wir wählen also die größte Nullstelle von U_{n+1}, nämlich

$$\rho_{1,n} = \lambda = \cos\left(\frac{\pi}{n+1}\right).$$

2.17 Satz: **(Jackson)** *Für jedes* $f \in C_{2\pi}$ *gilt*

$$d(f, \mathcal{T}_n) \le \left(1 + \frac{\pi^2}{2}\right) \omega_f \left(\frac{1}{n}\right).$$

Beweis: Der Satz folgt mit der optimalen Wahl von $\rho_{1,1} = \cos(\pi/(n+1))$ wegen

$$1 - \cos\left(\frac{\pi}{n+1}\right) = \frac{\pi^2}{2(n+1)^2} - \frac{\pi^4}{4!(n+1)^4} \pm \ldots \leq \frac{\pi^2}{2(n+1)^2}.$$

$\square$

2.18. Bemerkung: (1) Für differenzierbare Abbildungen $f \in C_{2\pi}$ folgt

$$d(f, \mathcal{T}_n) \leq \frac{c\|f'\|}{n}$$

mit $c = 1 + \pi^2/2$.

(2) Falls $f \in \mathrm{Lip}_\alpha M$ ist, so folgt

$$d(f, \mathcal{T}_n) \leq \frac{cM}{n^\alpha}.$$

2.19. Bemerkung: Jackson benutzte statt der Korovkin-Operatoren den Kern

$$K_n(t) = c\left(\frac{\sin nt/2}{\sin t/2}\right)^n.$$

2.20 Satz: *Sei f k-mal stetig differenzierbar, und sei $f^{(k)} \in \mathrm{Lip}_\alpha M$ für $0 < \alpha \leq 1$. Dann gilt*

$$d(f, \mathcal{T}_n) \leq \left(1 + \frac{\pi^2}{2}\right)^{k+1} \frac{M}{n^{k+\alpha}}.$$

Beweis: (1) Für $k = 0$ folgt des Satz unmittelbar aus dem obigen Satz von Jackson. Sei also $k \geq 1$.

(2) Wir setzen

$$\tilde{C}_{2\pi} = \{f \in C_{2\pi} : \textstyle\int_{-\pi}^{\pi} f = 0\},$$

und

$$\tilde{\mathcal{T}}_n = \mathcal{T}_n \cap \tilde{C}_{2\pi}.$$

$\tilde{\mathcal{T}}_n$ besteht dann aus allen trigonometrischen Polynomen, deren konstantes Glied wegfällt. Die Korovkin-Operatoren bilden den Raum $\tilde{C}_{2\pi}$ nach $\tilde{\mathcal{T}}_n$ ab. Dies bedeutet, dass wir für alle $f \in \tilde{C}_{2\pi}$

$$d(f, \tilde{\mathcal{T}}_n) \leq c\omega_f\left(\frac{1}{n}\right)$$

erhalten, sowie für differenzierbares $f \in \tilde{C}_{2\pi}$

$$d(f, \tilde{\mathcal{T}}_n) \leq \frac{c\|f'\|}{n},$$

wieder mit $c = 1 + \pi^2/2$.

(3) Für stetig differenzierbares $f \in C_{2\pi}$ und $w \in \mathcal{T}_n$ gilt

$$d(f, \mathcal{T}_n) = d(f - w, \mathcal{T}_n) \leq \frac{c\|f' - w'\|}{n}.$$

Also

$$d(f, \mathcal{T}_n) \leq \frac{c\, d(f', \tilde{\mathcal{T}}_n)}{n}.$$

Denn jedes $v \in \tilde{\mathcal{T}}_n$ lässt sich als $v = w'$ mit einem $w \in \mathcal{T}_n$ schreiben.

(4) Es gilt nun $f' \in \tilde{C}_{2\pi}$. Sei $w \in \mathcal{T}_n$ beliebig. Dann ist $w - \alpha \in \tilde{\mathcal{T}}_n$ für $\alpha = \int_{-\pi}^{\pi} w$. Es folgt für zweimal stetig differenzierbares $f \in \tilde{C}_{2\pi}$

$$d(f', \tilde{\mathcal{T}}_n) = d(f' - (w - \alpha), \tilde{\mathcal{T}}_n) \leq \frac{c\|f'' - w'\|}{n}$$

Also

$$d(f', \tilde{\mathcal{T}}_n) \leq \frac{c\, d(f'', \tilde{\mathcal{T}}_n)}{n}.$$

Da für $k \geq 1$ auch $f^{(k)} \in \tilde{C}_{2\pi}$ ist, erhält man mit (3) induktiv für k-mal stetig differenzierbares $f \in \tilde{C}_{2\pi}$,

$$d(f, \mathcal{T}_n) \leq \frac{c^k\, d(f^{(k)}, \tilde{\mathcal{T}}_n)}{n^k}.$$

(5) Wegen $f^{(k)} \in \mathrm{Lip}_\alpha M$ erhalten wir schließlich

$$d(f, \mathcal{T}_n) \leq \frac{c^k\, d(f^{(k)}, \tilde{\mathcal{T}}_n)}{n^k} \leq \frac{c^{k+1}\, M}{n^{k+\alpha}}.$$

□

2.21. Bemerkung: Für k-mal stetig differenzierbares $f \in C_{2\pi}$ folgt also

$$d(f, \mathcal{T}_n) \leq \frac{c^k\|f^{(k)}\|}{n^k}.,$$

da in diesem Fall der Satz mit $k - 1$ anstelle von k, $\alpha = 1$ und $M = \|f^{(k)}\|$ angewendet werden kann.

2.22 Satz: (1) *Für $f \in C[-1, 1]$ folgt*

$$d(f, \mathcal{P}_n) \leq \left(1 + \frac{\pi^2}{2}\right) \omega_f\left(\frac{1}{n}\right).$$

(2) *Falls f stetig differenzierbar ist und $n \geq 1$, so gilt*

$$d(f, \mathcal{P}_n) \leq \left(1 + \frac{\pi^2}{2}\right) d(f', \mathcal{P}_{n-1})$$

(3) *Falls f k-mal stetig differenzierbar ist und $f^{(k)} \in \mathrm{Lip}_\alpha M$, $0 \leq \alpha \leq 1$, so gilt für $n > k$*

$$d(f, \mathcal{P}_n) \leq \left(1 + \frac{\pi^2}{2}\right)^{k+1} \frac{M}{n(n-1) \cdot \ldots \cdot (n-(k-1))(n-k)^\alpha}.$$

Beweis: (1) Setzt man

$$\hat{f}(t) = f(\cos t)$$

so folgt

$$\omega_{\hat{f}}(\delta) \leq \omega_f(\delta)$$

wegen $|\cos(t_1) - \cos(t_2)| \leq |t_1 - t_2|$. Die Abbildung $f \to \hat{f}$ bildet $\mathcal{P}_n$ isometrisch in C_n, dem Raum der Kosinuspolynome ab. Es folgt

$$d(f, \mathcal{P}_n) = d(\hat{f}, C_n) \leq d(\hat{f}, \mathcal{T}_n) \leq c\,\omega_{\hat{f}}\left(\frac{1}{n}\right) \leq c\,\omega_f\left(\frac{1}{n}\right),$$

wieder mit $c = 1 + \pi^2/2$.

(2) Wenn $p \in \mathcal{P}_{n-1}$ die beste Approximation an f' ist, so können wir immer $p = q'$ mit $q \in \mathcal{P}_n$ schreiben (im Unterschied zum trigonometrischen Fall). Also

$$d(f, \mathcal{P}_n) = d(f - q, \mathcal{P}_n) \leq c\,\omega_{f-q}\left(\frac{1}{n}\right) \leq \frac{c\,\|f' - q'\|}{n} = \frac{c\,d(f', \mathcal{P}_{n-1})}{n}.$$

(3) Folgt per Induktion. □

2.23. Bemerkung: Für k-mal stetig differenzierbares f erhalten wir mit $k-1$ anstelle von k, $\alpha = 1$ und $M = \|f^{(k)}\|$

$$d(f, \mathcal{P}_n) \leq \frac{c^{k+1}}{n \cdot \ldots \cdot (n-k)} \|f^{(k)}\|$$

mit $c = 1 + \pi^2/2$. Man kann dieses Resultat verschärfen zu

$$d(f, \mathcal{P}_n) \leq \frac{\pi}{2} \frac{1}{(n+1)^k} \|f^{(k)}\|$$

für k-mal stetig differenzierbares f.

2.3 Umkehrsätze

2.24. Beispiel: Wir haben weiter oben mit Hilfe der Abschätzung von de la Vallée Poussin schon in einer Aufgabe gezeigt, dass im Fall der Funktion $f(t) = |\cos(t)|$

$$d(f, \mathcal{T}_n) \geq \frac{c}{n}$$

für alle $n \in \mathbb{N}$ mit einer Konstanten $c > 0$ gilt. Die Abschätzung von Jackson lässt sich in diesem Fall also im Wesentlichen nicht verbessern. Unser Ziel ist, derartige Abschätzungen für allgemeine Funktionen herzuleiten.

2.25 Satz: **(Bernstein)** *für $g \in \mathcal{T}_n$ gilt*

$$\|g'\| \leq n\|g\|.$$

Beweis: Man darf annehmen, dass

$$g'(0) = \|g'\| = 1.$$

gilt. Angenommen, $\|g\| < 1/n$. Dann hat

$$h(t) = \sin(nt) - ng(t)$$

$2n$ alternierende Punkte. Damit hat h' mindestens $2n$ verschiedene Nullstellen. Außerdem gilt

$$h'(0) = n - ng'(0) = 0,$$

sowie

$$h''(0) = -ng''(0) = 0,$$

weil g' ein Maximum in 0 hat. Also ist 0 doppelte Nullstelle und h' hat daher $2n+1$ Nullstellen mit Vielfachheit, was ein Widerspruch ist. □

2.26 Satz: **(Bernstein 1912)** *Sei $f \in C_{2\pi}$ und für $0 < \alpha \leq 1$, sowie $k \in \mathbb{N}_0$*

$$d(f, \mathcal{T}_n) = O\left(\frac{1}{n^{k+\alpha}}\right)$$

für alle $n \in \mathbb{N}$. Dann ist f k-mal stetig differenzierbar und

$$f^{(k)} \in \mathrm{Lip}_\alpha$$

im Fall $\alpha < 1$, bzw.

$$\omega_{f^{(k)}}(\delta) = O\left(\delta \log \frac{1}{\delta}\right)$$

im Fall $\alpha = 1$.

Beweis: Wir wählen $s_n \in \mathcal{T}_n$ mit

$$\|f - s_n\| \leq \frac{C}{n^{k+\alpha}}$$

Sei

$$v_0 = s_0, \quad v_n = s_{2^n} - s_{2^{n-1}}.$$

Aufgrund der Abschätzung von Bernstein gilt für $r \leq k$

$$\begin{aligned}
\|v_n^{(k)}\| &\leq 2^{rn}\|v_n\| \\
&\leq 2^{rn}(\|s_{2^n} - f\| + \|s_{2^{n+1}} - f\|) \\
&\leq \frac{C_1}{2^{n\alpha}}.
\end{aligned}$$

Also konvergiert $\sum_{n\in\mathbb{N}} v_n^{(r)}$ gleichmäßig für $1 \leq r \leq k$. Damit ist f k-mal stetig differenzierbar und

$$\|f^{(k)} - \sum_{\nu=0}^{n} v_\nu^{(k)}\| \leq \frac{C_2}{2^{n\alpha}}.$$

Wegen der Monotonie von $d(f, \mathcal{T}_n)$ gilt dann auch

$$d(f^{(k)}, \mathcal{T}_n) \leq \frac{C_3}{n^\alpha}.$$

Wir können also im Folgenden der Einfachheit halber $k = 0$ im Satz annehmen. Angenommen

$$|s - t| \leq \delta < \frac{1}{2}$$

und

$$2^{m-1} \leq \frac{1}{\delta} \leq 2^m.$$

Wir schätzen ab

$$|f(t) - f(s)| \leq \sum_{n=0}^{m-1} |v_n(t) - v_n(s)| + \sum_{n=m}^{\infty} |v_n(s) - v_n(t)|.$$

Nun gilt

$$\sum_{n=m}^{\infty} |v_n(s) - v_n(t)| \leq \frac{C_4}{2^{m\alpha}} \leq C_4 \delta^\alpha.$$

Außerdem

$$\begin{aligned}
\sum_{n=0}^{m-1} |v_n(t) - v_n(s)| &\leq \sum_{n=0}^{m-1} \|v_n'\| \delta \\
&\leq \delta \sum_{n=0}^{m-1} 2^n \|v_n\| \\
&\leq \delta \sum_{n=0}^{m-1} C_5 2^{n(1-\alpha)}.
\end{aligned}$$

Falls $\alpha < 1$, so folgt

$$\sum_{n=0}^{m-1} 2^{n(1-\alpha)} = \frac{2^{m(1-\alpha)}}{2^{1-\alpha} - 1} \leq C_6 \delta^{\alpha - 1}.$$

Es folgt die Behauptung für $\alpha < 1$. Im Fall $\alpha = 1$ erhält man also

$$|f(s) - f(t)| \leq (C_7 m + C_8)\delta.$$

Außerdem

$$(m-1)\log 2 \leq \log \frac{1}{\delta}.$$

Es folgt die Behauptung. □

2.27. Beispiel: Für die Funktion

$$f(t) = \sum_{k=1}^{\infty} \frac{\sin kt}{k^2}$$

gilt offenbar

$$d(f, \mathcal{T}_n) \leq \sum_{k=n+1}^{\infty} \frac{1}{k^2} \leq \frac{c}{n}.$$

Aber für $t = 1/n$

$$f(t) \geq \sum_{k=1}^{n} \frac{\sin kt}{k^2} - \sum_{k=n+1}^{\infty} \frac{1}{k^2} \geq c\frac{\log n}{n}.$$

Also ist $f \notin \mathrm{Lip}_1$. Dies zeigt, dass der obige Satz für $\alpha = 1$ nicht gilt.

2.28. Bemerkung: Wir haben im vorigen Abschnitt

$$f^{(k)} \in \mathrm{Lip}_1 \;\Rightarrow\; d(f, \mathcal{T}_n) = O\left(\frac{1}{n^{k+1}}\right)$$

hergeleitet. Die Umkehrung gilt aufgrund des obigen Beispiels nicht. Zygmund hat den Satz mit Hilfe des modifizierten Modulus

$$\tilde{\omega}_f(\delta) = \sup_{|h|<\delta} |f(t+h) - 2f(t) + f(t-h)|$$

verschärft. Es gilt dann

$$\tilde{\omega}_{f^{(k)}}\left(\frac{1}{n}\right) = O\left(\frac{1}{n}\right) \;\Leftrightarrow\; d(f, \mathcal{T}_n) = O\left(\frac{1}{n^{k+1}}\right).$$

2.29. Bemerkung: Aus dem Satz folgt, dass f genau dann unendlich oft differenzierbar ist, wenn

$$\lim_{n\to\infty} n^k d(f, \mathcal{T}_n) = 0$$

für alle $k \in \mathbb{N}$ gilt. Dies umfasst unser Ergebnis.

2.30 Aufgabe: Zeigen Sie, dass aus dem Resultat von Zygmund der Satz 2.20 folgt.

2.31 Satz: *Sei $p_n \in \mathcal{P}_n$. Dann gilt*

$$|p_n'(x)| \leq \frac{n}{\sqrt{1-x^2}} \|p_n\|_{[-1,1]}$$

für alle $x \in (-1, 1)$.

Beweis: Der Satz folgt unmittelbar aus der Bernsteinschen Abschätzung für $\mathcal{T}_n$, indem man $\hat{p}_n = p_n(\cos t)$ verwendet. $\square$

2.32 Aufgabe: (1) Zeigen Sie, dass die für $p \in \mathcal{P}_n$ mit $\|p\|_{[-1,1]} \leq 1$ in der Vorlesung gewonnene Abschätzung

$$|p'(x)| \leq \frac{n}{\sqrt{1-x^2}} \qquad \text{für alle } x \in (-1, 1)$$

für T_n scharf ist.

(2) Zeigen Sie für die Chebyshev-Polynome $T_n(x) = \cos(n \arccos(x))$

$$|T_n'(\pm 1)| = n^2.$$

(3) Zeigen Sie, dass für jedes andere Polynom $p \in \mathcal{P}_n$ mit $\|p\|_{[-1,1]} \leq 1$ die Ungleichung

$$|p'(\pm 1)| < n^2.$$

gilt. Verwenden Sie dazu dieselbe Technik wie beim Beweis der Bernsteinschen Ungleichung für trigonometrische Polynome.

(4) Folgern sie für $p \in \mathcal{P}_n$ mit $\|p\|_{[-1,1]} \leq 1$

$$|p'(x)| \leq \frac{n^2}{|x|} \qquad \text{für alle } 0 < |x| < 1.$$

indem Sie $\tilde{p}(t) = p(xt)$ betrachten.

(5) Berechnen Sie das Minimum dieser beiden Abschätzungen und folgern Sie für $p \in \mathcal{P}_n$ mit $\|p\|_{[-1,1]} \leq 1$

$$\|p'\|_{[-1,1]} \leq \sqrt{2} n^2.$$

2.33. Bemerkung: Man kann zeigen

$$\|p_n'\|_{[-1,1]} \leq n^2 \|p_n\|_{[-1,1]}.$$

Dies ist die **Markovsche Ungleichung**.

2.34 Satz: **(Bernstein 1912)** *Sei* $f \in C[-1,1]$ *und für* $0 < \alpha \leq 1$

$$d(f, \mathcal{P}_n) = O\left(\frac{1}{n^{k+\alpha}}\right)$$

für alle $n \in \mathbb{N}$. *Dann ist* f *k-mal stetig differenzierbar in* $(-1,1)$, *und auf abgeschlossenen Teilintervallen* I *von* $(-1,1)$ *gilt*

$$f^{(k)}|_I \in \mathrm{Lip}_\alpha$$

im Fall $\alpha < 1$, *sowie*

$$\omega_{f^{(k)}|_I}(\delta) = O\left(\delta \log \frac{1}{\delta}\right)$$

im Fall $\alpha = 1$.

Beweis: Der Beweis ist völlig analog zum obigen Beweis für $\mathcal{T}_n$. □

2.35. Bemerkung: Die Funktion $f(x) = \sqrt{x}$ zeigt, dass f nicht unbedingt in den Randpunkten differenzierbar sein muss, wenn

$$d(f, \mathcal{T}_n) = O\left(\frac{1}{n}\right)$$

gilt.

2.4 Analytische Funktionen

2.36 Satz: **(Hermite)** *Sei $U \subseteq \mathbb{C}$ offen $f : U \to \mathbb{C}$ analytisch in U und*

$$z_0, \ldots, z_n \in U$$

Interpolationspunkte. Dann gilt für das Hermitesche Interpolationspolynom p auf f in diesen Punkten

$$p(z) = \frac{1}{2\pi i} \oint_\gamma \frac{(\omega_n(z) - \omega_n(t))f(t)}{(z-t)\omega_n(t)}\,dt.$$

Dabei sei γ ein nullhomotoper Weg in U, der jeden der Punkte z, sowie $z_0, \ldots, z_n$ genau einmal umläuft und

$$\omega_n(t) = (t - z_0) \cdot \ldots \cdot (t - z_n).$$

Für den Fehler gilt

$$p(z) - f(z) = \frac{\omega_n(z)}{2\pi i} \oint_\gamma \frac{f(t)}{\omega_n(t)(z-t)}\,dt.$$

Beweis: Die Funktion

$$q_t(z) = \frac{\omega_n(z) - \omega_n(t)}{z - t}$$

ist ein Polynom n-ten Grades, dessen Koeffizienten Funktionen von t sind. Daher ist $p(z)$ in der Tat ein Polynom n-ten Grades. Nach der Formel von Cauchy gilt

$$f(z) = \frac{1}{2\pi i} \oint_\gamma \frac{f(t)}{z-t}\,dt.$$

Es folgt die Fehlerformel. □

2.37. Beispiel: Sei $z_0 = \ldots = z_n = 0$ und U eine offene Menge, die

$$D_r = \{z \in \mathbb{C} : |z| \le r\}$$

enthält. Dann gilt für $0 < \rho < r$

$$\begin{aligned}\|f - p_n\|_{D_\rho} &\le \frac{1}{2\pi}\rho^n \oint_{\partial D_r} \frac{|f(t)|}{|t|r^n}\,dt \\ &\le \|f\|_{D_r} \frac{r}{r-\rho}\left(\frac{\rho}{r}\right)^n,\end{aligned}$$

wobei $p_n \in \mathcal{P}_n$ das Hermitesche Interpolationspolynom an f in 0 ist. p_n ist also die abgebrochene Potenzreihe von f, und es ist bekannt, dass diese Reihe geometrisch auf allen Kreisen konvergiert, über die hinaus f analytisch fortsetzbar ist.

2.38. Bemerkung: Diese **Formel von Hermite** bleibt auch richtig, wenn γ ein Weg in $\overline{U}$ ist und f auf γ stetig. Dazu muss man lediglich wissen, dass auch die Formel von Cauchy für diesen Fall gilt. Wenn also f in D_r stetig und im Innern von D_r analytisch ist, so gilt die Abschätzung des Beispiels ebenfalls.

2.39 Aufgabe: Sei f auf D_r analytisch für ein $r > 1$. Zeigen Sie dann für die Interpolation in den n-ten Einheitswurzeln

$$\|f - p_n\|_{D_1} \leq \|f\|_{D_r} \frac{r}{r-1} \frac{2}{r^n - 1}.$$

2.40. Definition: Wir definieren die **Joukowski-Abbildung**

$$\phi(z) = \frac{1}{2}\left(z + \frac{1}{z}\right)$$

auf $\mathbb{C}$, sowie

$$E_r = \phi(\partial D_r).$$

2.41. Bemerkung: Man hat

$$\phi(re^{it}) = \frac{1}{2}\left(r + \frac{1}{r}\right)\cos(t) + i\frac{1}{2}\left(r - \frac{1}{r}\right)\sin(t).$$

Also ist E_r in der Tat eine Ellipse mit Halbachsen

$$\frac{1}{2}\left(r + \frac{1}{r}\right), \quad \frac{1}{2}\left(r - \frac{1}{r}\right).$$

Die Summe der Halbachsen ist daher r. Man berechnet, dass die Brennpunkt von E_r in ± 1 liegen. Der Rand des Einheitskreises wird wegen

$$\phi(e^{it}) = \cos(t) \in [-1, 1].$$

in der Tat auf $[-1, 1]$ abgebildet, allerdings nicht surjektiv, da $\phi(z) = \phi(\overline{z})$ gilt. Aus $\phi(z) = w$ folgt

$$z_{1,2} = w \pm \sqrt{w^2 - 1}.$$

Das Produkt dieser Punkte ist 1. Daher liegt einer im Einheitskreis, der andere außerhalb, und sie fallen nur für $w \in [-1, 1]$ auf den Rand des Einheitskreis. Die Joukowski-Abbildung bildet daher $\mathbb{C} \setminus D_1$ konform auf $\mathbb{C} \setminus [-1, 1]$ ab.

2.42 Satz: *Für das n-te Chebyshev-Polynom T_n gilt*

$$T_n(\phi(z)) = \phi(z^n)$$

für alle $z \in \mathbb{C} \setminus \{0\}$.

Beweis: Die Gleichung ist richtig für $|z| = 1$, denn für $z = e^{it}$ gilt

$$T_n(\phi(z)) = T_n(\cos(t)) = \cos(nt) = \phi(z^n).$$

Die Differenz $T_n(\phi(z)) - \phi(z^n)$ ist eine in $\mathbb{C} \setminus \{0\}$ analytische Funktion, die daher überall gleich 0 ist. □

2.43 Satz: **(Bernstein 1912)** *Sei $E = [-1,1]$ und $f : E \to \mathbb{C}$ analytisch im Innern von E_r mit $r > 1$ und stetig auf E_r, sowie $z_0, \ldots, z_n$ die Nullstellen von T_{n+1}. Dann gilt für das Interpolationspolynom p_n in diesen Punkten*

$$\|f - p_n\|_E \leq \frac{C_r \|f\|_{E_r}}{r^n}$$

mit einer Konstanten $C_r > 0$, die nur von r abhängt.

Beweis: In diesem Fall gilt

$$\omega_n = 2^n T_{n+1}(z).$$

Also

$$\|\omega_n\|_E = 2^n.$$

Andererseits gibt es für $w \in E_r$ ein t mit $|t| = r$ und $\phi(t) = w$. Es folgt für $n \geq 1$

$$\omega_n(\phi(t)) \geq 2^{n-1}\left(r^n - \frac{1}{r^n}\right) \geq 2^{n-1}\left(1 - \frac{1}{r}\right) r^n$$

Wir wenden die Fehlerformel nun auf den Weg

$$\gamma(t) = \phi(re^{it}), \quad 0 \leq t \leq 2\pi,$$

an, der ∂E_r parametrisiert. Es folgt

$$\|f - p_n\|_E \leq \frac{l(\gamma)}{2\pi} \frac{\|f\|_{E_r} \|\omega_n\|_E}{d(\gamma, E) \min_{t \in E_r} |\omega_n(t)|} \leq \frac{C_r \|f\|_{E_r}}{r^n}.$$

Dabei bezeichne $l(\gamma)$ die Länge des Weges γ und $d(\gamma, E)$ den Abstand von ∂E_r zu $[-1,1]$. $\square$

2.44 Aufgabe: Zeigen Sie

$$d(z, [-1,1]) \geq \max\{\mathrm{Im}\,(z), |z-1|, |z+1|\}.$$

Zeigen Sie außerdem geometrisch aus der Brennpunkteigenschaft, dass die Kreise mit Radius

$$\frac{1}{2}\left(r + \frac{1}{r}\right) - 1$$

um ± 1 ganz in E_r liegen. Folgern Sie für $1 \leq r \leq 2$

$$d(E_r, [-1,1]) \geq \frac{1}{4}(r-1)^2.$$

Diskutieren Sie dazu die Funktion

$$\frac{1/2(r + 1/r) - 1}{(r-1)^2}.$$

Folgern Sie für $1 \leq r \leq 2$

$$C_r = O((r-1)^2).$$

und für $r \geq 2$

$$C_r = O(r-1).$$

2.45 Satz: **(Bernstein-Walsh)** *Sei $E \subseteq \mathbb{C}$ kompakt mit zusammenhängendem Komplement und $\phi : \mathbb{C}_\infty \setminus D_1 \to \mathbb{C}_\infty \setminus E$ konform, stetig auf ∂E, mit $\phi(\infty) = \infty$, sowie für $r \geq 1$*

$$E_r = \{\phi(z) : |z| = r\}.$$

Dann gilt für alle $p_n \in \mathcal{P}_n$

$$\|p_n\|_{\partial E_r} \leq r^n \|p_n\|_E.$$

Beweis: Da ϕ in ∞ differenzierbar ist, existiert

$$\lim_{z \to \infty} \frac{\phi(z)}{z} = \phi'(\infty).$$

Da $p_n \in \mathcal{P}_n$ ist und $\phi(\infty) = \infty$, ist daher die Funktion

$$h(z) = \frac{p_n(\phi(z))}{z^n} = \frac{p_n(\phi(z))}{\phi(z)^n} \frac{\phi(z)^n}{z^n}$$

in ∞ beschränkt und daher dort holomorph ergänzbar. h ist auf $\mathbb{C} \setminus D_1$ analytisch. Aus dem Maximumsprinzip folgt

$$\left| \frac{p_n(\phi(z))}{\phi(z)^n} \right| = |h(z)| \leq \|h\|_{\partial E} = \|p_n\|_E.$$

Für $\phi(z) \in \partial E_r$ gilt $|z| = r$. Es folgt daher die Behauptung. □

2.46. Bemerkung: Im Fall $E = D_1$ ist $\phi(z) = z$ und wir erhalten

$$\|p_n\|_{D_r} \leq r^n \|p_n\|_{D_1}.$$

Im Fall $E = [-1, 1]$ folgt

$$\|p_n\|_{\partial E_r} \leq r^n \|p_n\|_{[-1,1]}.$$

2.47 Satz: *Sei E, ϕ und E_r wie im vorigen Satz, sowie $f \in E \to \mathbb{C}$ eine Funktion mit*

$$\limsup_{n \to \infty} d(f, \mathcal{P}_n)^{1/n} = \frac{1}{r}$$

für ein $r > 1$. Dann lässt sich f analytisch im Innern von E_r fortsetzen.

Beweis: Für $1 < \rho < r$ gibt es dann Polynome $p_n \in Pol_n$ mit

$$\|f - p_n\|_E \leq \frac{c}{\rho^n}.$$

Also

$$\|p_{n+1} - p_n\|_E \leq \frac{2c}{\rho^n}.$$

Es folgt aus dem Satz von Bernstein-Walsh für $1 < \tilde{\rho} < \rho$

$$\|p_{n+1} - p_n\|_{E_{\tilde{\rho}}} \leq \frac{2c\tilde{\rho}^n}{\rho^n}.$$

Aufgrund des Maximumsprinzips gilt diese Abschätzung auch im Innern U_ρ von E_ρ. Daher konvergieren diese Polynome dort gleichmäßig gegen eine analytische Fortsetzung von f. Es gilt

$$U_r = \bigcup_{1<\rho<r} U_\rho.$$

Also ist f ins Innere von E_r analytisch fortsetzbar. □

2.5 Rationale Approximation

2.48. Definition: Wir definieren

$$\mathcal{R}_{n,m} = \mathcal{R}_{n,m}[-1,1] = \left\{ \frac{p_n}{q_n} : p_n \in \mathcal{P}_n,\ q_n \in \mathcal{P}_n,\ q_n > 0 \text{ auf } [-1,1]. \right\}$$

Die Elemente von $\mathcal{R}_{n,m}$ nennt man **rationale Funktionen**.

2.49. Bemerkung: Die Darstellung von $r \in \mathcal{R}_{n,m}$ ist nicht eindeutig. Sie wird zum Beispiel eindeutig, indem man die Teilerfremdheit der Polynome fordert, sowie $\|q_n\| = 1$.

2.50. Bemerkung: Rationale Funktionen lassen sich in Kettenbrüche

$$r(x) = p_0(x) + \cfrac{a_0}{p_1(x) + \cfrac{a_1}{p_2(x) + \ldots}}$$

entwickeln. Dieser Kettenbruch bricht ab. Es gilt $p_0 \in \mathcal{P}_{n-m}$, $a_\nu \in \mathbb{R}$ und

$$\sum_{\nu \geq 1} \deg p_\nu = \deg q_m.$$

2.51 Satz: *Es existiert eine beste Approximation bezüglich $\mathcal{R}_{n,m}$ an jedes $f \in C[-1,1]$.*

Beweis: Wir wählen eine Folge $r_k = p_k/q_k \in \mathcal{R}_{n,m}$ mit $\|q_k\| = 1$, so dass p_k und q_k teilerfremd sind, sowie

$$\lim_{k to \infty} \|f - r_k\| = d(f, \mathcal{R}_{n,m}).$$

Da q_k, $k \in \mathbb{N}$ eine konvergente Teilfolge besitzt, können wir

$$\lim_{k \to \infty} q_k = q \in \mathcal{P}_m$$

annehmen. Es gilt $\|q\| = 1$, also $q(x) = 1$ für ein $x \in [-1,1]$. Sei nun $[a,b] \subseteq [-1,1]$ mit $a < b$ und

$$q(x) > \frac{1}{2} \qquad \text{für alle } x \in [a,b].$$

Dann ist die Folge p_k auf $[a, b]$ beschränkt und besitzt dort eine konvergente Teilfolge. Wir nehmen also an, dass p_k auf $[a, b]$ gleichmäßig gegen $p \in \mathcal{P}_n$ konvergiert. Folglich konvergiert p_k auch auf $[-1, 1]$ gleichmäßig gegen p (Äquivalenz von Normen auf endlich-dimensionalen Vektorräumen). Für $q(x) \neq 0$ gilt nun

$$\frac{p(x)}{q(x)} = \lim_{k\to\infty} \frac{p_k(x)}{q_k(x)}.$$

Weil $\|r_k\| = \|p_k/q_k\|$ beschränkt ist, existiert ein $c > 0$ mit

$$|p(x)| \leq cq(x), \qquad \text{für alle } x \in [-1, 1],\ q(x) \neq 0.$$

Es folgt, dass jede l-fache Nullstelle von q auch l-fache Nullstelle von p ist. Also

$$r = \frac{p}{q} \in \mathcal{R}_{n,m}.$$

Betrachten wir für $\delta > 0$

$$M_\delta = \{x \in [-1, 1] : q(x) \geq \delta\},$$

so konvergiert $r_k = p_k/q_k$ auf jedem M_δ gleichmäßig gegen $r = p/q$. Es folgt

$$\|f - r\|_{M_\delta} = \lim_{k\to\infty} \|f - r_k\|_{M_\delta} \leq d(f, \mathcal{R}_{n,m}).$$

Da $f - r$ eine stetige Funktion ist, folgt

$$\|f - r\|_{[-1,1]} = \lim_{\delta\to 0} \|f - r\|_{M_\delta} \leq d(f, \mathcal{R}_{n,m}).$$

Also ist r beste Approximation an f. □

2.52 Satz: *Sei $f \in C[-1, 1]$,*

$$r_{n,m}(x) = \frac{p_n(x)}{q_m(x)} \in \mathcal{R}_{n,m},$$

sowie

$$-1 \leq x_1 < \ldots < x_{n+m-d+2} \leq 1,$$

wobei

$$d = \min\{n - \deg p_n, m - \deg q_m\}$$

der ***Defekt*** *von $r_{n,m}$ sei, mit*

$$\operatorname{sign}\left(f(x_i) - r_{n,m}(x_i)\right) = \sigma(-1)^i$$

($\sigma = \pm 1$) gilt. Dann gilt

$$\min_i |f(x_i) - r_{n,m}(x_i)| \leq d(f, \mathcal{R}_{n,m}).$$

2.53. Bemerkung: Im Fall $m = 0$ entsteht als Spezialfall die Abschätzung von de la Vallée Poussin (Satz 1.60) für Polynome. In diesem Fall ist der Defekt automatisch 0.

Beweis: Angenommen

$$\|f - \tilde{r}_{n,m}\| < \min_i |f(x_i) - r_{n,m}(x_i)|$$

für $\tilde{r}_{n,m} = \tilde{p}_n/\tilde{q}_m \in \mathcal{R}_{n,m}$. Dann hat

$$h_{n,m} = \frac{p_n}{q_m} - \frac{\tilde{p}_n}{\tilde{q}_m} = \frac{p_n\tilde{q}_m - \tilde{p}_n q_m}{q_m\tilde{q}_m}$$

$n + m - d + 1$ Nullstellen. Es gilt

$$\begin{aligned} \deg(p_n\tilde{q}_m - \tilde{p}_n q_m) &\leq \max(\deg p_n + \det \tilde{q}_m, \deg \tilde{p}_n + \deg q_m) \\ &\leq \max(\deg p_n + m, \deg q_m + n) \\ &= n + m + \max(\deg p_n - n, \deg q_m - m) \\ &= n + m - d. \end{aligned}$$

Dies ist ein Widerspruch. □

2.54 Satz: **(Chebyshev 1899)** *Die beste Approximation $r_{n,m}$ bezüglich des Raumes $\mathcal{R}_{n,m}$ an $f \in C[-1,1]$ ist eindeutig bestimmt, und durch eine Alternante der Länge $n + m - d + 2$ charakterisiert, wobei d der Defekt von $r_{n,m}$ ist.*

Beweis: Falls $r_{n,m}$ eine Alternante der Länge $n + m - d + 2$ hat und $\tilde{r}_{n,m}$ denselben Approximationsfehler liefert, so können wir wie im Beweis des vorigen Satzes ein Polynom h des Grades $n + m - d$ finden, mit

$$h(x_1) \leq 0, h(x_2) \geq 0, \ldots$$

Es folgt $h = 0$ aus Satz 1.58. Also ist $r_{n,m}$ eindeutig bestimmte beste Approximation.

Wir nehmen an, dass $r_{n,m} = p_n/q_m$ eine beste Approximation an f bezüglich $\mathcal{R}_{n,m}$ ist. Wir müssen zeigen, dass $f - r_{n,m}$ eine Alternante der Länge $n+m-d+2$ hat. Offenbar dürfen wir $d = 0$ annehmen, da $r_{n,m}$ auch beste Approximation an f bezüglich $\mathcal{R}n - d, m - d$ ist. Wir dürfen auch annehmen, dass p_n und q_m teilerfremd sind. Der Beweis folgt nun im Wesentlichen der Idee des Beweises des Alternantensatzes 2.38 für Polynome. Seien dazu wieder

$$-1 \leq E_1 < \ldots < E_l \leq 1$$

Mengen von Extremalpunkten, in denen $f - r_{n,m}$ jeweils abwechselndes Vorzeichen hat. Also

$$f(x) - r_{n,m}(x) = \sigma(-1)^i \|f - r_{n,m}\|_K \qquad \text{für alle } x \in E_i,\ i = 1, \ldots, l$$

mit einem festen $\sigma = \pm 1$. Angenommen, es wäre $l < n + m + 2$. Dann suchen wir zu $\epsilon > 0$ eine rationale Funktion $\tilde{r}_{n,m} = \tilde{p}_n/\tilde{q}_m$, so dass

$$\operatorname{sign}(r_{n,m}(x) - \tilde{r}_{n,m}(x)) = -\sigma(-1)^i \qquad \text{für alle } x \in E_i,\ i = 1, \ldots, l,$$

sowie

$$\|r_{n,m} - \tilde{r}_{n,m}\| < \epsilon$$

ist. Wählt man $\epsilon > 0$ klein genug, so ist

$$f - \tilde{r}_{n,m} = (f - r_{n,m}) + (r_{n,m} - \tilde{r}_{n,m})$$

in einer offenen Umgebung U von $\bigcup_i E_i$ eine bessere Approximation, ebenso wie in $[-1,1] \setminus U$. Es gilt

$$r_{n,m} - \tilde{r}_{n,m} = \frac{p_n \tilde{q}_m - \tilde{p}_n q_m}{q_m \tilde{q}_m} = \frac{p_n(\tilde{q}_m - q_m) - (\tilde{p}_n - p_n) q_m}{q_m \tilde{q}_m}.$$

Ist dann $\|\tilde{q}_m - q_m\|$ klein genug, so ist der Nenner positiv auf $[-1,1]$. Es genügt daher, Polynome $g_m \in \mathcal{P}_m$ und $h_n \in \mathcal{P}_n$ zu finden, so dass

$$\operatorname{sign}\left(p_n(x) g_m(x) - q_m(x) h_n(x)\right) = (-1)^i \qquad \text{für alle } x \in E_i,\ i = 1, \ldots, l$$

ist. Wegen $l \leq m + n + 1$ kann zwischen den E_i Nullstellen plazieren und erhält ein Polynom $s \in \mathcal{P}_{n+m}$ mit

$$\operatorname{sign} s(x) = (-1)^i \qquad \text{für alle } x \in E_i,\ i = 1, \ldots, l$$

Wir zeigen, dass $s = p_n g_m - q_m h_n$ mit $g_m \in \mathcal{P}_m$ und $h_n \in \mathcal{P}_n$ geschrieben werden kann.

Da der Defekt $d = 0$ ist, hat entweder der Zähler p_n oder der Nenner q_m vollen Grad. Falls q_m vollen Grad m hat, so betrachten wir die Funktionen

$$p_n(x)x, \ldots p_n(x)x^m, q_m(x), q_m(x)x, \ldots q_m(x)x^n.$$

Wir haben zu zeigen, dass diese $n + m + 1$ Funktionen eine Basis von $\mathcal{P}_{n+m}$ bilden. Sei nun eine Linearkombination der Funktionen gleich 0. Dann haben wir

$$p_n g_m + q_m h_n = 0$$

mit Polynomen $g_m \in \mathcal{P}_m$ und $h_n \in \mathcal{P}_n$, mit $g_m(0) = 0$. Da p_n und q_m teilerfremd sind, folgt, dass das Polynom q_m das Polynom g_m teilt. Da q_m vollen Grad m hat, muss $g_m = c q_m$ sein. Dies ist aber wegen $g_m(0) = 0$ und $q_m(0) \neq 0$ nur für $c = 0$, also $g_m = 0$ möglich. Es folgt $q_m h_n = 0$, also $h_n = 0$.

Falls der Zähler p_n vollen Grad hat, so betrachten wir analog

$$p_n(x), p_n(x)x, \ldots p_n(x)x^m, q_m(x)(x - x_0), \ldots q_m(x)(x - x_0)^n,$$

wobei $p_n(x_0) \neq 0$ sei. Die lineare Unabhängigkeit folgt ganz äquivalent. □

2.55. Bemerkung: Mit dem Euklidschen Algorithmus für Polynome ist es für Polynome p, q möglich,

$$pg + qh = d$$

mit Polynomen g, h zu schreiben, wobei d der größte gemeinsame Teiler von p und q ist. Wie haben im obigen Beweis diese Polynome auf andere Art und Weise konstruiert.

2.56. Bemerkung: Newmann hat gezeigt, dass rationale Approximation deutlich besser sein kann als polynomiale Approximation. Es gilt für $f(x) = |x|$

$$d(f, \mathcal{R}_{n,n}) \geq \frac{1}{2} e^{-9\sqrt{n}}.$$

Dies ist um Größenordnungen besser als die Approximationsgüte c/n, die man mit Polynomen erreicht.

2.6 Segmentapproximation

2.57. Definition: Sei $f \in C[-1,1]$. Wir approximieren f stückweise mit Polynomen. Dazu seien **Knoten**

$$-1 = x_0 \leq \ldots \leq x_k = 1$$

gegeben, und

$$E_{k,n}(x_0,\ldots,x_k)(f) = \max_{1 \leq i \leq k} d(f|_{[x_{i-1},x_i]}, \mathcal{P}_n).$$

Wir approximieren also einfach f auf allen k Teilintervallen durch Polynome. Nun versuchen, wir optimale Knoten zu finden. Wir suchen also

$$E_{k,n}(f) = \inf\{E_{k,n}(x_0,\ldots,x_k)(f) : -1 = x_0 \leq \ldots \leq k_k = 1\}.$$

Dieses Problem heißt **Segmentapproximation** mit freien Knoten.

2.58. Beispiel: Sei $n = 0$ und f monoton wachsend. Dann können wir Knoten wählen, so dass

$$f(x_i) - f(x_{i-1}) = \frac{f(1) - f(-1)}{k} \qquad \text{für alle } i = 1,\ldots,k$$

gilt. Es folgt

$$E_{k,n}(x_0,\ldots,x_k)(f) = \frac{f(1) - f(-1)}{2k}.$$

Denn

$$d(f, \mathcal{P}_0) = \frac{\max f - \min f}{2}$$

auf jedem Intervall $[a,b]$. Um zu zeigen, dass diese Knotenverteilung optimal ist, beachten wir, dass für jede Knotenverteilung

$$\begin{aligned} E_{k,n}(x_0,\ldots,x_k)(f) &= \max_i \frac{f(x_i) - f(xi-1)}{2} \\ &\geq \frac{1}{k} \sum_{i=1}^{k} \frac{f(x_i) - f(x_{i-1})}{2} \\ &= \frac{f(1) - f(-1)}{2k} \end{aligned}$$

gilt. Es ist daher optimal, die Fehler auf allen Teilintervallen gleich zu wählen.

2.59. Definition: Eine Knotenverteilung heißt **ausgegliche Knotenverteilung** (oder nivellierte Knotenverteilung), wenn

$$E_{k,n}(x_0,\ldots,x_k)(f) = d(f|[x_{i-1},x_i], \mathcal{P}_n)$$

für alle $i = 1,\ldots,n$ gilt.

2.60 Satz: *Es gilt für jede Knotenverteilung*

$$\min_{1 \leq i \leq k} d(f|_{[x_{i-1},x_i]}, \mathcal{P}_n) \leq E_{k,n}(f).$$

Eine ausgeglichene Knotenverteilung ist optimal.

Beweis: Seien x_i und $\tilde{x}_i$, $i = 0, \ldots, k$, zwei Knotenverteilungen, und

$$m = \max\{\nu : x_i \geq \tilde{x}_i \text{ für } 0 \leq i \leq \nu\}.$$

Falls $m = k$, so folgt $x_{k-1} \geq \tilde{x}_{k-1}$, und wegen $x_k = 1 = \tilde{x}_k$ folgt

$$[x_{k-1}, x_k] \subseteq [\tilde{x}_{k-1}, \tilde{x}_k],$$

Falls $m < k$ ist, so folgt $x_{m-1} \geq \tilde{x}_{m-1}$ und $x_m < \tilde{x}_m$. Wir erhalten

$$[x_{m-1}, x_m] \subseteq [\tilde{x}_{m-1}, \tilde{x}_m],$$

Da der Approximationsfehler auf größeren Mengen größer wird, erhalten wir

$$\min_{1 \leq i \leq k} d(f|_{[x_{i-1}, x_i]}, \mathcal{P}_n) \leq \max_{1 \leq i \leq k} d(f|_{[\tilde{x}_{i-1}, \tilde{x}_i]}, \mathcal{P}_n) = E_{k,n}(\tilde{x}_0, \ldots, \tilde{x}_k)(f).$$

Dies gilt für alle Knotenverteilungen $\tilde{x}_i$, $i = 0, \ldots, k$. Es folgt die Behauptung. Für eine ausgeglichene Knotenverteilung gilt

$$\begin{aligned} E_{k,n}(f) &\geq \min_{1 \leq i \leq k} d(f|_{[x_{i-1}, x_i]}, \mathcal{P}_n) \\ &= \max_{1 \leq i \leq k} d(f|_{[x_{i-1}, x_i]}, \mathcal{P}_n) = E_{k,n}(x_0, \ldots, x_k)(f) \geq E_{k,n}(f). \end{aligned}$$

Dies ist die Behauptung. □

2.61 Aufgabe: Zeigen Sie anhand der Funktion

$$f(x) = \begin{cases} x + 1/2, & -1 \leq x \leq -1/2, \\ 0, & -1/2 < x \leq 1/2, \\ x - 1/2, & 1/2 < x \leq 1, \end{cases}$$

für den Fall $n = 0$, dass nicht jede optimale Knotenverteilung ausgeglichen ist.

2.62. Bemerkung: Setzen wir im obigen Satz für $-1 \leq a \leq b \leq 1$

$$L(a, b) = d(f|_{[a,b]}, \mathcal{P}_n)$$

so haben wir lediglich die Eigenschaft

$$[a, b] \subseteq [\tilde{a}, \tilde{b}] \ \Rightarrow \ L(a, b) \leq L(\tilde{a}, \tilde{b})$$

benutzt. Solche Funktionale auf Intervallen nennen wir monoton. Bezüglich L ausgeglichene Knotenverteilungen sind also bezüglich L optimal, wenn L monoton ist.

2.63. Beispiel: Wir können Funktionale L_f der Form

$$L_f(a, b) = (b - a)^n \sup_{x \in [a,b]} |f^{(n+1)}(x)|$$

verwenden. L_f ist monoton und leicht auszuwerten. Da die besten Approximationen interpolieren, besteht aufgrund der Formel für den Interpolationsfehler Hoffnung dass $L_f(a,b)$ und $d(f|_{a,b}, \mathcal{P}_n)$ von der gleichen Größenordnung sind.

2.64. Definition: Falls L monoton auf Intervallen ist und stetig von a und b abhängt, sowie $L(a,a) = 0$ für alle a ist, so nennen wir L bezüglich der Segmentapproximation zulässig.

2.65 Satz: *Das Intervallfunktional*

$$L_f(a,b) = d(f|_{[a,b]}, \mathcal{P}_n)$$

ist bezüglich der Segmentapproximation zulässig.

Beweis: Der problematische Teil ist zu zeigen, dass $L_f(a,b)$ stetig von a und b abhängt. Im Polynomfall kann man aber einfach den Alternantensatz und die Abschätzung von de la Vallée Poussin verwenden. Sei dazu p die beste Approximation an f auf $[a,b]$. Dann gilt aufgrund dieser Sätze

$$L_f(a,b) \geq L_f(a,b-\epsilon) \geq L_f(a,b) - \omega_{f-p}(\epsilon).$$

Denn eine Alternante von $f-p$ auf $[a,b]$ liefert auf $[a,b-\epsilon]$ die obige untere Abschätzung des Approximationsfehlers. Trivialerweise gilt

$$L_f(a,b) \leq L_f(a,b+\epsilon) \leq L_f(a,b) + \omega_{f-p}(\epsilon).$$

In jedem Fall erhalten wir

$$\lim_{\tilde{b}\to b} L_f(a,\tilde{b}) = L_f(a,b).$$

Analog für den linken Intervallrand a. □

2.66 Aufgabe: Zeigen Sie diesen Satz allgemein für die Approximation bezüglich beliebiger Unterräume.

2.67 Satz: *Sei L ein bezüglich der Segmentapproximation zulässiges Funktional. Dann gibt eine ausgeglichene, optimale Knotenverteilung.*

Beweis: Per Induktion nach k. $k=1$ ist klar. Wenn die Behauptung für k gilt, so gibt es auf $[1,x]$ für $x \in [0,1]$ eine ausgeglichene Knotenmenge

$$-1 \leq x_0 \leq \ldots \leq x_k = x$$

mit k Intervallen. Denn der Satz gilt per Transformation auf allen Intervallen, nicht nur auf $[-1,1]$. Es gilt aufgrund von Satz 2.60 und der darauf folgenden Bemerkung

$$L(x_{k-1},x) = g(x) \leq g(x+\epsilon) \leq L(x_{k-1}, x+\epsilon).$$

Andererseits für den Fall $x_{k-1} < x$

$$L(x_{k-1}, x-\epsilon) \leq g(x-\epsilon) \leq g(x) = L(x_{k-1},x).$$

Also ist g insgesamt stetig. Falls $x_{k-1} = x$, so ist $g(x) = 0$ und daher ohnehin $g(x-\epsilon) = g(x) = 0$. Die Funktion

$$h(x) = g(x) - L(x,1)$$

ist also ebenfalls stetig. Es gilt $h(0) \leq 0$ und $h(1) \geq 0$. Also existiert ein $x \in [-1, 1]$ mit $g(x) = L(x, 1)$. Damit haben wir eine ausgeglichene Knotenverteilung erhalten. $\square$

2.68. Bemerkung: Die oben erwähnte Rekursion ist nicht gut für die Berechnung geeignet. Stattdessen wählen wir eine Startunterteilung

$$-1 \leq x_{0,0} < \ldots < x_{k,0} = 1.$$

Dann gilt

$$a_0 = \min_i L(x_{i-1}, x_i) \leq E_{k,n} \leq \max_i L(x_{i-1}, x_i) = b_0.$$

Wir setzen

$$m_0 = \frac{a_0 + b_0}{2}$$

und konstruieren eine Unterteilung

$$-1 = x_{0,1} \leq \ldots \leq x_{k,1} = 1$$

mit

$$L(x_{i-1,1}, x_{i,1}) = m_0 \qquad \text{für alle } i \text{ mit } x_{i,1} < 1.$$

Falls $x_{i,1} = 1$ ist für ein $i < k$, so gilt $m_0 \geq E_{k,n}$. Wir setzen

$$a_1 = a, \quad b_1 = m_0.$$

Im anderen Fall ist $x_{k-1,1} < 1$. Dann gilt

$$m_0 \leq E_{k,n} \leq L(x_{k-1,1}, x_{k,1}).$$

Wir setzen

$$a_1 = m_0, \quad b_1 = \min\{b_0, L(x_{k-1,1}, x_{k,1})\}.$$

Dies ergibt eine Intervallschachtelung, die gegen $E_{k,n}$ konvergiert.

Index