6 Charakterisierung der Bestapproximation in $L^p(D)$ und $C(D)$

Im Hilbertraum $L^2(D)$ wird das Problem der Bestapproximation durch die Orthogonalprojektion gelöst: falls $M \subseteq L^2(D)$ ein abgeschlossener Teilraum von $L^2(D)$ ist, so ist die metrische Projektion

$$\Pi_M : L^2(D) \longrightarrow M, \qquad f \longmapsto \Pi_M(f),$$

eindeutig charakterisiert durch die Orthogonalitätsrelation

$$\int_D (f(x) - \Pi_M(f)(x))\,\overline{v(x)}\,dx = 0 \qquad \text{für alle } v \in M.$$

Wir entwickeln nun ähnliche Charakterisierungen für die Bestapproximation in weiteren Banachräumen.

Ein allgemeines Kriterium lässt sich mit Hilfe des Dualraums (siehe Funktionalanalysis I) aufstellen. Sei hierzu $(X, \|.\|_X)$ ein (reeller oder komplexer) normierter Raum. Der Skalarkörper wird mit $\mathbb{K}$ bezeichnet (also $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$). Der **algebraische Dualraum** von X ist die Menge aller linearen Funktionale

$$\{\lambda : X \longrightarrow \mathbb{K} : \ \lambda \text{ linear}\}.$$

Das lineare Funktional λ ist genau dann stetig, wenn seine **Operatornorm**

$$\|\lambda\|_{\mathrm{op}} := \sup\{|\lambda(f)| : \ f \in X \text{ mit } \|f\|_X = 1\}$$

endlich ist. Die Menge aller stetigen linearen Funktionale bildet den (topologischen) **Dualraum**

$$X' = \{\lambda : X \longrightarrow \mathbb{K} : \ \lambda \text{ linear}, \ \|\lambda\|_{\mathrm{op}} < \infty\}.$$

In der Funktionalanalysis wird gezeigt, dass $(X', \|.\|_{X'})$ ein vollständiger normierter Raum, also ein Banachraum ist. Aus Konsistenzgründen wird die Operatornorm von λ daher mit $\|\lambda\|_{X'}$ bezeichnet.

6.1 Beispiel. (a) Als einfachstes Beispiel betrachten wir den Banachraum $X = \mathbb{C}^n$ mit der p-Norm

$$\|x\|_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p}, \qquad x = (x_1, \dots, x_n)^T \in \mathbb{C}^n,$$

für $1 \leq p < \infty$ bzw. $\|x\|_\infty = \max_{1\leq k\leq n} |x_k|$. Fassen wir X als die Menge der Spaltenvektoren auf, so können wir die linearen Funktionale als die Zeilenvektoren $a = (a_1, \ldots, a_n) \in (C^n)^*$ verstehen, die mit dem üblichen Matrixprodukt auf die Elemente von X wirken,

$$a(x) = \sum_{k=1}^{n} a_k x_k \in \mathbb{C}, \qquad x \in \mathbb{C}^n.$$

Weil X endlich-dimensional ist, sind alle linearen Funktionale stetig. Mit dem konjugierten Index $q = \frac{p}{p-1}$ für $1 < p < \infty$ (und $q = \infty$ für $p = 1$ sowie $q = 1$ für $p = \infty$) ergibt die Hölder-Ungleichung und eine kleine Zusatzüberlegung

$$\|a\|_{\mathrm{op}} = \|a\|_q, \qquad a \in (\mathbb{C}^n)^*.$$

Vergisst man den Unterschied zwischen Spalten- und Zeilenvektoren, so ist $X' = (\mathbb{C}^n, \|.\|_q)$ der Dualraum von $X = (\mathbb{C}^n, \|.\|_p)$.

(Anmnerkung: Wir haben hier den Fall, dass X' Dualraum zu X ist und auch umgekehrt wieder X Dualraum zu X' ist. Dies ist der besonders günstige Fall eines reflexiven Banachraums.)

(b) **Dualraum von $L^p(D)$ für $1 < p < \infty$:** Zu einer Lebesgue-messbaren Menge $D \subset \mathbb{R}^n$ wird wie üblich der Banachraum

$$L^p(D) = \left\{ f : D \longrightarrow \mathbb{K} : \ f \text{ messbar}, \ \ \|f\|_p = \left(\int_D |f(x)|^p \, dx \right)^{\frac{1}{p}} < \infty \right\}$$

definiert. (Funktionen werden identifiziert, wenn sie fast überall übereinstimmen). Wir setzen $q = \frac{p}{p-1} \in (1, \infty)$. Dann definiert jede Funktion $h \in L^q(D)$ ein **stetiges** lineares Funktional

$$\lambda_h : L^p(D) \longrightarrow \mathbb{K}, \qquad \lambda_h(f) = \int_D f(x) \, \overline{h(x)} \, dx. \tag{6.1}$$

Die Stetigkeit (=Beschränktheit) folgt aus der Hölderungleichung

$$|\lambda_h(f)| \leq \|f\|_p \|h\|_q,$$

und man zeigt genauer, dass $\|\lambda_h\|_{\mathrm{op}} = \|h\|_q$ die Operatornorm des Funktionals λ_h ist. Es gilt weiterhin, dass hierdurch sämtliche linearen stetigen Funktionale auf $L^p(D)$ erfasst sind, d.h.

$$(L^p(D))' = L^q(D), \qquad 1 < p < \infty, \ \ \frac{1}{p} + \frac{1}{q} = 1.$$

(Anmerkung: Die L^p-Räume mit $1 < p < \infty$ sind reflexiv.)

(c) **Dualraum von $L^1(D)$:** $D \subset \mathbb{R}^n$ sei Lebesgue-messbar. Der Banachraum

$$L^1(D) = \left\{ f : D \longrightarrow \mathbb{K} : \; f \text{ messbar}, \quad \|f\|_1 = \int_D |f(x)|\, dx < \infty \right\}$$

hat den Dualraum

$$L^\infty(D) = \left\{ f : D \longrightarrow \mathbb{K} : \; f \text{ messbar}, \quad \|f\|_\infty = \operatorname*{ess\,sup}_{x \in D} |f(x)| < \infty \right\}$$

der „wesentlich beschränkten“ Funktionen. Hierbei ist das wesentliche Supremum von f gegeben durch

$$\operatorname*{ess\,sup}_{x \in D} |f(x)| = \inf\{c \in \mathbb{R} : |f(x)| \leq c \text{ für fast alle } x \in D\}.$$

Die Werte von f in Lebesgue-Nullmengen spielen also bei der Bildung des wesentlichen Supremums keine Rolle. Beispiele von L^∞-Funktionen sind beschränkte Treppenfunktionen.

(Anmerkung: Der Dualraum von $L^\infty(D)$ ist größer als $L^1(D)$, daher ist $L^1(D)$ nicht reflexiv. Ein Beispiel eines stetigen linearen Funktionals auf $L^\infty(\mathbb{R})$ findet man im Buch “Functional Analysis: Introduction to Further Topics in Analysis” von Elias M. Stein und Rami Shakarchi, Seite 23; nachzulesen unter `http://press.princeton.edu/titles/9627.html`, Chapter 1.)

Der Dualraum eines Hilbertraums wird durch den Rieszschen Darstellungssatz beschrieben.

6.2 Satz. *H sei ein Hilbertraum über $\mathbb{K}$. Zu jedem stetigen Funktional $\lambda : H \longrightarrow \mathbb{K}$ existiert ein eindeutig bestimmtes Element $v_\lambda \in H$, so dass*

$$\lambda(f) = \langle f, v_\lambda \rangle \qquad \textit{für alle} \;\; f \in H \tag{6.2}$$

gilt. Weiterhin gilt $\|\lambda\|_{\text{op}} = \|v_\lambda\|_H$. Also ist der Dualraum H' isometrisch isomorph zu H: alle linearen stetigen Funktionale auf H haben die Form (6.2).

Wir betrachten weiterhin den Raum $C(K)$ der stetigen Funktionen auf einem Kompaktum K mit der Maximumsnorm

$$\|f\|_\infty = \max_{x \in K} |f(x)|.$$

6.3 Satz. *Der Dualraum von $C(K)$ ist der Banachraum $\mathcal{M}(K)$ der regulären (komplexen oder reellen) Borelmaße auf K.*

Zu diesem Begriff: Die Borel-σ-Algebra von K ist die kleinste σ-Algebra, die alle (relativ) offenen Teilmengen von K enthält. Ein reelles oder komplexes Borel-Maß μ ordnet jeder Menge $M \subset K$ dieser σ-Algebra einen komplexen oder reellen Wert zu, man schreibt

$$\mu(M) = \int\limits_M d\mu.$$

Dabei gelten die Axiome $\mu(\emptyset) = 0$ und die σ-Additivität. Das reelle oder komplexe Borel-Maß μ ist regulär, wenn seine Variation $|\mu|$ die Eigenschaft

$$|\mu|(M) = \inf\{|\mu|(O);\ O \supset M,\ O \text{ offen}\} = \sup\{|\mu|(L) :\ L \subseteq M,\ L \text{ kompakt}\}$$

erfüllt, also „von außen“ und „von innen“ regulär ist. Weil K kompakt ist, ist

$$\|\mu\| := |\mu|(K) = \int\limits_K d|\mu| < \infty.$$

Hierdurch wird eine Norm definiert, mit der $\mathcal{M}(K)$ ein Banachraum ist. Durch ein reguläres Borel-Maß μ ist das beschränkte lineare Funktional

$$\lambda_\mu : C(K) \longrightarrow \mathbb{K}, \qquad \lambda_\mu(f) = \int\limits_K f\, d\mu,$$

mit der Operatornorm $\|\lambda_\mu\|_{\mathrm{op}} = \|\mu\|$ definiert.
Die obige Aussage besagt, dass hiermit sämtliche Elemente des Dualraums $(C(K))'$ vorliegen.

6.4 Beispiel. a) Zu $h \in L^1(K)$ ist $\mu_h(M) = \int_M h(x)\, dx$ ein reguläres Borel-Maß mit $\|\mu\| = \|f\|_1$. Das zugehörige beschränkte lineare Funktional ist

$$\lambda_h(f) = \int_K f(x)h(x)\, dx, \qquad f \in C(K).$$

(Anmerkung: μ_h ist absolutstetig bzgl. des Lebesgue-Maßes, weil $\mu_h(M) = 0$ für jede Lebesgue-Nullmenge M gilt.)

b) Das Dirac-Maß zu $x \in K$ ist

$$\delta_x(M) = \begin{cases} 1, & \text{falls } x \in M, \\ 0, & \text{sonst.} \end{cases}$$

(δ_x ist nicht absolut-stetig bzgl. des Lebesgue-Maßes.)
Häufig verwendete Borel-Maße sind Linearkombiantionen $\mu = \sum_{=1}^n \alpha_k \delta_{x_k}$ mit paarweise verschiedenen Punkten $x_1, \ldots, x_n \in K$ und (reellen oder komplexen) Gewichten $\alpha_1, \ldots, \alpha_n$. Das zugehörige Funktional ist eine Linearkombination von Punktauswertungen

$$\lambda_\mu(f) = \sum_{k=1}^n \alpha_k f(x_k), \qquad f \in C(K).$$

Beispiele kennt man von den Quadraturformeln aus der Numerik. Man berechnet leicht

$$\|\mu\| = \sum_{k=1}^{n} |\alpha_k|,$$

indem man eine Funktion $f \in C(K)$ mit $f(x_k) = \frac{\overline{\alpha_k}}{|\alpha_k|}$ und $\|f\|_\infty = 1$ konstruiert (z.B. durch stückweise lineare Interpolation; hierzu ist es wichtig, dass die x_k paarweise verschieden sind).

c) Weitere Beispiele für $C[0,1]$: Das lineare Funktional

$$\lambda(f) = \sum_{k=1}^{\infty} \frac{1}{k^2}\, f\left(\frac{1}{k}\right)$$

ist beschränkt mit $\|\lambda\|_{\mathrm{op}} = \pi^2/6$. Hingegen ist

$$\lambda(f) = \sum_{k=1}^{\infty} \frac{(-1)^k}{k}\, f\left(\frac{1}{k}\right)$$

nicht beschränkt (und daher nicht wohldefiniert auf $C[0,1]$).

Nach diesen Vorüberlegungen aus der Funktionalanalysis kommen wir zum eigentlichen Anliegen, der Charakterisierung eines Proximums. Hierbei spielt weder die Eindeutigkeit noch die globale Existenz von Proxima eine Rolle.

6.5 Satz. *Ein normierter Raum X, ein abgeschlossener Teilraum $M \subseteq X$ sowie $f \in X \setminus M$ seien gegeben.*
Ein Element $v^* \in M$ ist genau dann Proximum an f, wenn es ein lineares stetiges Funktional $\lambda : X \longrightarrow \mathbb{K}$ gibt mit

(i) $\|\lambda\|_{X'} = 1$,

(ii) $\lambda(v) = 0$ *für alle* $v \in M$, *d.h.* λ *ist „orthogonal" zu* M,

(iii) $\lambda(f) = \lambda(f - v^*) = \|f - v^*\|_X$.

Beweis. 1. Die Bedingungen (i)–(iii) sind hinreichend: Für alle $v \in M$ gilt

$$\|f - v^*\| \overset{(iii)}{=} \lambda(f) \overset{(ii)}{=} \lambda(f - v) \overset{(i)}{\leq} \|f - v\|,$$

also ist

$$\|f - v^*\| = \inf_{v \in M} \|f - v\| = E_M(f).$$

2. Die Bedingungen (i)–(iii) sind notwendig: Das Element $v^* \in M$ sei Proximum zu $f \in X \setminus M$. Wir definieren den Teilraum

$$M_f = M \oplus \operatorname{span}\{f\} = \{v + \alpha f : \ v \in M, \ \alpha \in \mathbb{K}\} \subseteq X.$$

Durch die Festlegung

$$\widetilde{\lambda}(f) = \|f - v^*\|, \qquad \widetilde{\lambda}(v) = 0 \ \text{ für } v \in M,$$

wird ein lineares Funktional $\widetilde{\lambda} : M_f \longrightarrow \mathbb{K}$ definiert. Um seine Norm zu bestimmen, berechnen wir für beliebiges $v + \alpha f \in M_f$

$$|\widetilde{\lambda}(v + \alpha f)| = |\widetilde{\lambda}(\alpha f)| = |\alpha| \ \|f - v^*\| = |\alpha| \ E_M(f).$$

Die elementare Eigenschaft $E_M(\alpha f) = |\alpha| E_M(f)$ des Approximationsgrades ergibt

$$|\widetilde{\lambda}(v + \alpha f)| = E_M(\alpha f) \leq \|\alpha f - (-v)\| = \|v + \alpha f\|.$$

Also ist $\|\widetilde{\lambda}\| \leq 1$. Zusammen mit $\widetilde{\lambda}(f - v^*) = \widetilde{\lambda}(f) = \|f - v^*\|$ folgt $\|\widetilde{\lambda}\| = 1$.

Nach dem Satz von Hahn-Banach (Funktionalanalysis I) existiert ein stetiges lineares Funktional $\lambda : X \longrightarrow \mathbb{K}$ mit $\|\lambda\| = \|\widetilde{\lambda}\| = 1$ und $\lambda(g) = \widetilde{\lambda}(g)$ für alle $g \in M_f$. Dieses Funktional erfüllt (i)–(iii). □

Im ersten Beweisteil wird eine Form der “Orthogonalität” $\lambda \perp M$ ausgenutzt, die für Elemente $\lambda \in X'$ des Dualraums in Analogie zur Orthogonalität in Hilberträumen definiert wird. Diese Orthogonalität allein (also ohne die Eigenschaft (iii)) liefert eine untere Abschätzung des Approximationsgrades.

Folgerung: Es sei $\lambda \in X'$ mit $\|\lambda\| = 1$ und $\lambda(v) = 0$ für alle $v \in M$ gegeben. Dann gilt für alle $f \in X$ die Beziehung

$$E_M(f) \geq |\lambda(f)|.$$

Beweis. $E_M(f) = \inf\limits_{v \in M} \|f - v\| = \inf\limits_{v \in M} \|\lambda\| \ \|f - v\| \geq \inf\limits_{v \in M} |\lambda(f - v)| = |\lambda(f)|.$ □

In Abschnitt 6.1 behandeln wir die Charakterisierung der Bestapproximation im Raum $L^p(D)$, in Abschnitt 6.2 folgen entsprechende Aussagen für $C(K)$. Wir betrachten jeweils nur den Fall reeller Funktionen, setzen also $\mathbb{K} = \mathbb{R}$.

6.1 Charakterisierung der Bestapproximation in $L^p(D)$

Zunächst wird der Fall der strikt konvexen Räume $L^p(D)$ mit $1 < p < \infty$ behandelt.

6.6 Satz. *Es sei $1 < p < \infty$, $M \subseteq L^p(D)$ ein abgeschlossener Teilraum und $f \in L^p(D) \setminus M$. Das Element $v^* \in M$ ist genau dann Proximum an f, wenn*

$$\int_D |f(x) - v^*(x)|^{p-1} \operatorname{sign}(f(x) - v^*(x)) \; v(x)\, dx = 0 \tag{6.3}$$

für alle $v \in M$ gilt.

Beweis. 1. Zur Existenz des Integrals: Wir setzen $q = \frac{p}{p-1}$. Für beliebige $f, v^* \in L^p(D)$ ist die Funktion

$$h := |f - v^*|^{p-1} \operatorname{sign}(f - v^*)$$

ein Element von $L^q(D)$, denn

$$\|h\|_q = \left(\int_D |f - v^*|^{(p-1)q}\, dx \right)^{\frac{1}{q}} = \left(\int_D |f - v^*|^p\, dx \right)^{\frac{p-1}{p}} = \|f - v^*\|_p^{p-1}.$$

2. Die Bedingung (6.3) ist hinreichend: Das lineare Funktional

$$\lambda : L^p(D) \longrightarrow \mathbb{K}, \qquad \lambda(g) = \frac{1}{\|f - v^*\|_p^{p-1}} \int_D h(x) g(x)\, dx$$

ist stetig. Seine Norm ist

$$\|\lambda\|_{\text{op}} = \frac{1}{\|f - v^*\|_p^{p-1}} \; \|h\|_q = 1$$

und es gilt

$$\begin{aligned} \lambda(f - v^*) &= \frac{1}{\|f - v^*\|_p^{p-1}} \int_D h(x)(f(x) - v^*(x))\, dx \\ &= \frac{1}{\|f - v^*\|_p^{p-1}} \int_D |f(x) - v^*(x)|^p\, dx = \|f - v^*\|_p. \end{aligned}$$

Also erfüllt λ die Eigenschaften (i) und (iii) aus Satz 6.5. Falls zusätzlich die Bedingung (6.3) gilt, ist auch (ii) in Satz 6.5 erfüllt, und damit ist v^* ein Proximum an f.

3. Die Bedingung (6.3) ist notwendig: Wir verwenden ein Argument der Variationsrechnung.

Es sei v^* ein Proximum an f und $v \in M$ sei beliebig vorgegeben. Die reelle Funktion

$$k : \mathbb{R} \longrightarrow \mathbb{R}, \qquad k(\varepsilon) = \|f - v^* + \varepsilon v\|_p^p = \int_D |f(x) - v^*(x) + \varepsilon v(x)|^p\, dx$$

hat bei $\varepsilon = 0$ ihr absolutes Minimum. Da k stetig und differenzierbar ist (s. Analysis 2: Parameterintegrale), muss $k'(0) = 0$ gelten. Einfache Rechnung ergibt

$$k'(\varepsilon) = p \int_D |f(x) - v^*(x) + \varepsilon v(x)|^{p-1} \operatorname{sign}(f(x) - v^*(x) + \varepsilon v(x))\; v(x)\, dx,$$

und durch Einsetzen von $\varepsilon = 0$ folgt die Identität (6.3). $\square$

6.7 Bemerkung. Die Bedingung (6.3) wird in Anlehnung an den Fall $p = 2$ ebenfalls als **Orthogonalitätsrelation** bezeichnet.

Würde man die Orthogonalitätsrelation aus Satz 6.5 auf den Fall $p = 1$ erweitern, so würde sich

$$\int_D \operatorname{sign}(f(x) - v^*(x))\; v(x)\, dx = 0 \qquad \text{für alle } \; v \in M \subseteq L^1(D) \tag{6.4}$$

ergeben. Diese Bedingung ist allerdings zu stark, wie das folgende Beispiel zeigt.

6.8 Beispiel. Die Funktion

$$f : [0,1] \longrightarrow \mathbb{R}, \qquad f(x) = \begin{cases} 1 & \text{, für } 0 \le x < \frac{1}{2}, \\ 0 & \text{, für } \frac{1}{2} \le x \le 1, \end{cases}$$

liegt in $L^p(0,1)$ für alle $1 \le p \le \infty$. Wir betrachten die Bestapproximation durch konstante Funktionen, also $M = \mathcal{P}_0$.

(a) Für jeden Wert $1 < p < \infty$ besitzt f ein eindeutiges Proximum, nämlich die konstante Funktion $v^*(x) \equiv \frac{1}{2}$. Den Nachweis kann man direkt (s. Übung) oder mit der Orthogonalitätsrelation (6.3) führen: für jede konstante Funktion $v \equiv a$ ist

$$\int_0^1 |f(x) - v^*(x)|^{p-1} \operatorname{sign}(f(x) - v^*(x))\; v(x)\, dx = a \left(\frac{1}{2}\right)^{p-1} \left(\int_0^{\frac{1}{2}} dx - \int_{\frac{1}{2}}^1 dx \right) = 0.$$

(b) Für $p = 1$ ist jede konstante Funktion $v^*(x) \equiv c$ mit $c \in [0,1]$ Proximum an f (s. Übung). Für $v^* \equiv c$ mit $0 < c < 1$ gilt tatsächlich die Orthogonalitätsrelation

$$\int_0^1 \operatorname{sign}(f(x) - v^*(x))\; v(x)\, dx = a \left(\int_0^{\frac{1}{2}} dx - \int_{\frac{1}{2}}^1 dx \right) = 0.$$

Allerdings erhalten wir für das Proximum $v^* \equiv 0$

$$\int_0^1 \operatorname{sign}(f(x) - v^*(x))\; v(x)\, dx = a \int_0^{\frac{1}{2}} dx \neq 0,$$

sofern $v \in \mathcal{P}_0 \setminus \{0\}$ gewählt wird.

Wir zeigen nun, dass eine abgeschwächte Relation die Proxima in $L^1(D)$ charakterisiert. Eine besondere Rolle spielt dabei die Menge

$$A_0 = A_0(f - v^*) = \{x \in D : \ f(x) - v^*(x) = 0\} \tag{6.5}$$

der Nullstellen von $f - v^*$. (Genau genommen ist diese Menge nur bis auf eine Lebesgue-Nullmenge bestimmt, da ja die einzelne Punktauswertung einer L^1-Funktion gar nicht zulässig ist.) Wir sehen im obigen Beispiel bereits, dass die strenge Orthogonalitätsrelation (6.4) verletzt werden kann, wenn A_0 ein positives Maß hat.

6.9 Satz. *Es sei $M \subseteq L^1(D)$ ein abgeschlossener Teilraum und $f \in L^1(D) \setminus M$. Das Element $v^* \in M$ ist genau dann Proximum an f, wenn mit der Menge $A_0 = A_0(f - v^*)$ in* (6.5)

$$\left| \int\limits_{D \setminus A_0} \operatorname{sign}(f(x) - v^*(x)) \ v(x) \, dx \right| \le \int\limits_{A_0} |v(x)| \, dx \tag{6.6}$$

für alle $v \in M$ gilt.

Beweis. 1. Die Bedingung (6.6) ist hinreichend: Für alle $v \in M$ gilt

$$\|f - v^* + v\|_1 = \int\limits_{D \setminus A_0} |f(x) - v^*(x) + v(x)| \, dx + \int\limits_{A_0} |v(x)| \, dx.$$

Im ersten Integral verwenden wir die reelle Ungleichung

$$|a + b| \ge |a| + \operatorname{sign}(a) \, b, \qquad a, b \in \mathbb{R},$$

und erhalten insgesamt

$$\begin{aligned} \|f - v^* + v\|_1 \quad &\ge \quad \int\limits_{D \setminus A_0} |f(x) - v^*(x)| \, dx + \int\limits_{D \setminus A_0} \operatorname{sign}(f(x) - v^*(x)) \ v(x) \, dx \\ &\quad + \int\limits_{A_0} |v(x)| \, dx \\ &= \quad \|f - v^*\|_1 + \int\limits_{D \setminus A_0} \operatorname{sign}(f(x) - v^*(x)) \ v(x) \, dx + \int\limits_{A_0} |v(x)| \, dx \\ &\overset{(6.6)}{\ge} \quad \|f - v^*\|_1. \end{aligned}$$

Also ist v^* Proximum an f.

2. Die Bedingung (6.6) ist notwendig: Weil $L^\infty(D)$ der Dualraum von $L^1(D)$ ist, existiert nach Satz 5.1 eine Funktion $h \in L^\infty(D)$ mit

(i) $\|h\|_\infty = 1$,

(ii) $\displaystyle\int_D h(x)v(x)\,dx = 0$ für alle $v \in M$, und

(iii) $\displaystyle\int_D h(x)(f(x) - v^*(x))\,dx = \|f - v^*\|_1$.

Die letzte Identität ist äquivalent zu

$$\int_{D\setminus A_0} h(x)(f(x) - v^*(x))\,dx = \int_{D\setminus A_0} |f(x) - v^*(x)|\,dx.$$

Wegen $\|h\|_\infty = 1$ folgt hieraus sofort

$$h(x) = \operatorname{sign}(f(x) - v^*(x)) \qquad \text{für fast alle } \ x \in D \setminus A_0.$$

(Das „fast alle" lässt Nullmengen zu, in denen h andere Werte annimmt. Diese spielen für die Integration keine Rolle.) Weiter folgt mit (i) und (ii)

$$\begin{aligned} 0 \;&=\; \int_{A_0} h(x)v(x)\,dx + \int_{D\setminus A_0} \operatorname{sign}(f(x) - v^*(x))\; v(x)\,dx \\ &\begin{cases} \leq +\displaystyle\int_{A_0} |v(x)|\,dx + \int_{D\setminus A_0} \operatorname{sign}(f(x) - v^*(x))\; v(x)\,dx \\ \geq -\displaystyle\int_{A_0} |v(x)|\,dx + \int_{D\setminus A_0} \operatorname{sign}(f(x) - v^*(x))\; v(x)\,dx. \end{cases} \end{aligned}$$

Umstellung der so erhaltenen Ungleichungen ergibt

$$-\int_{A_0} |v(x)|\,dx \leq \int_{D\setminus A_0} \operatorname{sign}(f(x) - v^*(x))\; v(x)\,dx \leq \int_{A_0} |v(x)|\,dx.$$

Dies ist genau die Beziehung (6.6). □

6.2 Charakterisierung der Bestapproximation in $C(K)$

Die Menge K sei kompakt. Wir behandeln nun die Bestapproximation im Banachraum $(C(K), \|.\|_\infty)$ der stetigen und reellwertigen Funktionen. Wie bereits erwähnt ist der Dualraum gegeben durch den Raum der regulären reellen Borel-Maße auf K.
Für zwei Elemente f und v^* definieren wir die Menge

$$A_\infty = A_\infty(f - v^*) = \{x \in K : \ |f(x) - v^*(x)| = \|f - v^*\|_\infty\} \tag{6.7}$$

aller **Extremalpunkte** von $f - v^*$; dies sind alle Punkte, an denen $|f - v^*|$ sein absolutes Maximum annimmt. (Skizze!)
Wir teilen die Charakterisierung auf in zwei Teile. Die erste Aussage gilt für beliebige abgeschlossene Teilräume von $C(K)$, die zweite gilt nur im endlich-dimensionalen Fall.

6.10 Satz. *K sei kompakt, $M \subseteq C(K)$ sei ein abgeschlossener Teilraum und $f \in C(K) \setminus M$ sei gegeben. Für ein Element $v^* \in M$ sind äquivalent:*

(a) v^* ist Proximum an f.

(b) Es existiert ein reguläres reelles Borel-Maß μ mit

$$(i)\quad \|\mu\| = 1, \qquad (ii)\quad \int_K v\, d\mu = 0 \quad \text{für alle } v \in M, \qquad (iii)\quad \int_K f\, d\mu = \|f - v^*\|_\infty.$$

(c) Kolmogorov-Kriterium (1948): Für alle $v \in M$ gilt

$$\max_{x \in A_\infty(f - v^*)} (f(x) - v^*(x))\ v(x) \geq 0.$$

6.11 Bemerkung. Das Kolmogorov-Kriterium kann auch in den Formen

$$\max_{x \in A_\infty(f - v^*)} \operatorname{sign}(f(x) - v^*(x))\ v(x) \geq 0$$

oder

$$\max_{x \in A_\infty(f - v^*)} \operatorname{sign}[(f(x) - v^*(x))\ v(x)] \geq 0$$

angegeben werden. Dabei ist $\operatorname{sign}(a) = 0$ für $a = 0$ zu setzen. Man beachte auch die Beziehung

$$f(x) - v^*(x) = \sigma(x)\ \|f - v^*\|_\infty \qquad \text{für alle } \ x \in A_\infty(f - v^*),$$

wobei $\sigma(x) \in \{-1, 1\}$ geeignet zu wählen ist.

Beweis. Die Äquivalenz von (a) und (b) ist eine Wiederholung der Aussage von Satz 6.5 für den Fall $X = C(K)$. Wir zeigen nun, dass (a) äquivalent ist zu (c).

1. Sei v^* ein Proximum an $f \in C(K) \setminus M$. Weil M abgeschlossen ist, gilt $E := E_M(f) > 0$. Die Extremalmenge $A_\infty(f - v^*) = (f - v^*)^{-1}(\{-E, E\}) \subseteq K$ ist abgeschlossen (als Urbild der abgeschlossenen Menge $\{-E, E\}$), also ebenfalls kompakt. Wir führen nun die folgende Annahme zum Widerspruch: **Es existiert ein $v \in M$ mit**

$$\max_{x \in A_\infty(f - v^*)} (f(x) - v^*(x))\ v(x) < 0. \tag{6.8}$$

Als Widerspruch weisen wir nach, dass dann $\|f - v^* + \lambda v\|_\infty < E_M(f)$ für hinreichend kleines $\lambda > 0$ gilt. Dazu wählen wir $\varepsilon > 0$ mit

$$\max_{x \in A_\infty(f - v^*)} (f(x) - v^*(x))\ v(x) < -2\varepsilon$$

sowie eine offene Menge $U \supset A_\infty(f - v^*)$ mit

$$\max_{x \in U} (f(x) - v^*(x))\; v(x) < -\varepsilon.$$

Setzen wir $C := \|v\|_\infty$, so ist $C > 0$ wegen (6.8). Für alle $0 < \lambda < \frac{\varepsilon}{C^2}$ und alle $x \in U$ folgt

$$\begin{aligned} |f(x) - v^*(x) + \lambda v(x)|^2 &= |f(x) - v^*(x)|^2 + 2\lambda \underbrace{(f(x) - v^*(x))v(x)}_{<-\varepsilon} + \lambda^2 |v(x)|^2 \\ &< E^2 - 2\lambda\varepsilon + \lambda^2 C^2 < E^2 - 2\lambda\varepsilon + \lambda\varepsilon < E^2. \end{aligned}$$

Weiterhin ist $K \setminus U$ kompakt und disjunkt zur Extremalmenge $A_\infty(f - v^*)$. Also existiert $\delta > 0$ mit

$$\max_{x \in K \setminus U} |f(x) - v^*(x)| \le E - \delta.$$

Für alle $0 < \lambda < \frac{\delta}{2C}$ und alle $x \in K \setminus U$ folgt

$$|f(x) - v^*(x) + \lambda v(x)| \le |f(x) - v^*(x)| + \lambda |v(x)| < E - \delta + \frac{\delta}{2} < E.$$

Damit wäre $v^* - \lambda v$ eine bessere Approximation an f, ein Widerspruch.

2. Nun sei $v^* \in M$ so gegeben, dass die Kolmogorov-Bedingung (c) erfüllt ist. Zu beliebigem $v \in M$ wählen wir einen Punkt $x_0 \in A_\infty(f - v^*)$ so, dass

$$(f(x_0) - v^*(x_0))v(x_0) \ge 0$$

gilt. Dann erhalten wir (ähnlich wie in 1.)

$$|f(x_0) - v^*(x_0) + v(x_0)|^2 = \underbrace{|f(x_0) - v^*(x_0)|^2}_{= E^2} + 2 \underbrace{(f(x_0) - v^*(x_0))v(x_0)}_{\ge 0} + |v(x_0)|^2 \ge E^2.$$

Also ist v^* Proximum an f. □

Das Kolmogorov-Kriterium hat eine schöne geometrische Interpretation für einen endlich-dimensionalen Teilraum $M \subseteq C(K)$. Im Fall $\dim(M) = n$ ordnet man der Extremalmenge $A_\infty(f - v^*)$ eine gewisse Punktmenge $Z \subseteq \mathbb{R}^n$ zu und geht dann zur **konvexen Hülle** dieser Punktmenge über.

6.12 Definition. *Die* **konvexe Hülle** *einer nichtleeren Menge $Z \subseteq \mathbb{R}^n$ ist die Menge*

$$\operatorname{conv}(Z) = \left\{ y = \sum_{k=1}^{r} p_k z_k;\ r \in \mathbb{N},\ z_1, \dots, z_r \in Z,\ p_1, \dots, p_r \in [0,1] \text{ und } \sum_{k=1}^{r} p_k = 1 \right\}. \tag{6.9}$$

Mit anderen Worten: conv(Z) *ist die Menge aller* **Konvexkombinationen** *von Elementen aus Z. Dies ist gleichzeitig die kleinste konvexe Menge, die Z enthält.*

Eine Aussage zur Anzahl der Summanden in (6.9) gibt der wichtige Satz von Carathéodory.

6.13 Satz (Carathéodory)**.** *In Definition 6.12 darf $r \leq n+1$ verlangt werden; d.h. für jeden Punkt $y \in \operatorname{conv}(Z)$ existieren $r \leq n+1$ Punkte $z_1, \dots, z_r \in Z$ und Zahlen $p_1, \dots, p_r \in [0,1]$ mit $\sum\limits_{k=1}^{r} p_k = 1$ sowie*

$$y = \sum_{k=1}^{r} p_k z_k.$$

Auf diesem Hintergrund liefert der folgende Satz eine geometrische Charakterisierung der Bestapproximation in $C(K)$.

6.14 Satz (Rivlin-Shapiro)**.** *K sei kompakt, $M \subseteq C(K)$ sei ein endlich-dimensionaler Teilraum und $f \in C(K) \setminus M$ sei gegeben. Weiter sei $\{\varphi_1, \dots, \varphi_n\}$ eine Basis von M, $\dim(M) = n$. Für $v^* \in M$ sind äquivalent:*

(a) v^* ist Proximum an f.

(d) Es existieren $r \leq n+1$ paarweise verschiedene $x_1, \dots, x_r \in A_\infty(f - v^*)$ und Zahlen $p_1, \dots, p_r \in [0,1]$ mit $\sum\limits_{k=1}^{r} p_k = 1$, so dass

$$\sum_{k=1}^{r} p_k \operatorname{sign}(f(x_k) - v^*(x_k))\ v(x_k) = 0$$

für alle $v \in M$ gilt.

(e) Für $x \in A_\infty(f - v^*)$ definieren wir den Punkt

$$z(x) = \operatorname{sign}(f(x) - v^*(x)) \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_n(x) \end{pmatrix} \in \mathbb{R}^n.$$

Die konvexe Hülle der Menge

$$Z = \{z(x);\ x \in A_\infty(f - v^*)\} \subseteq \mathbb{R}^n$$

enthält den Nullpunkt.

6.15 Bemerkung. Mit Teil (d) lässt sich für den Fall $\dim(M) = n$ das lineare Funktional λ in Satz 6.5 beschreiben. Setzen wir $\sigma_k = \operatorname{sign}(f(x_k) - v^*(x_k)) \in \{-1, 1\}$ für $1 \leq k \leq r$, so definiert

$$\lambda(g) = \sum_{k=1}^{r} p_k\, \sigma_k\, g(x_k)$$

ein reguläres Borel-Maß mit den Eigenschaften

$$(i) \qquad \|\lambda\| = \sum_{k=1}^{r} |p_k \sigma_k| = \sum_{k=1}^{r} p_k = 1,$$

$$(ii) \qquad \lambda(v) = 0 \quad \text{für alle} \;\; v \in M,$$

$$(iii) \qquad \lambda(f) = \lambda(f - v^*) = \sum_{k=1}^{r} p_k \sigma_k (f(x_k) - v^*(x_k)) = \sum_{k=1}^{r} p_k \underbrace{|f(x_k) - v^*(x_k)|}_{= \|f - v^*\|_\infty} = \|f - v^*\|_\infty.$$

Beweis. Wir führen einen Ringschluss (e)$\Longrightarrow$ (d)$\Longrightarrow$ (a) $\Longrightarrow$ (e) durch, bei dem die Aussage (a) durch Aussage (c) in Satz 6.10 ersetzt wird.
Falls 0 ein Element der konvexen Hülle von Z ist, so ist 0 bereits eine Konvexkombination von $r \leq n+1$ Punkten aus Z (Carathéodory). Die zugehörigen Elemente aus $A_\infty(f - v^*)$ seien mit $x_1, \ldots, x_r$ und die Faktoren der Konvexkombination mit $p_1, \ldots, p_r$ bezeichnet. Lesen wir die Vektorgleichung

$$0 = \sum_{k=1}^{r} p_k z(x_k)$$

zeilenweise, so erhalten wir

$$0 = \sum_{k=1}^{r} p_k \operatorname{sign}(f(x_k) - v^*(x_k)) \; \varphi_i(x_k), \qquad 1 \leq i \leq n.$$

Dies ist äquivalent zu

$$0 = \sum_{k=1}^{r} p_k \operatorname{sign}(f(x_k) - v^*(x_k)) \; v(x_k), \qquad v \in M, \tag{6.10}$$

also folgt die Aussage (d) aus (e).
Weiterhin impliziert (6.10), dass für beliebiges $v \in M$ und mindestens ein k die Relation

$$\operatorname{sign}(f(x_k) - v^*(x_k)) \; v(x_k) \geq 0$$

gilt. Daher folgt die Aussage (c) in Satz 6.10 aus (d).
Schließlich ist zu zeigen, dass (c) in Satz 6.10 auch (e) impliziert. Dies führen wir anhand eines Widerspruchsbeweises durch. Falls 0 kein Element der konvexen Hülle der Menge Z ist, gibt es eine Hyperebene $H \subseteq \mathbb{R}^n$ mit $H \cap Z = \emptyset$; d.h. es existiert $c = (c_1, \ldots, c_n) \neq 0$ mit

$$\langle c, z \rangle < 0 \qquad \text{für alle} \;\; z \in Z.$$

Für alle $x \in A_\infty(f - v^*)$ ist also

$$\langle c, z(x) \rangle = \sum_{i=1}^{n} c_i \operatorname{sign}(f(x) - v^*(x)) \; \varphi_i(x) < 0.$$

Schreiben wir dies mit Hilfe der Linearkombination $v = \sum_i c_i \varphi_i \in M$, so lautet dies

$$\operatorname{sign}(f(x) - v^*(x))\; v(x) < 0 \qquad \text{für alle} \;\; x \in A_\infty(f - v^*),$$

ein Widerspruch zu (c) in Satz 6.10. □

Die Charakterisierung des Proximums in $M \subseteq C(K)$ mit Hilfe der Punktauswertung in Stellen der Extremalmenge $A_\infty(f - v^*)$ wird eine wichtige Rolle bei der numerischen Berechnung des Proximums spielen, die wir in Abschnitt 6.3.4 (Remez-Algorithmus) behandeln werden.

Zusammenfassung: Wir haben in Satz 6.6 eine abstrakte Charakterisierung von Proxima in normierten Räumen kennengelernt. Diese ließ sich für die Funktionenräume $L^p(D)$, $1 \leq p < \infty$, und $C(K)$ nochmals präzisieren.

6.3 Bestapproximation mit Haarschen Räumen

Wir setzen die Untersuchungen von Abschnitt 6.2 zur Bestapproximation in Teilräumen von $C(K)$, K kompakt, fort. Dabei betrachten wir den Fall reellwertiger Funktionen ($\mathbb{K} = \mathbb{R}$) und endlich-dimensionaler Teilräume

$$M = \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}, \qquad \dim(M) = n \in \mathbb{N}.$$

Die Funktionen $\varphi_i : K \longrightarrow \mathbb{R}$ bilden eine Basis von M.

Die Existenz von Proxima zu beliebigem $f \in C(K)$ wurde bereits in Kapitel 2 (Existenzsatz) geklärt. Die Eindeutigkeit des Proximums ist jedoch nicht immer gegeben, die Maximumsnorm ist ja nicht strikt konvex. Wir betrachten im ersten Teilabschnitt Beispiele Haarscher Räume und beweisen den Satz von Haar (Satz 6.22), dass $M \subseteq C(K)$ **genau dann** eine Eindeutigkeitsmenge ist, wenn M ein Haarscher Raum ist. Der zweite Teilabschnitt ist eher von technischer Natur. Er dient dem Beweis des Satzes von Krein (Satz 6.24), dass jeder Haarsche Raum eine strikt positive (also nullstellen-freie) Funktion enthält. Dies geht dann im letzten Teilabschnitt in den Beweis des Alternantensatzes von Chebyshev ein (Satz 6.33).

6.3.1 Haarsche Räume und der Eindeutigkeitssatz von Haar

In der Numerik I lernt man bereits den folgenden Begriff kennen.

6.16 Definition. *K sei kompakt.*

(a) *Ein n-dimensionaler Teilraum $M \subseteq C(K)$ heißt* **Haarscher Raum**, *wenn jedes $v \in M \setminus \{0\}$ höchstens $n-1$ Nullstellen besitzt.*

(b) *Linear unabhängige Funktionen* $\varphi_1, \ldots, \varphi_n \in C(K)$ *heißen* **Chebyshev-System** *oder* **Haarsches System**, *wenn der Raum* $M = \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}$ *ein Haarscher Raum ist.*

Nach der obigen Sprechweise ist jede Basis eines Haarschen Raums also ein Chebyshev-System.

6.17 Beispiel. (i) Es sei $K = [a, b] \subseteq \mathbb{R}$ ein kompaktes Intervall. Der Raum

$$\mathcal{P}_n = \operatorname{span}\{e_k : [a, b] \longrightarrow \mathbb{R};\ e_k(x) = x^k,\ 0 \leq k \leq n\}$$

ist ein Haarscher Raum der Dimension $n+1$, die Monome e_k, $0 \leq k \leq n$, bilden also ein Chebyshev-System. Dies ist im **Fundamentalsatz der Algebra** begründet: Jedes von 0 verschiedene Polynom vom Grad n hat genau n komplexe Nullstellen (unter Berücksichtigung der Vielfachheit). Als Konsequenz kann $p \not\equiv 0$ höchstens n paarweise verschiedene Nullstellen im Intervall $[a, b]$ besitzen.

(ii) Es sei $K = [a, b] \subseteq \mathbb{R}$ ein kompaktes Intervall. Zu n paarweise verschiedenen Zahlen $a_i \in \mathbb{R}$, $1 \leq i \leq n$, definieren wir die Funktionen $\varphi_i(x) = e^{a_i x}$ und den Raum

$$M_n = \operatorname{span}\{\varphi_i :\ 1 \leq i \leq n\} = \left\{ v = \sum_{k=1}^{n} c_k e^{a_i \cdot} :\ c_1, \ldots, c_n \in \mathbb{R} \right\}$$

von Exponentialsummen. Dieser Raum ist ein Haarscher Raum der Dimension n.

Beweis per Induktion nach n: Für $n = 1$ hat $v(x) = ce^{ax}$ mit $c \neq 0$ und beliegibem $a \in \mathbb{R}$, keine Nullstelle, also ist M_1 ein Haarscher Raum der Dimension 1.

Sind nun $a_1, \ldots, a_n \in \mathbb{R}$ paarweise verschieden, so betrachten wir eine Linearkombination

$$v = \sum_{i=1}^{n} c_i e^{a_i \cdot}, \qquad c_i \in \mathbb{R},\ \ (c_1, \ldots, c_n) \neq 0.$$

O.B.d.A. sei $c_n \neq 0$. Die hieraus gebildete Funktion g mit

$$g(x) = \left(e^{-a_1 x} v(x)\right)' = \left(c_1 + \sum_{i=2}^{n} c_i e^{(a_i - a_1)x}\right)' = \sum_{i=2}^{n} c_i (a_i - a_1) e^{(a_i - a_1)x}$$

liegt im Haarschen Raum M_{n-1}, der von $\psi_i(x) = e^{(a_i - a_1)x}$ mit $2 \leq i \leq n$ aufgespannt wird. Die hierbei auftretenden Exponenten $a_i - a_1$, $2 \leq i \leq n$, sind paarweise verschieden und der Koeffizientenvektor $(c_2(a_2 - a_1), \ldots, c_n(a_n - a_1)) \in \mathbb{R}^{n-1}$ ist ungleich 0 wegen $c_n \neq 0$. Nach der Induktionsvoraussetzung hat g höchstens $n - 2$ Nullstellen. Mit dem Satz von Rolle folgt, dass $e^{-a_1 \cdot} v$ höchstens $n - 1$ Nullstellen hat, und dies gilt auch für v selbst.

(iii) Es sei $K = \mathbb{T} = \mathbb{R}/\mathbb{Z}$ der Definitionsbereich der 1-periodischen Funktionen. Wir verwenden reelle Basisfunktionen $c_k : \mathbb{T} \to \mathbb{R}$ und $s_k : \mathbb{T} \to \mathbb{R}$ mit $c_k(x) = \cos(2\pi kx)$, $s_k(x) = \sin(2\pi kx)$, $k \in \mathbb{N}_0$. Der Vektorraum

$$\mathcal{T}_n = \operatorname{span}(\{c_k;\ 0 \leq k \leq n\} \cup \{s_k :\ 1 \leq k \leq n\}$$

der reellen trigonometrischen Polynome vom Grad n ist ein Haarscher Raum der Dimension $2n + 1$, seine Basisfunktionen c_k, s_k bilden ein Chebyshev-System, vgl. Satz 4.16.

Auch wenn es momentan seltsam erscheint, so spielen Chebyshev-Systeme auf **endlichen Punktmengen** K eine Rolle bei der Suche nach Proxima. Deshalb stellen wir die folgenden Überlegungen an.

6.18 Bemerkung. Falls der Definitionsbereich K endlich ist, und zwar genau $\#K = n$ Elemente besitzt, so ist jede Funktion $f : K \longrightarrow \mathbb{R}$ stetig; also ist $C(K)$ isomorph zum $\mathbb{R}^n$ und insbesondere gilt $\dim(C(K)) = n$. Jede Funktion $f \in C(K) \setminus \{0\}$ verschwindet an höchstens $n - 1$ Stellen, also ist $C(K)$ selbst ein Haarscher Raum der Dimension n. Die Haarsche Bedingung für Räume der Dimension $\geq n$ ist also nur interessant, wenn der Definitionsbereich K mindestens $n + 1$ Punkte enthält.

Zur Charakterisierung von Chebyshev-Systemen verwendet man bereits in der Numerik I die sog. **Vandermonde-Determinante**. Für Funktionen $\varphi_1, \ldots, \varphi_n \in C(K)$ und Punkte $x_1, \ldots, x_n \in K$ setzen wir

$$D\begin{pmatrix} \varphi_1, \ldots, \varphi_n \\ x_1, \ldots, x_n \end{pmatrix} := \det \begin{pmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \vdots & \vdots & & \vdots \\ \varphi_1(x_n) & \varphi_2(x_n) & \cdots & \varphi_n(x_n) \end{pmatrix}. \tag{6.11}$$

6.19 Lemma. *Die Funktionen $\varphi_1, \ldots, \varphi_n \in C(K)$ seien linear unabhängig. Sie bilden genau dann ein Chebyshev-System, wenn für beliebige paarweise verschiedene Punkte $x_1, \ldots, x_n \in K$ gilt:*

$$D\begin{pmatrix} \varphi_1, \ldots, \varphi_n \\ x_1, \ldots, x_n \end{pmatrix} \neq 0.$$

Beweis. Die Existenz einer Linearkombination

$$v = \sum_{k=1}^{n} c_k \varphi_k, \qquad (c_1, \ldots, c_n) \neq 0,$$

die an n paarweise verschiedenen Stellen $x_1, \ldots, x_n \in K$ verschwindet, ist äquivalent dazu, dass die zugehörige Vandermonde-Matrix einen nichttrivialen Kern besitzt, also

$$D\begin{pmatrix} \varphi_1, \ldots, \varphi_n \\ x_1, \ldots, x_n \end{pmatrix} = 0$$

erfüllt ist. $\square$

6.20 Bemerkung. Chebyshev-Systeme der Dimension n sind also dadurch charakterisiert, dass sie die eindeutige und universelle Lösbarkeit des Lagrange-Interpolationsproblems zu **beliebigen** paarweise verschiedenen Knoten $x_1, \ldots, x_n \in K$ garantieren.

Alfréd Haar [4] gibt in seiner Arbeit von 1918 eine weitere Charakterisierung Haarscher Räume $M \subseteq C(K)$: Dies sind genau die Eindeutigkeitsmengen der besten Approximation in $C(K)$. Bevor wir seinen Satz beweisen, benötigen wir eine Aussage über die Mindestanzahl der Punkte der Extremalmenge

$$A_\infty(f - v^*).$$

Diese Anzahl wird noch prägnanter im Alternantensatz 6.33 auftreten.

6.21 Lemma. *K sei kompakt und es gelte $\#K \geq n+1$. Weiter sei $M \subseteq C(K)$ ein Haarscher Raum der Dimension n. Falls $v^* \in M$ Proximum an $f \in C(K)$ ist, so gilt*

$$\#A_\infty(f - v^*) \geq n + 1. \tag{6.12}$$

Beweis. Wir nehmen an, dass A_∞ nur $s \leq n$ Elemente $x_1, \ldots, x_s \in K$ enthält. Dann ist $K \setminus A_\infty \neq \emptyset$ und hieraus folgt

$$E_M(f) = \|f - v^*\|_\infty > 0.$$

Wir wählen zusätzliche Punkte $x_{s+1}, \ldots, x_n \in K \setminus A_\infty$. Dann hat das Lagrange-Interpolationsproblem

$$v(x_k) = -\operatorname{sign}(f(x_k) - v^*(x_k)), \qquad 1 \leq k \leq n,$$

eine eindeutige Lösung $v \in M$. Für diese Funktion v berechnen wir leicht

$$\max_{x \in A_\infty(f - v^*)} (f(x) - v^*(x))\, v(x) = \max_{1 \leq k \leq s} (f(x_k) - v^*(x_k))\, v(x_k) = -E_M(f) < 0.$$

Dies ist ein Widerspruch zum Kolmogorov-Kriterium in Satz 6.10(c). □

6.22 Satz (Haar 1918)**.** *K sei kompakt und es gelte $\#K \geq n$. Ein n-dimensionaler Teilraum $M \subseteq C(K)$ ist genau dann ein Haarscher Raum, wenn zu jedem $f \in C(K)$ genau ein Proximum $v^* \in M$ existiert.*

[4] geboren am 11. Oktober 1885 in Budapest; gestorben am 16. März 1933 in Szeged, ungarischer Mathematiker. Haar leistete Beiträge zur Approximationstheorie, Maßtheorie (Haar-Maß auf lokalkompakten Gruppen) und partiellen Differentialgleichungen. Nach ihm wurde auch das Haar-Wavelet (s. Kap. 4) benannt.

Beweis. Wir setzen $n := \dim(M)$ und geben eine Basis $\{\varphi_1, \ldots, \varphi_n\}$ von M vor. Der Fall $\#K = n$ ist wegen $M = C(K)$ klar, siehe Bemerkung 6.18. Also brauchen wir nur den Fall $\#K \geq n + 1$ (inklusive unendlicher kompakter Mengen K) zu betrachten.

1. Die Eindeutigkeit des Proximums ist notwendig: M sei ein Haarscher Raum und $v_1, v_2 \in M$ seien Proxima zu $f \in C(K)$. Um $v_1 = v_2$ zu beweisen, genügt es zu zeigen, dass $v_1 - v_2$ mindestens n Nullstellen besitzt.

   Weil die Menge $\Pi_M(f)$ der Proxima konvex ist, ist auch $v := \frac{1}{2}(v_1 + v_2)$ Proximum an f. Die zugehörige Extremalmenge $A_\infty(f - v)$ hat mindestens $n + 1$ Elemente (Lemma 6.21), und für jedes $x \in A_\infty(f - v)$ gilt

   $$E_M(f) = |f(x) - v(x)| \leq \frac{1}{2}\Bigg(\underbrace{|f(x) - v_1(x)|}_{\leq E_m(f)} + \underbrace{|f(x) - v_2(x)|}_{\leq E_M(f)}\Bigg) \leq E_M(f).$$

   In den beiden Abschätzungen muss also Gleichheit gelten, dies ist nur für

   $$f(x) - v_1(x) = f(x) - v_2(x) = \sigma(x) E_M(f) \qquad \text{mit } \ \sigma(x) \in \{-1, 1\}$$

   möglich. Also gilt $v_1(x) = v_2(x)$ für alle $x \in A_\infty(f - v)$ und hieraus folgt $v_1 \equiv v_2$.

2. Die Eindeutigkeit des Proximums ist hinreichend: Wir treffen die Annahme, dass die Basis $\{\varphi_1, \ldots, \varphi_n\}$ von M kein Chebyshev-System ist. Nun wollen wir ein $f \in C(K)$ konstruieren, das mehrere Proxima in M besitzt.

   Nach Lemma 6.19 existieren paarweise verschiedene Punkte $x_1, \ldots, x_n \in K$ mit

   $$D\begin{pmatrix} \varphi_1, \ldots, \varphi_n \\ x_1, \ldots, x_n \end{pmatrix} = 0. \tag{6.13}$$

2.a Konstruiere $v_1 \in M$ mit $\|v_1\|_\infty = 1$ und $v_1(x_k) = 0$, $1 \leq k \leq n$:
Gleichung (6.13) beschreibt die lineare Abhängigkeit der Spalten der Matrix in (6.11). Also existiert ein Vektor $(a_1, \ldots, a_n) \neq 0$ mit

$$\sum_{i=1}^{n} a_i \varphi_i(x_k) = 0, \qquad 1 \leq k \leq n. \tag{6.14}$$

O.B.d.A. wird $(a_1, \ldots, a_n)$ so gewählt, dass

$$v_1 := \sum_{i=1}^{n} a_i \varphi_i \in M \quad \text{die Bedingung} \quad \|v_1\|_\infty = 1$$

erfüllt. Also löst v_1 das homogene Interpolationsproblem $v_1(x_k) = 0$ für $1 \leq k \leq n$.

2.b Konstruiere ein lineares Funktional λ wie in Satz 6.5(b):
Ebenso erlaubt die lineare Abhängigkeit der Zeilen in (6.11) die Wahl von $(c_1, \ldots, c_n) \neq 0$ mit

$$\sum_{k=1}^{n} c_k \varphi_i(x_k) = 0, \qquad 1 \leq i \leq n.$$

O.B.d.A. wird $\sum\limits_{k=1}^{n} |c_k| = 1$ gewählt. Dann hat das lineare Funktional

$$\lambda : C(K) \longrightarrow \mathbb{R}, \qquad \lambda(g) := \sum_{k=1}^{n} c_k g(x_k) \tag{6.15}$$

die Norm 1 und erfüllt $\lambda(\varphi_i) = 0$ für alle $1 \leq i \leq n$, also auch $\lambda(v) = 0$ für alle $v \in M$. Dies sind die Eigenschaften (i) und (ii) des Funktionals in Satz 6.5(b).

2.c Konstruiere $f \in C(K)$ mit unendlich vielen Proxima:
Wir wählen (mit dem Satz von Tietze aus der Topologie) $f_1 \in C(K)$ mit

$$\|f_1\|_\infty = 1, \qquad f_1(x_k) = \operatorname{sign} c_k, \quad 1 \leq k \leq n.$$

Dann setzen wir

$$f(x) := (1 - |v_1(x)|)\ f_1(x),$$

also ist

$$|f(x_k)| = |f_1(x_k)| = 1, \quad 1 \leq k \leq n,$$

und

$$|f(x)| = \underbrace{(1 - |v_1(x)|)}_{\geq 0} |f_1(x)| \leq 1, \quad x \in K.$$

Nun zeigen wir mit dem Funktional λ und Satz 6.5: **Jede Funktion** $v_\alpha = \alpha v_1 \in M$ **mit** $0 \leq \alpha \leq 1$ **ist Proximum an** f**.** (Insbesondere ist also die Nullfunktion ein Proximum.)
Die Eigenschaften (i) und (ii) in Satz 6.5(b) wurden bereits in der Konstruktion 2.b von λ realisiert. Für die dritte Eigenschaft

(iii) $\lambda(f) = \|f - v_\alpha\|_\infty$

in Satz 6.5(b) berechnen wir zuerst

$$\lambda(f) = \sum_{k=1}^{n} c_k f(x_k) = \sum_{k=1}^{n} c_k f_1(x_k) = \sum_{k=1}^{n} |c_k| = 1. \tag{6.16}$$

Die Norm $\|f - v_\alpha\|_\infty = 1$ erhalten wir mit

$$|f(x) - v_\alpha(x)| \leq (1 - |v_1(x)|) \underbrace{|f_1(x)|}_{\leq 1} + \alpha|v_1(x)| \leq 1 - |v_1(x)| + \alpha|v_1(x)| \leq 1$$

und

$$\max_{1\le k\le n} |f(x_k) - \underbrace{v_\alpha(x_k)}_{=0}| = \max_{1\le k\le n} |f_1(x_k)| = 1.$$

Damit ist auch die Eigenschaft (iii) von λ in Satz 6.5(b) gezeigt. Also hat f unendlich viele Proxima $v_\alpha \in M$. □

Die Haarsche Bedingung ist eine sehr starke Einschränkung an den Teilraum M. Man beachte, dass ja die Elemente von M schon durch die Werte an endlich vielen Stellen eindeutig bestimmt sind (Interpolationseigenschaft). Dies liefert sogar eine noch viel stärkere Einschränkung, die 1956 von Mairhuber und (unabhängig) 1958 von Curtis bewiesen wurde.

6.23 Satz (Mairhuber, Curtis)**.** *Die Menge K sei kompakt. Falls ein Haarscher Raum $M \subseteq C(K)$ der Dimension ≥ 2 existiert, so ist K homöomorph zu einer Teilmenge von $\mathbb{R}$ oder $\mathbb{T} = \mathbb{R}/\mathbb{Z}$.*

Dieser Satz hat einen schönen geometrischen Beweis. Im wesentlichen Schritt zeigt man, dass K keine Teilmenge der Form eines „Verschiebebahnhofs“ enthalten kann: es gibt keine stetige Abbildung

$$\gamma : [0,1] \longrightarrow K \times K, \qquad \gamma(t) = (x_1(t), x_2(t)),$$

mit $x_1(t) \ne x_2(t)$ für alle $t \in [0,1]$ und $x_1(1) = x_2(0)$, $x_2(1) = x_1(0)$. Denn sonst wäre auf dem Weg von

$$d(0) = D\begin{pmatrix} \varphi_1, \dots, \varphi_n \\ x_1(0), x_2(0), x_3, \dots, x_n \end{pmatrix}$$

nach

$$d(1) = D\begin{pmatrix} \varphi_1, \dots, \varphi_n \\ x_1(1), x_2(1), x_3, \dots, x_n \end{pmatrix} = -d(0)$$

eine Vandermonde-Determinante unzulässigerweise 0. (Hierbei sind feste Punkte $x_3, \dots, x_n$ geeignet zu wählen.)

6.3.2 Existenz positiver Funktionen

Wir haben in den bisherigen Beispielen Haarsche Räume betrachtet, die jeweils Funktionen ohne Nullstellen enthalten (z.B. die konstanten Funktionen in $\mathcal{P}_n$ und $\mathcal{T}_n$ oder die Funktionen e^{ax} bei den Exponentialsummen). Es kostet einige Mühe zu zeigen, dass jeder Haarsche Raum $M \subseteq C[a,b]$ oder $C(\mathbb{T})$ eine strikt positive Funktion enthält.

6.24 Satz (Krein)**.** *Es sei $K = [a,b]$ oder $K = \mathbb{T}$. Jeder Haarsche Raum $M \subseteq C(K)$ reeller Funktionen enthält eine strikt positive Funktion.*

Um dies einzusehen, verwenden wir die Notation einfacher und doppelter Nullstellen.

6.25 Definition. *Es sei $K \subseteq \mathbb{R}$ kompakt oder $K = \mathbb{T}$ und $f \in C(K)$.*

*(a) Ein Punkt $\xi \in K$ mit $f(\xi) = 0$ heißt **einfache** Nullstelle von f, wenn entweder ξ ein Randpunkt von K ist oder f in ξ das Vorzeichen wechselt, d.h. es gilt*

$$f(\xi - t)f(\xi + t) < 0 \qquad \text{für alle } t \in [\xi - \delta, \xi + \delta].$$

(b) Ein innerer Punkt $\xi \in K$ mit $f(\xi) = 0$ heißt **doppelte** *Nullstelle von f, wenn*

$$f(\xi - t)f(\xi + t) > 0 \qquad \text{für alle } t \in [\xi - \delta, \xi + \delta] \text{ und } \delta > 0.$$

gilt.

Die Haarsche Bedingung lässt sich nun formulieren, indem man die Vielfachheit der Nullstellen in die Zählung einbezieht.

6.26 Lemma. *Es sei $K \subseteq \mathbb{R}$ kompakt oder $K = \mathbb{T}$. Der Raum $M \subseteq C(K)$ mit $\dim(M) = n$ ist genau dann ein Haarscher Raum (gemäß Definition 6.16), wenn für jedes $v \in M \setminus \{0\}$ die Anzahl $j \in \mathbb{N}_0$ der einfachen Nullstellen und die Anzahl $k \in \mathbb{N}_0$ der doppelten Nullstellen die Ungleichung*

$$j + 2k \leq n - 1$$

erfüllt; mit anderen Worten: jedes Element $v \in M \setminus \{0\}$ hat höchstens $n - 1$ Nullstellen unter Berücksichtigung der Vielfachheit.

Beweis. Die neue Bedingung impliziert die Haarsche Bedingung in Definition 6.16. Also ist nur zu zeigen, dass die Bedingung in Definition 6.16 auch die neue Bedingung impliziert. Dazu nehmen wir an, ein $v \in M \setminus \{0\}$ habe

$j \geq 0$ einfache Nullstellen $\xi_1, \ldots, \xi_j \in K$,

$k \geq 0$ doppelte Nullstellen $\eta_1, \ldots, \eta_k \in K$,

und es gelte

$$j + 2k \geq n.$$

Die Menge $A_0(v) = \{\xi_1, \ldots, \xi_j\} \cup \{\eta_1, \ldots, \eta_k\}$ ist die Menge aller Nullstellen von v. Die Bedingung in Definition 6.16 verlangt

$$\#A_0(v) = j + k \leq n - 1,$$

insbesondere folgt $k \geq 1$. Um jede doppelte Nullstelle η_i wählen wir ein Intervall $[\eta_i - \delta_i, \eta_i + \delta_i] \subseteq K$ mit $\delta_i > 0$, das keine weitere Nullstelle von v enthält. Wir setzen

$$c_i := \text{sign } v(\eta_i - \delta_i) = \text{sign } v(\eta_i + \delta_i)$$

und

$$C := \min_{1 \leq i \leq k} \{|v(\eta_i - \delta_i)|, \ |v(\eta_i + \delta_i)|\}.$$

Durch Hinzunahme beliebiger Punkte

$$\theta_1, \dots, \theta_{n-j-k} \in K \setminus A_0(v)$$

lässt sich die Interpolationsfunktion $q \in M$ mit

$$q(\eta_i) = c_i, \ \ 1 \leq i \leq k, \qquad q(\xi_i) = 0, \ \ 1 \leq i \leq j, \qquad q(\theta_i) = 0, \ \ 1 \leq i \leq n - j - k,$$

bestimmen. Die Funktion $v_\alpha = v - \alpha q \in M$ hat für $0 < \alpha < \frac{C}{\|q\|_\infty}$ die Funktionswerte

$$v_\alpha(\xi_i) = 0, \ \ 1 \leq i \leq j, \qquad v_\alpha(\eta_i) = -\alpha c_i \neq 0, \ \ 1 \leq i \leq k,$$

und je zwei Vorzeichenwechsel

$$c_i = \text{sign}\, v_\alpha(\eta_i - \delta_i) = -\text{sign}\, v_\alpha(\eta_i) = \text{sign}\, v_\alpha(\eta_i + \delta_i).$$

Dies ergibt je zwei Nullstellen von v_α im Intervall $[\eta_i - \delta_i, \eta_i + \delta_i]$. Damit hätte v_α mindestens $j + 2k \geq n$ Nullstellen, ein Widerspruch zur Haarschen Bedingung in Definition 6.16. □

Es gibt einen Trick, eine Funktion $v \in M$ mit fest vorgegebener Vorzeichenverteilung zu konstruieren. Sei dazu $M \subseteq C(K)$, mit $K = [a, b]$ oder $K = \mathbb{T}$, ein Haarscher Raum der Dimension n. Zu gegebenen

$$x_1 < x_2 < \ldots < x_{n-1}$$

definieren wir die Funktion

$$v(x) = D\begin{pmatrix} \varphi_1, \dots, \varphi_n \\ x_1, \dots, x_{n-1}, x \end{pmatrix} = \det \begin{pmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \vdots & \vdots & & \vdots \\ \varphi_1(x_{n-1}) & \varphi_2(x_{n-1}) & \cdots & \varphi_n(x_{n-1}) \\ \varphi_1(x) & \varphi_2(x) & \cdots & \varphi_n(x) \end{pmatrix}. \tag{6.17}$$

Die Entwicklung der Determinante nach der letzten Zeile zeigt, dass $v \in M$ gilt. Per Konstruktion ist $v(x_i) = 0$ für alle $1 \leq i \leq n - 1$ und $v(x) \neq 0$ für alle anderen $x \in K \setminus \{x_1, \dots, x_{n-1}\}$. Nach Lemma 6.26 muss in jedem der Punkte x_i, der im Inneren von K liegt, ein Vorzeichenwechsel von v vorliegen. Also hat v in den Intervallen

$$(x_1, x_2), \ \ (x_2, x_3), \dots \ \ (x_{n-2}, x_{n-1})$$

abwechselndes Vorzeichen, d.h.

$$(-1)^{i+1}\sigma\ v(x) > 0 \quad \text{für alle } \ x \in (x_i, x_{i+1}), \qquad 1 \le i \le n-2,$$

wobei σ das Vorzeichen von v im Intervall (x_1, x_2) angibt. Hinzu kommen noch im Fall $K = [a, b]$

$$\begin{aligned}
-\sigma\ v(x) &> 0 && \text{für alle } \ x \in (a, x_1), && \text{falls } x_1 > a,\\
(-1)^n \sigma\ v(x) &> 0 && \text{für alle } \ x \in (x_{n-1}, b), && \text{falls } x_{n-1} < b,
\end{aligned}$$

und im Fall $K = \mathbb{T}$ (wegen des Zusammenklebens der Intervallenden)

$$-\sigma\ v(x) > 0 \qquad \text{für alle } \ x \in (0, x_1) \cup (x_{n-1}, 1).$$

Eine kurze Beobachtung lässt sich direkt hieran anschließen.

6.27 Lemma. *Auf $K = \mathbb{T}$ existieren nur Haarsche Räume ungerader Dimension.*

Beweis. Damit in x_{n-1} ein Vorzeichenwechsel von v in der obigen Konstruktion vorliegt, muss $-\sigma \neq (-1)^{n-1}\sigma$ gelten, also muss n ungerade sein. $\square$

Der vollständige Beweis des Satzes von Krein benutzt den erwähnten Trick geschickt, um die folgenden zwei Aussagen herzuleiten. Die sehr technischen Beweise werden kurz zusammengefasst.

6.28 Lemma. *Sei $K = [a, b]$ oder $K = \mathbb{T}$ und $M \subseteq C(K)$ ein Haarscher Raum der Dimension n. Zu jedem Paar $(j, k) \in \mathbb{N}_0^2$ mit $j + 2k = n - 1$ sowie paarweise verschiedenen Punkten $\xi_1, \ldots, \xi_j \in K$, $\eta_1, \ldots, \eta_k$ im Inneren von K existiert ein $v \in M$ mit einfachen Nullstellen $\xi_1, \ldots, \xi_j$ sowie doppelten Nullstellen $\eta_1, \ldots, \eta_k$. Insbesondere wechselt v sein Vorzeichen nur in $\xi_1, \ldots, \xi_j$.*

Beweis. Man macht aus den geforderten doppelten Nullstellen η_i je zwei einfache Nullstellen $\eta_i \pm \delta$, ordnet die so entstandene Punktmenge zu

$$x_1 < x_2 < \ldots < x_{n-1}$$

und bildet die Funktion v_δ wie in (6.17) mit den $n-1$ einfachen Nullstellen $\xi_1, \ldots, \xi_j$ sowie $\eta_i \pm \delta$, $1 \le i \le k$. Durch Normalisierung erhält man

$$w_\delta = c_\delta\ v_\delta = \sum_{i=1}^{n} \alpha_i(\delta)\varphi_i$$

mit $\sum\limits_i |\alpha_i(\delta)| = 1$. Das übliche Kompaktheitsargument für die Koeffizientenvektoren in $\mathbb{R}^n$ liefert für $\delta = 1/m \longrightarrow 0$ eine konvergente Teilfolge

$$w_{1/m_\ell} \longrightarrow w = \sum_{i=1}^{n} \alpha_i\varphi_i$$

mit $\sum_i |\alpha_i| = 1$. Also ist $w \not\equiv 0$ und hat wie alle Folgenglieder $w_{1/m}$ einfache Nullstellen $\xi_1, \ldots, \xi_j$. Weiterhin führt das „Zusammenlaufen“ der beiden einfachen Nullstellen $\eta_i \pm 1/m$ zu einer doppelten Nullstelle von w in η_i. $\square$

6.29 Lemma. *Sei $K = [a, b]$ oder $K = \mathbb{T}$ und $M \subseteq C(K)$ ein Haarscher Raum der Dimension n. Zu jedem $0 \le j \le n-1$ (nur gerades j im Fall $K = \mathbb{T}$) sowie paarweise verschiedenen Punkten $\xi_1, \ldots, \xi_j$ im Inneren von K existiert ein $v \in M$ mit einfachen Nullstellen in $\xi_1, \ldots, \xi_j$ und keinen weiteren Nullstellen.*

Aus diesem Lemma folgt der Satz von Krein, indem wir $j = 0$ setzen und das Vorzeichen von v an einer Stelle $x \in K$ positiv wählen.

Beweis. Die Stellen $\xi_1, \ldots, \xi_j$ seien streng monoton geordnet.
Im Fall $j = n - 1$ wählen wir

$$v(x) = D\begin{pmatrix} \varphi_1, \ldots, \varphi_n \\ \xi_1, \ldots, \xi_{n-1}, x \end{pmatrix}.$$

Der Fall $j = n - 2$ tritt nur für $K = [a, b]$ auf und dann ist $a < \xi_1 < \ldots < \xi_{n-2} < b$. Wir wählen zwei Funktionen

$$\begin{aligned} v_a(x) &= D\begin{pmatrix} \varphi_1, \ldots, \varphi_n \\ a, \xi_1, \ldots, \xi_{n-2}, x \end{pmatrix}, \\ v_b(x) &= \pm D\begin{pmatrix} \varphi_1, \ldots, \varphi_n \\ \xi_1, \ldots, \xi_{n-2}, b, x \end{pmatrix}, \end{aligned}$$

wobei das Vorzeichen von v_b so zu wählen ist, dass $\operatorname{sign} v_a(\xi_1 - \delta) = \operatorname{sign} v_b(\xi_1 - \delta)$ für kleine $\delta > 0$ ist. Mit den Überlegungen zu (6.17) folgt dann

$$\operatorname{sign} v_a(x) = \operatorname{sign} v_b(x) \qquad \text{für alle } \; x \in (a, b) \setminus \{\xi_1, \ldots, \xi_{n-2}\}.$$

Deshalb hat $v = v_a + v_b$ außer den einfachen Nullstellen in $\xi_1, \ldots, \xi_{n-2}$ keine weiteren Nullstellen.
Im Fall $j = n - 1 - 2k$ mit $k \ge 1$ wählen wir zwei beliebige Sätze von k Punkten η_i und τ_i, $1 \le i \le k$, im Inneren von K, so dass alle Punkte $\xi_1, \ldots, \xi_j, \eta_1, \ldots, \eta_k, \tau_1, \ldots, \tau_k$ paarweise verschieden sind. Die nach Lemma 6.28 gebildeten Funktionen v_1 (zu ξ_i und η_i) sowie $\pm v_2$ (zu ξ_i und τ_i) haben ihre Vorzeichenwechsel nur in $\xi_1, \ldots, \xi_j$, ihre Summe $v = v_1 + v_2$ hat nur $\xi_1, \ldots, \xi_j$ als einfache Nullstellen und keine weiteren Nullstellen in K.
Der Fall $j = n - 2 - 2k$ mit $k \ge 1$ tritt nur für $K = [a, b]$ auf und dann ist $a < \xi_1 < \ldots < \xi_{n-2-2k} < b$. Wie im vorherigen Fall geben wir zwei Sätze von jeweils k doppelten Nullstellen vor, bilden aber v_1 mit der zusätzlichen einfachen Nullstelle a und v_2 mit b. Der Rest folgt wie zuvor. $\square$

6.3.3 Charakterisierung des Proximums

Die Charakterisierung von Proxima im Polynomraum $\mathcal{P}_n$ durch **Extremalalternanten** wurde bereits 1857 von Tschebyschew [5] dargelegt. Zunächst beschreiben wir zwei wesentliche Begriffe.

6.30 Definition. *Es sei $K \subseteq \mathbb{R}$ kompakt oder $K = \mathbb{T}$ sowie $M \subseteq C(K)$ ein Teilraum. Für $f \in C(K) \setminus M$ und $v \in M$ heißen die Punkte*

$$x_0 < x_1 < \ldots < x_r$$

(a) eine **Alternante der Länge** $r + 1$ *von $f - v$, wenn gilt*

$$(-1)^k \sigma \ (f(x_k) - v(x_k)) > 0 \qquad \text{für} \ \ 0 \leq k \leq r,$$

(b) eine **Extremal-Alternante der Länge** $r + 1$ *von $f - v$, wenn gilt*

$$(-1)^k \sigma \ (f(x_k) - v(x_k)) = \|f - v\|_\infty \qquad \text{für} \ \ 0 \leq k \leq r.$$

Hierbei ist $\sigma = \operatorname{sign} \ (f(x_0) - v(x_0))$.

Mit Alternanten lässt sich eine untere Schranke für $E_M(f)$ bestimmen.

6.31 Satz (de la Vallée-Poussin). *Es sei $\widetilde{K} = [a, b]$ oder $\widetilde{K} = \mathbb{T}$ sowie $M \subseteq C(\widetilde{K})$ ein Haarscher Raum der Dimension n. Weiter seien $K \subseteq \widetilde{K}$ kompakt mit $\#K \geq n + 1$ sowie $f \in C(K)$ und $v \in M$ gegeben. Dann gelten für jede Alternante*

$$x_0 < x_1 < \ldots < x_n \qquad \text{in} \ \ K$$

der Länge $n + 1$ von $f - v$ die Ungleichungen

$$\min_{0 \leq k \leq n} |f(x_k) - v(x_k)| \leq E_M(f) = \inf_{w \in M} \|f - w\|_{\infty,K} \leq \|f - v\|_{\infty,K}.$$

6.32 Bemerkung. Die Einschränkung des Approximationsproblems auf Teilmengen $K \subseteq [a, b]$ bzw. $K \subseteq \mathbb{T}$ wird für viele praktische Zwecke benötigt. Die Haarsche Bedingung ist aber auf dem gesamten Intervall $[a, b]$ bzw. $\mathbb{T}$ zu erfüllen.

[5]Pafnuti Lwowitsch Tschebyschew (auch Tschebyscheff, Tschebyschow, Tschebyschev und insbesondere im Englischen als Chebyshev transkribiert); geb. 16. Mai 1821, gest. 8. Dezember 1894 war einer der bedeutendsten russischen Mathematiker des 19. Jahrhunderts. Er hat die Wahrscheinlichkeitstheorie fundamental neu aufgebaut, Interpolation und Approximation sowie Algebra und Zahlentheorie erforscht. Mit seinem Namen verbindet man die Chebyshev-Polynome, Chebyshev-Ungleichung, und mancherorts wird die Maximumsnorm als Chebyshev-Norm bezeichnet.

Beweis. Angenommen es existiert $v^* \in M$ mit

$$E := \|f - v^*\|_{\infty,K} < \min_{0 \le k \le n} |f(x_k) - v(x_k)|.$$

Mit $\sigma = \text{sign}\,(f(x_0) - v(x_0))$ folgt für $0 \le k \le n$ die Beziehung

$$(-1)^k \sigma\ (v^*(x_k) - v(x_k)) = (-1)^k \sigma\ [(f(x_k) - v(x_k)) - (f(x_k) - v^*(x_k))] > 0.$$

Die Funktion $v^* - v \in M \setminus \{0\}$ hat nach dem Zwischenwertsatz mindestens je eine Nullstelle in (x_k, x_{k+1}), $0 \le k \le n-1$. (Achtung: Hier geht ein, dass $v^* - v$ auf $[a,b]$ bzw. $\mathbb{T}$ definiert ist, nicht nur auf K.) Dies ist ein Widerspruch zur Haarschen Bedingung.
Damit ist die untere Schranke für $E_M(f)$ gezeigt. Die obere Schranke ist klar. $\square$

Für eine **Extremal**alternante liefert der Satz von de la Vallée-Poussin bereits die Einschließung

$$\min_{0 \le k \le n} |f(x_k) - v(x_k)| = \|f - v\|_{\infty,K} \le E_M(f) \le \|f - v\|_{\infty,K},$$

also folgt Gleichheit und v ist Proximum an f. Die Umkehrung ist im Alternantensatz enthalten.

6.33 Satz (Alternantensatz). *Es sei $\widetilde{K} = [a,b]$ oder $\widetilde{K} = \mathbb{T}$ sowie $M \subseteq C(\widetilde{K})$ ein Haarscher Raum der Dimension n. Weiter sei $K \subseteq \widetilde{K}$ kompakt mit $\#K \ge n+1$ sowie $f \in C(K)$.*
Die Funktion $v^* \in M$ ist genau dann Proximum an f auf K, wenn eine Extremalalternante (der Länge $n+1$)

$$x_0 < x_1 < \ldots < x_n \qquad \textit{in } \; K \tag{6.18}$$

von $f - v^*$ existiert.

Beweis. Wir brauchen nur noch die Existenz einer Extremalalternante (6.18) der Länge $n+1$ nachzuweisen, wenn v^* Proximum an f ist. Sei dazu $f \in C(K) \setminus M$. Dann ist

$$E_M(f) = \max_{x \in K} |f(x) - v^*(x)| > 0$$

und die Extremalmenge $A_\infty(f - v^*)$ enthält nach Lemma 6.21 mindestens $n+1$ Elemente. Wir betrachten nun das lineare Funktional $\lambda : C(K) \longrightarrow \mathbb{R}$ aus Satz 6.14 (d),

$$\lambda(g) = \sum_{k=0}^{r-1} p_k \ \text{sign}\,(f(x_k) - v^*(x_k))\ g(x_k)$$

mit $r \le n+1$ Punkten $x_0 < x_1 < \ldots < x_{r-1}$ aus $A_\infty(f - v^*)$, Zahlen $p_k \ge 0$ und $\sum\limits_{k=0}^{r-1} p_k = 1$.

1. Zunächst zeigen wir, dass bei diesem Funktional

$$r = n + 1 \quad \text{und} \quad p_k > 0 \;\; \text{für} \;\; 0 \le k \le n$$

vorliegen müssen. Wäre eine dieser Eigenschaften verletzt, würde $v \in M$ existieren mit $\lambda(v) \neq 0$: wähle eine Interpolationsfunktion, die an den Stellen x_k mit $p_k > 0$ den Wert $\sigma_k := \operatorname{sign}(f(x_k) - v^*(x_k))$ hat (dies sind höchstens n Stellen).

2. Dann zeigen wir, dass die Vorzeichen $\sigma_k = \operatorname{sign}(f(x_k) - v^*(x_k))$ alternieren. Wäre dies nicht so, dann ließen sich $s \le n$ Mengen aufeinander folgender Indizes mit gleichem Vorzeichen

$$\sigma_0 = \cdots = \sigma_{k_1}, \quad \sigma_{k_1+1} = \cdots = \sigma_{k_2}, \quad \ldots, \quad \sigma_{k_{s-1}+1} = \cdots = \sigma_n,$$

bilden. Die zugehörigen x_k trennen wir mit Punkten $\xi_i \in \widetilde{K}$ gemäß

$$x_0, \ldots, x_{k_1} < \xi_1 < x_{k_1+1}, \ldots, x_{k_2} < \xi_2 < \cdots < \xi_{s-1} < x_{k_{s-1}+1}, \ldots, x_n.$$

Die Punkte ξ_i liegen im Inneren von $\widetilde{K}$. Lemma 6.29 erlaubt nun die Wahl von $v \in M \setminus \{0\}$, dessen einzige Nullstellen (in der zusammenhängenden Menge $\widetilde{K}$) die Punkte $\xi_1, \ldots, \xi_{s-1}$ sind und die in diesen Punkten das Vorzeichen wechselt. Für dieses v ergibt sich wieder der Widerspruch $\lambda(v) \neq 0$.

Insgesamt haben wir gezeigt, dass die Punkte x_k des Funktionals λ eine Extremalalternante der Länge $n + 1$ sind. $\square$

Zusammenfassung: Die Bedeutung Haarscher Räume als Eindeutigkeitsmengen für die Approximation in $C(K)$ wurde in Satz 6.22 dargelegt. Das Kolmogorov-Kriterium konnte für Haarsche Räume in Form des Alternantensatzes verstärkt werden.

6.3.4 Remez-Algorithmus

Wir behandeln in diesem Kapitel einen Algorithmus, benannt nach dem russischen Mathematiker Evgeny Yakovlevich Remez (1896–1975), der zur numerischen Berechnung der Bestapproximation in einem Haarschen Raum eingesetzt wird. Sei dazu $M \subset C(K)$ ein Haarscher Raum der Dimension n sowie $f \in C(K) \setminus M$. Der Algorithmus konstruiert iterativ Alternanten

$$x_0^{(k)} < x_1^{(k)} < \cdots < x_n^{(k)}, \qquad k = 0, 1, 2, \ldots \tag{6.19}$$

einer Fehlerfunktion $f - v_k$ mit $v_k \in M$, die an den Stellen $x_j^{(k)}$ den gleichen Absolutbetrag annimmt, also

$$d_k = (-1)^j (f - v_k)(x_j^{(k)}) \tag{6.20}$$

erfüllt. Wir sagen, dass die Fehlerfunktion $f - v_k$ in der Punktmenge (6.19) *äqui-oszilliert*. Dann folgt mit dem Satz von de la Vallée-Poussin

$$|d_k| \leq E_M(f) \leq D_k := \|f - v_k\|_{\infty,K}. \tag{6.21}$$

Die Zahlenfolge $(|d_k|)_{k \geq 0}$ wird in dieser Konstruktion sogar monoton wachsen und

$$\lim_{k\to\infty} |d_k| = E_M(f) = \lim_{k\to\infty} D_k \tag{6.22}$$

erfüllen. Dann wird gezeigt, dass die Folge der Elemente $v_k \in M$ gleichmäßig gegen das eindeutige Proximum v^* von f konvergiert.

Die Bestätigung, dass es zu genau $n + 1$ gegebenen paarweise verschiedenen Punkten

$$B = \{x_0 < x_1 < \cdots < x_n\} \subset K$$

eine Funktion $v \in M$ gibt, deren Fehlerfunktion $f - v$ wie in (6.20) äqui-oszilliert, liefert bereits der Alternantensatz. Wir nennen im Folgenden v das "diskrete Proximum" zu $f|_B$.

6.34 Satz. *Es sei $\widetilde{K} = [a, b]$ oder $\widetilde{K} = \mathbb{T}$ sowie $M \subset C(\widetilde{K})$ ein Haarscher Raum der Dimension n. Weiter sei $K \subset \widetilde{K}$ kompakt mit $\#K \geq n + 1$ sowie $f \in C(K)$. Dann existieren zu jeder Teilmenge*

$$B = \{x_0 < x_1 < \cdots < x_n\} \subset K$$

eindeutig bestimmte Elemente $v \in M$ und $d \in \mathbb{R}$ mit

$$f(x_j) - v(x_j) = (-1)^j d. \tag{6.23}$$

Beweis. Die Einschränkung $f\,|_B\colon B \to \mathbb{R}$ von f ist ein Element von $C(B)$. Das eindeutige Proximum $v \in M$ (eigentlich $v\,|_B$ mit $v \in M$) ist durch die Extremalalternante

$$f(x_j) - v(x_j) = (-1)^j \sigma \|f - v\|_{\infty,B}, \qquad 0 \le j \le n,$$

mit $\sigma \in \{-1, 1\}$ charakterisiert. Die Aussage folgt mit

$$d = \sigma \|f - v\|_{\infty,B} = \pm E_{M|_B}(f\,|_B).$$

□

Die Berechnung des diskreten Proximums v zu $f|_B$ sowie der Zahl d erfolgt mit Mitteln der Linearen Algebra. Als Nebenprodukt wird sogar das lineare Funktional in Satz 6.14(d) gefunden, durch das das diskrete Proximum charakterisiert ist. Dies wird wie folgt präzisiert.

6.35 Folgerung. *Die Voraussetzungen in Satz 6.34 seien erfüllt. Weiter sei* $\{\phi_1, \dots, \phi_n\}$ *eine Basis von* M*. Dann ist die zur Teilmenge*

$$B = \{x_0 < x_1 < \dots < x_n\} \subset K$$

gebildete Matrix

$$W := \begin{bmatrix} \phi_1(x_0) & \cdots & \phi_n(x_0) & 1 \\ \phi_1(x_1) & \cdots & \phi_n(x_1) & -1 \\ \vdots & & \vdots & \vdots \\ \phi_1(x_n) & \cdots & \phi_n(x_n) & (-1)^n \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$$

regulär. Weiterhin gilt:

(a) Die Lösung $[c_1, \dots, c_n, d]^T \in \mathbb{R}^{n+1}$ *des linearen Gleichungssystems*

$$W\ [c_1, \dots, c_n, d]^T = [f(x_0), \dots, f(x_n)]^T \tag{6.24}$$

liefert sowohl das diskrete Proximum $v = \sum_i c_i v_i$ *zu* $f\,|_B$ *als auch den Approximationsgrad* $E_{M|_B}(f\,|_B) = |d|$.

(b) Die Lösung $\alpha = [\alpha_0, \dots, \alpha_n] \in \mathbb{R}^{n+1}$ *des linearen Gleichungssystems*

$$\alpha W = [0, \dots, 0, 1] \tag{6.25}$$

erfüllt die Beziehungen $(-1)^j \alpha_j > 0$ *und* $\sum_j (-1)^j \alpha_j = 1$*. Sie definiert das lineare Funktional*

$$\lambda_B : C(B) \to \mathbb{R}, \quad \lambda_B(g) = \sum_{j=0}^{n} \alpha_j g(x_j), \tag{6.26}$$

mit den Eigenschaften

$$(i)\ \|\lambda_B\| = 1, \qquad (ii)\ \lambda_B(v\ |_B) = 0 \text{ für alle } v \in M,$$

und

$$(iii)\ \lambda_B(f\ |_B) = d = \operatorname{sign}(d)\ E_{M|_B}(f\ |_B).$$

Also ist $\lambda = \operatorname{sign}(d)\ \lambda_B$ *ein Funktional in Satz 6.14(d) zur Charakterisierung des diskreten Proximums* v *an* $f\ |_B$.

Beweis. Die Gleichung (6.23) in Satz 6.34 wird geschrieben als

$$v(x_j) + (-1)^j d = f(x_j), \qquad 0 \le j \le n.$$

Mit der Basis $\{\phi_1, \dots, \phi_n\}$ von M und $v = \sum_{i=1}^n c_i\phi_i$ lässt sich dies als lineares Gleichungssystem

$$W\ [c_1, \dots, c_n, d]^T = [f(x_0), \dots, f(x_n)]^T$$

schreiben. Die eindeutige Lösbarkeit für beliebige rechte Seiten ist gleichbedeutend mit der Invertierbarkeit von W. Außerdem folgt sofort die Aussage (a).
Zum Beweis von (b) vergleichen wir das lineare Funktional λ_B in (6.26) mit dem Funktional aus Satz 6.14(d), das auch im Beweis des Alternantensatzes auftrat. Wegen $\#B = n+1$ und Lemma 6.6 muss gelten $A_\infty(f|_B - v|_B) = B$. Also lautet das Funktional zur Extremalalternante von $f|_B - v|_B$

$$\lambda : C(B) \to \mathbb{R}, \qquad \lambda(g) = \sum_{j=0}^n p_j \operatorname{sign}\left(f(x_j) - v(x_j)\right)\ g(x_j).$$

Hierbei sind alle $p_j > 0$, $\sum_j p_j = 1$ und die Sequenz $\sigma_j := \operatorname{sign}\left(f(x_j) - v(x_j)\right)$, $0 \le j \le n$, hat alternierendes Vorzeichen. Die Zahlen $\beta_j := (-1)^j p_j = \sigma_0 p_j \sigma_j$ sind dann eine Lösung des linearen Gleichungssystems

$$[\beta_0, \dots, \beta_n]W = [0, \dots, 0, 1],$$

denn die ersten n Gleichungen gelten wegen $\lambda(\phi_i) = 0$ und die letzte Gleichung wegen $\sum_j (-1)^j \beta_j = \sum_j p_j = 1$. Die Invertierbarkeit von W liefert $\alpha_j = \beta_j$, und hieraus folgen die geforderten Eigenschaften der Zahlen α_j sowie $\lambda_B = \sigma_0\lambda$, also auch die Eigenschaften von λ_B. □

6.36 Bemerkung. Das lineare Gleichungssystem

$$W\ [c_1, \dots, c_n, d]^T = [f(x_0), \dots, f(x_n)]^T$$

lässt sich im Fall $M = \mathcal{P}_{n-1}$ vereinfachen. Hierbei wird zuerst die Komponente $d \in \mathbb{R}$ berechnet und anschließend die Interpolationsaufgabe

$$v(x_j) = f(x_j) + (-1)^j d, \qquad 0 \leq j \leq n-1,$$

gelöst. (Die Interpolationsbedingung an der Stelle x_n ergibt sich daraus automatisch!) Zur Berechnung von d benötigt man die *dividierten Differenzen*

$$\begin{aligned} s_B &:= [x_0, \ldots, x_n \mid s] \neq 0 \quad \text{mit} \quad s \in C(B), \;\; s(x_j) = (-1)^j, \\ f_B &:= [x_0, \ldots, x_n \mid f] \end{aligned}$$

und erhält $d = \frac{f_B}{s_B}$. Denn: Das diskrete Proximum $v\,|_B$ erfüllt

$$f(x_j) = v(x_j) + ds(x_j), \qquad 0 \leq j \leq n.$$

Die dividierte Differenz n-ter Ordnung annulliert v, also gilt $f_B = ds_B$. Die Beziehung $s_B \neq 0$ ist wieder eine Konsequenz der Haarschen Bedingung. (s_B ist der Höchstkoeffizient des Interpolationspolynoms vom Grad n zu den Daten $s(x_j) = (-1)^j$. Dieses Polynom muss den exakten Grad n haben, da es n Nullstellen besitzt, aber nicht identisch verschwindet.)

Bemerkenswert ist ebenfalls, dass das lineare Funktional in Folgerung 6.35 gegeben ist durch das Vielfache der dividierten Differenz

$$\lambda_B(g) = \frac{1}{s_B}\,[x_0, \ldots, x_n \mid g].$$

Dies liefert eine explizite Darstellung der Koeffizienten α_j in (6.26), nämlich

$$\alpha_j = \frac{1}{s_B} \prod_{\substack{\ell=0 \\ \ell \neq j}}^{n} \frac{1}{x_j - x_\ell}.$$

Als weitere Konsequenz aus der Folgerung 6.35 halten wir eine Form von "Definitheit" des Funktionals

$$\mu(v) = \max_{0 \leq j \leq n} (-1)^j\, v(x_j)$$

fest.

6.37 Lemma. *$M \subset C(\widetilde{K})$, mit $\widetilde{K} = [a,b]$ oder $\mathbb{T}$, sei ein Haarscher Raum der Dimension n. Dann existiert zu Punkten*

$$x_0 < x_1 < \cdots < x_n \qquad \textit{in} \qquad \widetilde{K}$$

ein $0 < \gamma < 1$ so, dass für alle $v \in M$

$$\gamma\, \|v\|_{\infty,\widetilde{K}} \leq \max_{0 \leq j \leq n} (-1)^j\, v(x_j) \leq \|v\|_{\infty,\widetilde{K}} \tag{6.27}$$

gilt.

Bemerkung: Die Eigenschaft (6.27) sieht aus wie die Äquivalenz von Normen. Würde man an Stelle des Funktionals μ das Funktional

$$\beta(v) = \max_{0 \le j \le n} |v(x_j)|$$

verwenden, so wäre hierdurch tatsächlich eine äquivalente Norm auf M definiert: die Haarsche Bedingung ergibt die Definitheit, und die weiteren Axiome der Norm sind erfüllt. Die Äquivalenz zur Maximum-Norm folgt aus der Äquivalenz aller Normen auf M. In unserem Fall ist durch μ aber keine Norm definiert, da im Allgemeinen $v \in M$ existiert mit $\mu(-v) \neq \mu(v)$.

Beweis. Die obere Abschätzung ist trivial.
Für die untere Abschätzung verwenden wir Folgerung 6.35. Sei dazu $\{\phi_1, \dots, \phi_n\}$ eine Basis von M und W die $(n+1) \times (n+1)$ Matrix zur gegebenen Punktmenge $B = \{x_0, \dots, x_n\}$. Weiter sei $\alpha = [\alpha_0, \dots, \alpha_n]$ der Lösungsvektor des Gleichungssystems $\alpha W = [0, \dots, 0, 1]$. Insbesondere kennen wir das Vorzeichenverhalten $(-1)^j \alpha_j > 0$ für $0 \le j \le n$. Für beliebiges $v = \sum_i c_i \phi_i \in M \setminus \{0\}$ gilt dann

$$\sum_{j=0}^{n} \alpha_j v(x_j) = \alpha \; W \; [c_1, \dots, c_n, 0]^T = [0, \dots, 0, 1] \; [c_1, \dots, c_n, 0]^T = 0. \tag{6.28}$$

Die Annahme $\max_{0 \le j \le n} (-1)^j v(x_j) \le 0$ führt mit der Haarschen Bedingung auf

$$(-1)^j v(x_j) < 0 \quad \text{für mindestens zwei} \quad 0 \le j \le n.$$

Dann ist aber

$$\sum_{j=0}^{n} \alpha_j v(x_j) = \sum_{j=0}^{n} \underbrace{(-1)^j \alpha_j}_{>0} \underbrace{(-1)^j v(x_j)}_{\le 0} < 0,$$

im Widerspruch zu (6.28). Also gilt

$$\max_{0 \le j \le n} (-1)^j v(x_j) > 0 \quad \text{für alle} \quad v \in M \setminus \{0\}. \tag{6.29}$$

Die linearen Funktionale $\mu_j : M \to \mathbb{R}$, $\mu_j(v) = (-1)^j v(x_j)$ sind offensichtlich stetig, deshalb ist auch $\mu : M \to \mathbb{R}$, $\mu(v) = \max_j \mu_j(v)$ stetig. Auf der kompakten Menge $M_1 = \{v \in M;\ \|v\| = 1\}$ nimmt μ sein Minimum an, und wegen (6.29) gilt

$$\gamma := \min_{v \in M_1} \mu(v) > 0.$$

Die positive Homogenität $\mu(cv) = c\mu(v)$ (mit $c > 0$) liefert nun

$$\gamma \|v\| \le \mu(v) \quad \text{für alle} \quad v \in M. \quad \square$$

Die bisherigen Ergebnisse dieses Abschnitts sind die wesentlichen Bestandteile der Beweise von zwei allgemeinen Resultaten Ã¼ber Haarsche Räume. Die folgenden zwei Sätze unterstreichen nochmals die Bedeutung der Haarschen Bedingung fÃ¼r spezielle weitere Eigenschaften der Bestapproximation. Sie sollten daher mit den allgemeinen Ergebnissen aus Kapitel 2 verglichen werden.

6.38 Satz (Newman, Shapiro, 1963: Starke Eindeutigkeit des Proximums)**.**
Es sei $\widetilde{K} = [a, b]$ oder $\mathbb{T}$ und $M \subset C(\widetilde{K})$ sei ein Haarscher Raum. Weiter sei $K \subset \widetilde{K}$ mit $\#K \geq n + 1$. Dann existiert zu $f \in C(K)$ eine Konstante $\gamma = \gamma(f) > 0$ mit

$$\|f - v\|_{\infty,K} \geq \|f - \Pi_M(f)\|_{\infty,K} + \gamma\|v - \Pi_M(f)\|_{\infty,K}.$$

Beweis. Wir setzen $n = \dim M$, $v^* = \Pi_M(f)$ und wählen eine Extremalalternante

$$x_0 < x_1 < \cdots < x_n \qquad \text{in} \qquad A_\infty(f - v^*).$$

Die Vorzeichen

$$\sigma_j = \sigma(-1)^j = \operatorname{sign}(f(x_j) - v^*(x_j)), \qquad 0 \leq j \leq n,$$

alternieren. Für beliebiges $v \in M$ und $0 \leq j \leq n$ erhalten wir die Abschätzung

$$\|f - v\|_{\infty,K} \geq \sigma_j(f - v)(x_j) = \underbrace{\sigma_j(f - v^*)(x_j)}_{= \|f - v^*\|_{\infty,K}} + \sigma_j(v^* - v)(x_j).$$

Also gilt

$$\|f - v\|_{\infty,K} \geq \|f - v^*\|_{\infty,K} + \max_{0 \leq j \leq n} \sigma_j(v^* - v)(x_j).$$

Mit Lemma 6.37 folgt die Existenz der Konstante γ in der Behauptung. $\square$

Als weiteres Ergebnis halten wir eine Eigenschaft der metrischen Projektion Π_M auf einen Haarschen Raum fest.

6.39 Satz (Freud 1958: Lokale Lipschitz-Stetigkeit der metrischen Projektion)**.**
Die Voraussetzungen aus Satz 6.38 seien erfüllt und $f \in C(K) \setminus M$ sei gegeben. Dann gilt mit $\gamma = \gamma(f)$ aus Satz 6.38 die Abschätzung

$$\|\Pi_M(f) - \Pi_M(g)\|_{\infty,K} \leq \frac{2}{\gamma}\,\|f - g\|_{\infty,K} \quad \textit{für alle} \quad g \in C(K).$$

Beweis. Mit γ aus Satz 6.38 und Umstellen der Ungleichung der starken Eindeutigkeit folgt

$$\begin{aligned}
\gamma\|\Pi_M(f) - \Pi_M(g)\| &\leq \|f - \Pi_M(g)\| - \|f - \Pi_M(f)\| \\
&\leq \|f - g\| + \|g - \Pi_M(g)\| - \|f - \Pi_M(f)\| \\
&\leq \|f - g\| + \|g - \Pi_M(f)\| - \|f - \Pi_M(f)\| \\
&\leq \|f - g\| + \|f - g\| = 2\|f - g\|. \quad \square
\end{aligned}$$

Wir können nun den Remez-Algorithmus in seiner einfachsten Form formulieren und die lineare Konvergenz beweisen.

6.40 Bemerkung. Die Formulierung des Algorithmus sieht etwas technisch aus, lässt sich aber mit einem sehr einfachen Spiel beschreiben. Die aufsteigende Punktmenge stellen wir uns vor als eine "Menschenkette" von Personen mit aufsteigender Körpergröße und abwechselnd roter und blauer Mütze. Die Farbe repräsentiert dabei das Vorzeichen der Fehlerfunktion $f - v_k$, die Punkte bilden ja eine Alternante. Im Fall $K = \mathbb{T}$ bilden die Personen sogar einen Kreis, weil sich die kleinste und größte Person an die Hand nehmen. (Aha, die Anzahl ist dann gerade.) Nun wird eine Person mit der Körpergröße ξ von außen dazugenommen (Schritt 2), ihre Mütze richtet sich nach dem Vorzeichen von $(f - v_k)(\xi)$. Sie wird in die Punktmenge B an die richtige Stelle "eingewechselt", je nach Farbe der Mütze fliegt der linke oder rechte Nachbar raus (Schritt 4). Damit ist für $K = \mathbb{T}$ alles vorbereitet, um das neue diskrete Proximum v_{k+1} auszurechnen. (Evtl. müssen danach die Mützen getauscht werden :)
Für $K = [a, b]$ gibt es zwei Ausnahmefälle: gehört die neue Person wegen der Körpergröße ξ an den Anfang der Kette, hat aber nicht die gleiche Farbe wie der bisherige Anfang, so fliegt die Person am Ende der Kette raus. Das gleiche kann dem Kettenanfang passieren, wenn die neue Person an das Ende der Kette gehört.

Die algorithmische Formulierung lautet wie folgt. Die Implementierung für die Approximation mit $\mathcal{P}_n$ oder $\mathcal{T}_n$ steht auf der Vorlesungs-Homepage.

6.41 Algorithmus (Remez-Algorithmus mit einfachem Austauschschritt (1934))**.**

Gegeben: $M \subset C(\widetilde{K})$ Haarscher Raum mit Basis $\{\phi_1, \dots, \phi_n\}$,
$K \subset \widetilde{K}$ mit $\#K \geq n + 1$ und $f \in C(K)$.

Start: Wähle eine Punktmenge $B_0 = \{x_0^{(0)} < x_1^{(0)} < \cdots < x_n^{(0)}\}$.

Iteration: Für $k = 0, 1, 2, \dots$

1. berechne das diskrete Proximum $v_k \in M$ und $d_k \in \mathbb{R}$ mit
$$v_k(x_j^{(k)}) + (-1)^j d_k = f(x_j^{(k)}), \qquad 0 \leq j \leq n,$$
durch Lösung des linearen Gleichungssystems (6.24);
2. berechne (eine gute Näherung von)
$$D_k = \|f - v_k\|_{\infty,K}$$
und einen Punkt ξ, an dem das Maximum oder Minimum angenommen wird;

3. falls $D_k - |d_k| < \epsilon$, so ist B_k (fast) Extremalalternante und v_k eine geeignete Näherung an das Proximum;
4. sonst bestimme B_{k+1} durch Austauschen eines $x_\ell^{(k)}$ gegen ξ nach folgender Vorzeichenregel:
   - falls $x_j^{(k)} < \xi < x_{j+1}^{(k)}$ gilt, so ersetze $x_\ell^{(k)}$ mit dem Index

     $$\ell = \begin{cases} j, & \text{falls } \ (f - v_k)(x_j^{(k)}) \cdot (f - v_k)(\xi) > 0, \\ j+1, & \text{falls } \ (f - v_k)(x_{j+1}^{(k)}) \cdot (f - v_k)(\xi) > 0; \end{cases}$$

     falls $\widetilde{K} = \mathbb{T}$, so werden hier auch die Fälle $x_n^{(k)} < \xi < x_0^{(k)} + 1$ und $x_n^{(k)} - 1 < \xi < x_0^{(k)}$ durch Setzen von $\ell = 0$ oder $\ell = n$ einbezogen. Sodann ist die neue Punktmenge

     $$B_{k+1} = \{x_0^{(k)}, \ldots, x_{\ell-1}^{(k)}, \xi, x_{\ell+1}^{(k)}, \ldots, x_n^{(k)}\},$$

   - im Fall $\widetilde{K} = [a, b]$ können zwei Sonderfälle $\xi < x_0^{(k)}$ und $\xi > x_n^{(k)}$ auftreten. Im ersten setzen wir

     $$B_{k+1} = \begin{cases} \{\xi, x_1^{(k)}, \ldots, x_n^{(k)}\}, & \text{falls } \ (f - v_k)(x_0^{(k)}) \cdot (f - v_k)(\xi) > 0, \\ \{\xi, x_0^{(k)}, \ldots, x_{n-1}^{(k)}\}, & \text{falls } \ (f - v_k)(x_0^{(k)}) \cdot (f - v_k)(\xi) < 0; \end{cases}$$

     Im zweiten Sonderfall $\xi > x_n^{(k)}$ verfahren wir entsprechend, ersetzen also auch entweder $x_n^{(k)}$ oder $x_0^{(k)}$.
5. und fahre mit der neuen Punktmenge B_{k+1} fort.

Die Glieder der Folgen $(d_k)_{k\geq 0}$ und $(D_k)_{k\geq 0}$ erfüllen

$$|d_k| \leq E_M(f) = \min_{v \in M} \|f - v\|_{\infty,K} \leq D_k, \qquad k \geq 0.$$

Hinter der ersten Ungleichung steckt der Satz von de la Vallée-Poussin. Bei der praktischen Durchführung beachte man, dass die Sonderfälle zu einem Vorzeichenwechsel in der Folge (d_k) führen können. Wir zeigen nun, dass monotone und lineare Konvergenz $|d_k| \to E_M(f)$ vorliegt und dass auch die Folge v_k linear (bzgl. der Maximum-Norm auf K) gegen das Proximum v^* konvergiert.

6.42 Satz (Lineare Konvergenz des Remez-Algorithmus).
Der Remez-Algorithmus mit einfachem Austauschschritt konvergiert linear. Genauer: zu $f \in C(K) \setminus M$ existieren Konstanten $0 < c < 1$ und $\gamma > 0$, die nur von f, der gewählten Basis von M und der Start-Punktmenge B_0 abhängen, mit

(i) $E_M(f) - |d_k| \leq (1-c)^k D_0$,

(ii) $\|v_k - v^*\|_{\infty,K} \leq (1-c)^k \,\dfrac{D_0}{c\gamma}$.

Außerdem ist die Folge der $|d_k|$ streng monoton wachsend.

Beweis. Der Kern des Beweises besteht darin, die Existenz von $0 < c < 1$ zu zeigen mit der Eigenschaft

$$|d_{k+1}| \geq |d_k| + c(D_k - |d_k|). \tag{6.30}$$

Dann folgen nämlich die Behauptungen der Monotonie von $(|d_k|)_{k\geq 0}$ sowie beide Aussagen (i) und (ii):

(i) Es gilt

$$\begin{aligned} E_M(f) - |d_{k+1}| &\leq E_M(f) - |d_k| - c(D_k - |d_k|) \\ &\leq E_M(f) - |d_k| - c(E_M(f) - |d_k|) \qquad (\text{wegen } \; D_k \geq E_M(f)) \\ &= (1-c)(E_M(f) - |d_k|). \end{aligned}$$

Durch Iteration folgt

$$E_M(f) - |d_k| \leq (1-c)^k (E_M(f) - |d_0|) \leq (1-c)^k E_M(f) \leq (1-c)^k D_0.$$

(ii) Mit der Konstanten $\gamma = \gamma(f)$ aus dem Satz von Newman, Shapiro (Satz 6.38) gilt

$$\begin{aligned} \gamma\|v_k - v^*\| &\leq \|f - v_k\| - \|f - v^*\| \\ &= D_k - E_M(f) \;\leq\; D_k - |d_k| \\ &\leq \frac{1}{c}\,(|d_{k+1}| - |d_k|) \qquad (\text{wegen } (6.30)) \\ &\leq \frac{1}{c}\,(E_M(f) - |d_k|) \leq \frac{(1-c)^k}{c}\,D_0. \end{aligned}$$

Um die Beziehung (6.30) zu zeigen, verwenden wir das Funktional λ_B mit $B = B_{k+1}$ in (6.26). Die Koeffizienten des Funktionals seien mit $\alpha_j^{(k+1)}$ bezeichnet. Wir erinnern nochmals an die Beziehungen

$$(-1)^j \alpha_j^{(k+1)} > 0$$

in Folgerung 6.35. Mit $\sigma = \operatorname{sign} d_k$ und einem $\rho \in \{-1, 0, 1\}$ gelten im k-ten Austauschschritt

$$\begin{array}{lll} x_\ell^{(k+1)} = \xi, & |(f - v_k)(\xi)| = D_k, & \operatorname{sign}(f - v_k)(\xi) = (-1)^{\ell+\rho}\,\sigma, \\ x_j^{(k+1)} = x_{j+\rho}^{(k)}, & (f - v_k)(x_{j+\rho}^{(k)}) = (-1)^{j+\rho} d_k & \text{für } \; j \neq \ell, \end{array}$$

wobei $0 \leq \ell \leq n$ die Stelle des "Einwechselns" von ξ bezeichnet. Der Fall $\rho \neq 0$ kann nur in den Ausnahmefällen $\xi < x_0^{(k)}$ und $\xi > x_n^{(k)}$ auftreten.
Wegen $\lambda_{B_{k+1}}(v) = 0$ für alle $v \in M$ erhalten wir

$$\begin{aligned}
d_{k+1} &= \lambda_{B_{k+1}}(f) = \lambda_{B_{k+1}}(f - v_k) = \sum_{j=0}^{n} \alpha_j^{(k+1)}(f - v_k)(x_j^{(k+1)}) \\
&= \alpha_\ell^{(k+1)}(f - v_k)(\xi) + \sum_{\substack{j=0 \\ j \neq \ell}}^{n} \alpha_j^{(k+1)}(f - v_k)(x_{j+\rho}^{(k)}) \\
&= \alpha_\ell^{(k+1)}\ (-1)^{\ell+\rho}\ \sigma\ D_k + \sum_{\substack{j=0 \\ j \neq \ell}}^{n} \alpha_j^{(k+1)}\ (-1)^{j+\rho} d_k \\
&= (-1)^\rho \left\{ |\alpha_\ell^{(k+1)}|\ \sigma\ D_k + d_k \sum_{\substack{j=0 \\ j \neq \ell}}^{n} |\alpha_j^{(k+1)}| \right\} \\
&= (-1)^\rho \sigma \left\{ |\alpha_\ell^{(k+1)}|\,(D_k - |d_k|) + |d_k| \sum_{j=0}^{n} |\alpha_j^{(k+1)}| \right\} \\
&= (-1)^\rho \sigma \left\{ |\alpha_\ell^{(k+1)}|\,(D_k - |d_k|) + |d_k| \right\} \qquad \text{(wegen } \sum_j |\alpha_j^{(k+1)}| = 1).
\end{aligned}$$

Dies ergibt die Gleichung

$$|d_{k+1}| = |d_k| + |\alpha_\ell^{(k+1)}|\,(D_k - |d_k|)\,, \tag{6.31}$$

aus der die strikte Monotonie der Folge $(|d_k|)_{k \geq 0}$ folgt. Sobald wir gezeigt haben, dass eine untere Schranke $0 < c < 1$ von $|\alpha_\ell^{(k+1)}|$ existiert, die unabhängig von k und ℓ ist, ist die Ungleichung (6.30) bewiesen. Damit ist dann der Konvergenzsatz 6.42 komplett bewiesen. □

6.43 Lemma. *Die Funktion $f \in C(K)\backslash M$ und die Startmenge B_0 des Remez-Algorithmus seien gegeben. Dann gibt es $0 < c < 1$ und $C > 0$ mit:*

(i) $x_{j+1}^{(k)} - x_j^{(k)} \geq C$ für alle $k \geq 0$ und $0 \leq j \leq n-1$ (sowie $x_0^{(k)} + 1 - x_n^{(k)} \geq C$ im Fall $\widetilde{K} = \mathbb{T}$).

(ii) $|\alpha_j^{(k)}| \geq c$ für alle $0 \leq j \leq n$, $k \geq 0$.

Beweis. Die Mengen B_k werden im Fall $\widetilde{K} = [a, b]$ als Punkte im Simplex

$$\mathcal{S} = \{[y_0, \ldots, y_n] \in \mathbb{R}^{n+1};\ a \leq y_0 \leq y_1 \leq \cdots \leq y_n \leq b\} \subset \mathbb{R}^{n+1}$$

angesehen. Diese Menge ist abgeschlossen und beschränkt, also kompakt. Im Fall $\widetilde{K} = \mathbb{T}$ ist

$$\mathcal{S} = \{[y_0, \ldots, y_n] \in \mathbb{T}^{n+1};\ 0 \leq y_0 \leq y_1 \leq \cdots \leq y_n < 1\} \subset \mathbb{T}^{n+1}$$

ebenfalls kompakt.

(i) Angenommen, es existiert eine Teilfolge $[x_0^{(k_m)}, \dots, x_n^{(k_m)}]$, $m \geq 0$, die gegen einen Randpunkt von $\mathcal{S}$ konvergiert, für die also

$$\lim_{m\to\infty} x_j^{(k_m)} = \xi_j, \qquad 0 \leq j \leq n,$$

und

$$\xi_0 \leq \xi_1 \leq \cdots \leq \xi_n$$

mit mindestens einem Gleichheitszeichen (oder $\xi_0 + 1 = \xi_n$ im Fall $\widetilde{K} = \mathbb{T}$) gilt. Dann existiert eine Funktion $v \in M$, die die höchstens n Interpolationsbedingungen

$$v(\xi_j) = f(\xi_j), \qquad 0 \leq j \leq n,$$

erfüllt. Wegen der gleichmäßigen Stetigkeit von $f - v$ gilt

$$|(f-v)(x) - (f-v)(y)| < |d_1| \qquad \text{für alle} \qquad |x - y| < \delta$$

mit hinreichend kleinem $\delta > 0$. (Wir verwenden hier $|d_1| > 0$ als das "epsilon" in der gleichmäßigen Stetigkeit.) Wählen wir nun m so groß, dass

$$|x_j^{(k_m)} - \xi_j| < \delta \qquad \text{für} \qquad 0 \leq j \leq n$$

gilt, so erhalten wir zusammen mit der bereits bewiesenen Monotonie der $|d_k|$

$$|(f-v)(x_j^{(k_m)})| = |(f-v)(x_j^{(k_m)}) - (f-v)(\xi_j)| < |d_1| < |d_{k_m}| \qquad \text{für} \qquad 0 \leq j \leq n,$$

im Widerspruch zur Definition von d_{k_m}. Deshalb ist die Annahme falsch. Damit ist die Existenz von $C > 0$ in der Behauptung (i) gezeigt.

(ii) Die im Fall $\widetilde{K} = [a, b]$ zu $C > 0$ gebildete Teilmenge

$$\mathcal{S}_C = \{[y_0, \dots, y_n] \in \mathbb{R}^{n+1};\ a \leq y_0 < y_1 < \cdots < y_n \leq b,\ y_{j+1} - y_j \geq C \text{ für } 0 \leq j \leq n-1\}$$

von $\mathcal{S}$ ist kompakt. Im Fall $\widetilde{K} = \mathbb{T}$ wird entsprechend

$$\begin{aligned} \mathcal{S}_C = \{[y_0, \dots, y_n] \in \mathbb{T}^{n+1}; \quad & 0 \leq y_0 < y_1 < \cdots < y_n < 1, \\ & y_{j+1} - y_j \geq C \text{ für } 0 \leq j \leq n-1,\ \ y_0 + 1 - y_n \geq C\} \end{aligned}$$

gebildet. Die zu einem Punkt $B \in \mathcal{S}_C$ gebildete Matrix W in Folgerung 6.35 ist regulär. Auf dem Kompaktum $\mathcal{S}_C$ nimmt $|\det W|$ sein Minimum und Maximum an,

$$0 < \kappa_1 \leq |\det W| \leq \kappa_2.$$

Ebenso sind die Unter-Determinanten von W der Form

$$D\begin{pmatrix} \phi_1, \dots, \phi_n \\ y_0, \dots, y_{j-1}, y_{j+1}, \dots, y_n \end{pmatrix},$$

die in der Cramer'schen Regel zur Berechnung der Lösung von $\alpha W = [0, \ldots, 0, 1]$ auftreten, ungleich Null. Ihr Absolutbetrag nimmt auf dem Kompaktum $\mathcal{S}_C$ sein Minimum $\tau > 0$ an. Damit erhalten wir die globale Abschätzung

$$|\alpha_j| \geq \tau/\kappa_2 =: c > 0$$

für die Koeffizienten α_j, die sich in Folgerung 6.35 für eine Punktmenge mit dem Mindestabstand C wie in $\mathcal{S}_C$ ergeben. $\square$

Zusammenfassung: Die Bestapproximation im diskreten Fall $\#K = n + 1$ wird durch Lösen eines linearen Gleichungssystems bestimmt. Die Matrix dieses Gleichungssystems liefert auch das lineare Funktional aus Satz 6.14(d). Aus dieser Beobachtung lassen sich die Aussagen von Newman, Shapiro und von Freud über die starke Eindeutigkeit und die lokale Lipschitzstetigkeit der metrischen Projektion herleiten. Weiterhin bildet sie die Basis für den Konvergenzbeweis des Remez-Algorithmus.