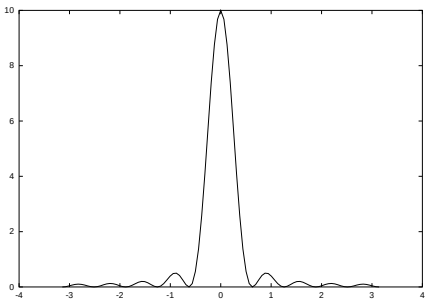
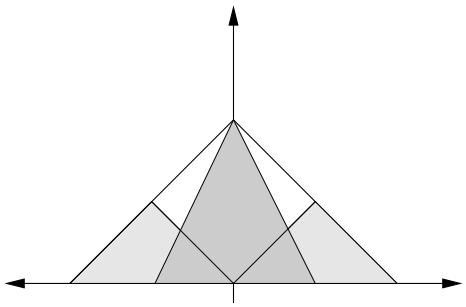


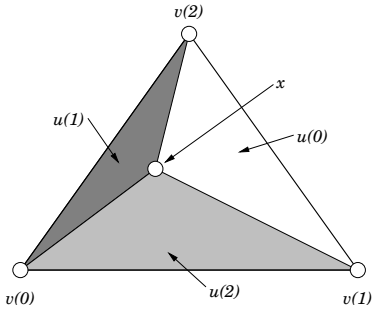


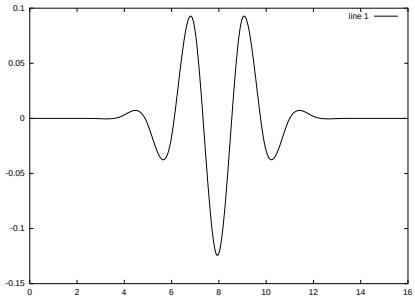
Approximationstheorie

Vorlesung, zuerst gehalten im Wintersemester 2001/2002, überarbeitet im WS 2009/10

Tomas Sauer

Version 2.0
Letzte Änderung: 16.3.2010





Statt einer Leerseite . . . | 0

Es ist eine große Stärkung beim Studieren, wenigstens für mich, alles, was man liest, so deutlich zu fassen, daß man eigne Anwendungen davon oder gar Zusätze dazu machen kann. Man wird am Ende dann geneigt zu glauben, man hätte alles selbst erfinden können, und so was macht Mut. So wie nichts mehr abschreckt als Gefühl von Superiorität im Buch.

Georg Christoph Lichtenberg

Although this may seen as a parodox, all exact science is dominated by the idea of approximation.

Bertrand Russell

0 Inhaltsverzeichnis

1 Was ist Approximationstheorie **3**
1.1 Grundsätzliche Fragen . . . 3
1.2 Ein historisches Beispiel: Summation von Fourierreihen . . . 4
1.3 Fazit . . . 13
1.4 Approximation von Funktionen und Normen . . . 13

2 Polynomapproximation – Dichtheitsaussagen **15**
2.1 Der Satz von Weierstraß . . . 15
2.2 Der Satz von Stone . . . 18
2.3 Der Satz von Bishop . . . 22
2.4 Müntz–Sätze . . . 25

3 Approximation in linearen Räumen **36**
3.1 Approximation durch lineare Räume . . . 37
3.2 Das Kolmogoroff–Kriterium und extremale Signaturen . . . 41
3.3 Haar–Räume und Alternanten . . . 49
3.4 Der Remez–Algorithmus . . . 59

4 Mehr über Bernsteinpolynome **67**
4.1 Ableitungen von Bernsteinpolynomen . . . 67
4.2 Simultanapproximation . . . 69
4.3 Shape preservation . . . 72
4.4 Der Preis: Saturation . . . 75
4.5 Multivariate Bernsteinpolynome . . . 82

5 Approximationsordnung **92**
5.1 Ein Satz von Bernstein . . . 93
5.2 Trigonometrische Polynome I: Stetige Funktionen . . . 96
5.3 Trigonometrische Polynome II: Jackson–Sätze . . . 96
5.4 Trigonometrische Polynome III: Bernstein–Sätze . . . 100
5.5 Trigonometrische Polynome IV: Differenzierbare Funktionen . . . 104
5.6 Trigonometrische Polynome V: Die Zygmund–Klasse . . . 107
5.7 Algebraische Polynome . . . 108

6 Approximation mit translationsinvarianten Räumen **114**
6.1 Translationsinvariante Räume . . . 115
6.2 Ein bißchen Fourieranalysis . . . 119
6.3 Polynomreproduktion und die Strang–Fix–Bedingungen . . . 127

6.4 Approximationsordnung 133

7 Wavelets **138**
7.1 Multiresolution Analysis 138
7.2 Orthogonale Skalierungsfunktionen 143
7.3 Wavelets für orthonormale Skalierungsfunktionen 146
7.4 Approximation mit Wavelets 151

8 Der Satz von Kolmogoroff **158**
8.1 Nomographie, Hilberts 13. Problem und Kolmogoroffs Lösung 158
8.2 Von Würfeln und Intervallen 162
8.3 Der Beweis 167
8.4 Neuronale Netze 176

And first, so that all may understand what is the peril, the tale [. . .] shall be told from the beginning even to this time present.

J. R. R. Tolkien, *The Lord of the Rings*

1 Was ist Approximationstheorie

Bevor wir uns mit der Approximationstheorie und ihren wesentlichen Resultaten beschäftigen, ist es zuerst einmal sinnvoll, uns die wesentlichen Fragestellungen anzusehen, mit denen sich die Approximationstheorie beschäftigt, und zwar nicht nur "abstrakt" theoretisch, sondern vor allem anhand eines klassischen Beispiels, nämlich der (Nicht-)Konvergenz von Fourierreihen.

1.1 Grundsätzliche Fragen

Untersuchungsobjekt der Approximationstheorie ist die Darstellung "komplizierter" Objekte (meist Funktionen) durch einfachere Objekte, die sich mit endlicher Information darstellen lassen. Die Standardsituation ist ein normierter Raum X mit Norm $\|\cdot\|$ und ein meist endlichdimensionaler Teilraum $P \subset X$, dessen Elemente die "einfachen" Funktionen sind.

Beispiel 1.1 *Die beiden gebräuchlichsten Fälle von Approximationsräumen sind*

1. $X = C[a,b]$*, der Vektorraum aller stetigen Funktionen* $f : [a,b] \to \mathbb{R}$ *mit der Norm*

$$\|f\| = \|f\|_\infty = \max_{x\in[a,b]} |f(x)|, \qquad f \in C[a,b],$$

   und $P = \Pi_n$*, der* $(n+1)$*–dimensionale Vektorraum aller* algebraischen Polynome *vom Grad* $\leq n$*.*

2. $X = C(\mathbb{T})$ *der Vektorraum aller stetigen,* 2π*–periodischen Funktionen und*

$$P = P_n = \text{span}\,\{1, \sin x, \cos x, \ldots, \sin nx, \cos nx\}$$

   der $(2n+1)$*–dimensionale Vektorraum der* trigonometrischen Polynome *vom Grad* $\leq n$*.*

Die Fragen, mit denen sich die Approximationstheorie beschäftigt, können nun folgendermaßen formuliert werden:

- Gibt es, zu gegebenem $x \in X$ eine *Bestapproximation* $f^* \in P$, das heißt, eine Funktion, die
$$\|x - f^*\| \leq \|x - f\|, \qquad f \in P, \tag{1.1}$$
erfüllt?

- Wann ist diese Bestapproximation eindeutig?

- Kann man, für Folgen $P_0 \subseteq P_1 \subseteq P_2 \subseteq \cdots$ von Approximationsräumen qualitative und quantitative Aussagen über den Fehler
$$E_n(x) := \inf_{f \in P_n} \|x - f\|, \qquad x \in X,$$
machen, beispielsweise die *Dichtheitsaussage*
$$\lim_{n\to\infty} E_n(x) = 0, \qquad x \in X$$
oder sogar genauere, *quantitative*, Aussagen über die *Ordnung* von $E_n(x)$ in Abhängigkeit von n und x.

- Kann man Funktionen, genauer *Funktionsklassen* $Y \subset X$ durch die *asymptoische* Approximationsordnung, also via
$$Y = Y_\varphi := \left\{ x \in X \ : \ 0 < \lim_{n\to\infty} \frac{E_n(x)}{\varphi(n)} < \infty \right\}, \qquad \varphi : \mathbb{N}_0 \to \mathbb{R},$$
charakterisieren?

- Kann man solche Bestapproximationen *charakterisieren*?

- Kann man solche Bestapproximationne *konstruieren* oder *berechnen*?

Solche Fragen sind nicht nur von rein mathematischem Interesse! Kenntnisse der Struktur und Genauigkeit von Approximationen sind immer dann wichtig, wenn man versucht, eine komplizierte Funktion mittels einfacher "Ansatzfunktionen" dazustellen und solche Probleme sind in vielen Bereichen der angewandten Mathematik üblich und verbreitet. Aber bevor wir uns jetzt der "theoretischen Praxis" widmen und ein bestenfalls akademisches Spielbeispiel generieren, vertagen wir die Anwendungen auf später, wenn wir mehr über Bestapproximation und deren Struktur wissen.

1.2 Ein historisches Beispiel: Summation von Fourierreihen

Die Geburtsstunde der Approximationstheorie läßt sich erfreulich einfach und genau lokalisieren, nämlich auf ein (zu seiner Zeit schockierendes) Gegenbeispiel und eine positive Aussage, die zeigte, in welchem Sinne alles dann doch wieder "funktioniert". Dazu brauchen wir aber zuerst ein bißchen Information über Fourierreihen und deren Konvergenz.

Definition 1.2 *Mit $\mathbb{T}$ bezeichnen wir den eindimensionalen* Torus $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$*, den wir normalerweise durch die Intervalle $[0, 2\pi]$ oder $[-\pi, \pi]$ darstellen werden, wobei die Endpunkte miteinander identifiziert werden.*

Nur um das klarzustellen: $\mathbb{T}$ ist *nicht* das Intervall $I = [0, 2\pi]$ oder das Intervall $I = [-\pi, \pi]$, sondern eines dieser Intervalle mit einer Addition *modulo* 2π, das heißt, für $x, y \in I$ ist $x + y$ derjenige Wert z, so daß

$$z - (x + y) \in 2\pi\,\mathbb{Z}.$$

Auf diese Art und Weise erhält $\mathbb{T}$ nämlich eine additive Gruppenstruktur, die das Intervall nicht hat.

Definition 1.3 *Mit $C(\mathbb{T})$ bezeichnen wir den Vektorraum der stetigen reellwertigen*[1] *Funktionen auf $\mathbb{T}$, mit der Norm*

$$\|f\| = \|f\|_\infty := \max_{x\in\mathbb{T}} |f(x)|\,,$$

den wir wegen der obigen Bemerkung auch mit

$$C_{2\pi}(\mathbb{R}) := \{f \in C(\mathbb{R}) \;:\; f\,(x + 2\pi) = f(x),\, x \in \mathbb{R}\}\,,$$

dem Vektorraum der stetigen, 2π–periodischen Funktionen auf $\mathbb{R}$ mit der Norm

$$\|f\| = \|f\|_\infty := \max_{x\in\mathbb{T}} |f(x)| = \sup_{x\in\mathbb{R}} |f(x)|$$

identifizieren können.

Für die Funktionen in $C(\mathbb{T})$ gibt es nun eine ganz besondere klassische Darstellung, nämlich die *Fourierreihen*, die von Fourier[2] zur Behandlung der Wärmeleitungsgleichung entwickelt wurden. Wir definieren sie hier nur für stetige Funktionen, ein Großteil des Interesses an und Spaßes mit Fourierreihen beruht aber mit Sicherheit auch auf der Behandlung anderer, beispielsweise quadratintegrierbarer Funktionen. Eine schöne und einfache Einführung in die Fourieranalysis sind [28, 35].

Definition 1.4 *Zu einer Funktion $f \in C(\mathbb{T})$ definiert man die reellen bzw. komplexen* Fourierkoeffizienten

$$a_k = a_k(f) := \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\cos kt\,dt, \qquad b_k = b_k(f) := \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\sin kt\,dt, \qquad k \in \mathbb{N}_0, \tag{1.2}$$

[1] Im wesentlichen werden wir uns in dieser Vorlesung mit *reeller* Approximationstheorie befassen, weswegen alle Funktionen normalerweise als reellwertig angenommen werden. Das heißt aber nicht, wie wir gleich sehen werden, daß komplexe Zahlen tabu sein sollen.

[2] Jean Baptiste Fourier, 1768–1830, französicher Mathematiker und Politiker. Er war nicht nur Mitglied der "Acadèmie des Sciences", sondern (vorher) auch Teilnehmer an der Ägypten–Expedtion von Napoleon (Bonaparte) als wissenschaftlicher Berater und Gouverneur des Departement Isére mit Hauptstadt Grenoble. In den beiden letzteren Eigenschaften trug er nicht unwesentlich (als Förderer von Champollion) zur Entzifferung der Hieroglyphen bei, siehe [20].

bzw.

$$c_k = c_k(f) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\, e^{-ikt}\, dt, \qquad k \in \mathbb{Z}, \tag{1.3}$$

und die reelle bzw. komplexe Fourierreihe *als*

$$\mathscr{F}(f)(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) = \sum_{k=-\infty}^{\infty} c_k\, e^{ikx}, \qquad x \in \mathbb{T}. \tag{1.4}$$

Übung 1.1 Zeigen Sie, daß die reelle und die komplexe Darstellung in (1.4) übereinstimmen. ♢

Eigentlich sind solche Reihen ja nur sinnvoll, wenn sie konvergieren, sei es nun *punktweise* für ein festes $x \in \mathbb{T}$ oder *gleichmäßig*, das heißt in der Norm von $C(\mathbb{T})$. Nur leider sind Fourierreihen in dieser Hinsicht ziemlich zickig: Es ist und war wohlbekannt, daß Stetigkeit *nicht* notwendig für die Konvergenz der Fourierreihen ist, und es war Du Bois–Reymond[3], der 1873 für einen Schock in der Analysis sorgte, als er ein Beispiel einer stetigen Funktion angab, deren Fourierreihe an einer Stelle divergiert. Bis dahin wurden nämlich Reihen generell mehr oder weniger blauäugig verwendet[4], das heißt, lediglich formal und ohne sich um Konvergenz zu kümmern – ein durchaus nicht ungebräuchliches Vorgehen, das man beispielsweise auch bei Gauß [24] findet.
Um dieses Resultat "vernünftig" zu formulieren, definieren wir die n–te *Partialsumme* der Fourierreihe von $f \in C(\mathbb{T})$ als das trigonometrische Polynom

$$\sigma_n(f)(x) := \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) = \sum_{k=-n}^{n} c_k\, e^{ikx}, \qquad x \in \mathbb{T}, \tag{1.5}$$

und erhalten das folgende Resultat.

Satz 1.5 *Es gibt eine Funktion $f \in C(\mathbb{T})$, so daß*

$$\lim_{n\to\infty} \|f - \sigma_n(f)\| = \infty. \tag{1.6}$$

Bevor wir diesen Satz beweisen, erzählen wir aber erst unsere Geschichte fertig. Es ist also im allgemeinen nicht möglich, eine Funktion $f \in C(\mathbb{T})$ durch ihre Partialsummen darzustellen – genauer, möglich ist es schon, nur nicht sonderlich sinnvoll, denn dieser Prozess liefert am Ende nicht die Funktion f, sondern irgendwas anderes. In diesem Zusammenhang war es dann

[3] Paul David Gustav Du Bois–Reymond, 1831–1889, deutscher Mathematiker; sein Bruder Emil war einer der bekanntesten Physiologen seiner Zeit. Studium der Mathematik in Berlin und der Medizin in Zürich, Professuren in Heidelberg, Freiburg, Tübingen und Berlin, Erfinder des Begriffs "Integralgleichung".

[4] Man erinnere sich nur an den Abelschen Beweis, daß der Grenzwert der alternierenden Reihe $\sum_j (-1)^j$ gerade $\frac{1}{2}$ ist.

schon tröstlich, daß Weierstraß[5] in 1885 [85] zeigen konnte, daß, auch wenn es mit den Partiasummen nicht funktioniert, man jede Funktion $f \in C(\mathbb{T})$ *beliebig genau* durch trigonometrische Polynome annähern, also *approximieren* kann. Und schon war die Approximationstheorie geboren[6].

Definition 1.6 *Ein* trigonometrisches Polynom *der Ordnung* $n \in \mathbb{N}_0$ *ist eine Funktion* $f \in C(\mathbb{T})$ *der Form*

$$f(x) = \frac{a_0}{2} + \sum_{j=1}^{n} a_j \cos jx + b_j \sin jx, \qquad x \in \mathbb{T}.$$

Satz 1.7 *(Weierstraß'scher Approximationssatz für trigonometrische Polynome)*
Zu jeder Funktion $f \in C(\mathbb{T})$ *und jedem* $\varepsilon > 0$ *gibt es ein trigonometrisches Polynom* p*, so daß*

$$\|f - p\| < \varepsilon. \tag{1.7}$$

So, jetzt aber an die Arbeit! Wir müssen diese beiden Sätze natürlich auch beweisen und dafür brauchen wir noch ein bißchen Terminologie, nämlich den Begriff der Faltung und des Kerns.

Definition 1.8 *(Faltungen und Kerne)*

1. *Für* $f, g \in C(\mathbb{T})$ *ist die* Faltung $f * g$ *definiert als*

$$f * g := \int_{-\pi}^{\pi} f(t)\, g(\cdot - t)\, dt = \int_{-\pi}^{\pi} f(\cdot - t)\, g(t)\, dt. \tag{1.8}$$

2. *Eine Funktion*[7] $K \in C(\mathbb{T})$ *heißt* Kern *zu einem (Faltungs-)Operator* κ *wenn*

$$\kappa(f) = f * K, \qquad f \in C(\mathbb{T}).$$

3. *Der Kern* D_n *zu dem Operator*

$$\sigma_n(f) = f * D_n \qquad n \in \mathbb{N}_0, \quad f \in C(\mathbb{T}),$$

heißt n*–ter* Dirichlet[8]–Kern.

[5] Karl Weierstraß, 1815–1897, gilt auch als Erfinder der "Epsilontik", also der Form, wie Analysis heute normalerweise präsentiert wird: "Sei $\varepsilon > 0 \ldots$"

[6] Na gut, so einfach ist's natürlich nicht! Immerhin hat Tchebycheff schon Bestapproximationen berechnet, bevor der Approximationssatz von Weierstraß veröffentlicht wurde.

[7] Wir haben es hier nur mit stetigen Kernen zu tun. Stetigkeit ist aber *nicht* essentiell und kann beispielsweise zu Betragsintegrierbarkeit abgemildert werden.

[8] Johann Peter Gustav Lejeune Dirichlet, 1805–1859, deutsch–französischer Mathematiker belgischer Abstammung, Schwager von Felix Mendelssohn.

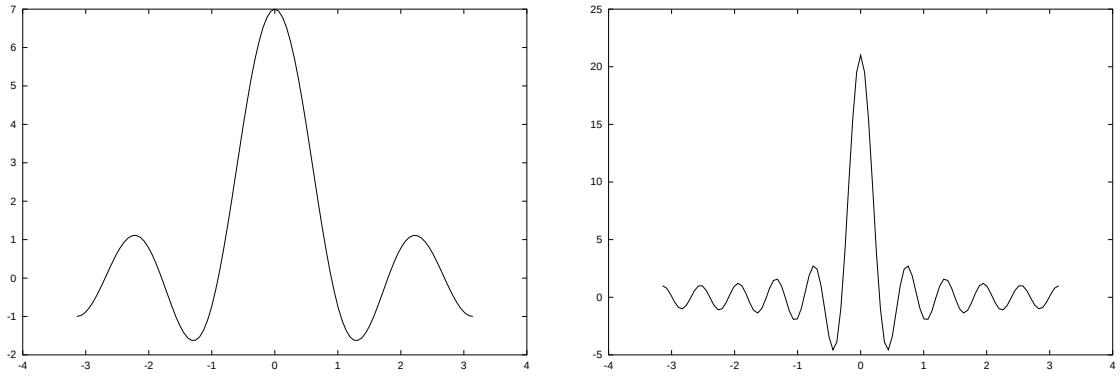


Abbildung 1.1: Die beiden Dirichletkerne D_3 und D_{10}. Man sieht sehr schön, daß diese Kerne stetig sind und ihr Maximum an der Stelle $t = 0$ annehmen.

Jetzt beginnen wir mit einem *modernen* Beweis von Satz 1.5, der auf Funktionalanalysis, genauer, auf dem *Uniform Boundedness Principle*[9] beruht.

Beweis von Satz 1.5: Wir werden zeigen, daß die Operatornorm

$$\|\sigma_n\| = \sup_{f \neq 0} \frac{\|\sigma_n(f)\|}{\|f\|}$$

divergiert, das heißt, daß

$$\lim_{n\to\infty} \|\sigma_n\| = \infty \tag{1.9}$$

ist. Daraus folgt sofort mit dem *Uniform Boundedness Principle*, siehe [43, 4.7-3, p. 249], daß es eine Funktion $f \in C(\mathbb{T})$ geben muß, so daß

$$\lim_{n\to\infty} \|\sigma_n(f)\| = \infty$$

und somit folgt, da $\|f\| < \infty$, die Gültigkeit von (1.6).
Um (1.9) zu beweisen, bemerken wir zuerst, daß der Dirichletkern als

$$D_n(t) = \frac{1}{2\pi} \frac{\sin\left(n + \frac{1}{2}\right) t}{\sin \frac{1}{2} t}, \qquad n \in \mathbb{N}_0, \tag{1.10}$$

geschrieben werden kann, was übrigens auch zeigt, daß $D_n \in C(\mathbb{T})$ ist. Da D_n obendrein symmetrisch ist, ist für beliebiges $f \in C(\mathbb{T})$

$$\|\sigma_n(f)\| \geq |\sigma_n(f)(0)| = |f * D_n(0)| = \left| \int_{-\pi}^{\pi} f(t) \underbrace{D_n(0-t)}_{=D_n(t)} dt \right| = \left| \int_{-\pi}^{\pi} f(t) D_n(t) dt \right|$$

[9] Ist eine Folge T_n von linearen Operatoren *punktweise* beschränkt, das heißt, ist $\sup_n \|T_n f\| < \infty$ für alle f, dann ist die Folge auch global beschränkt: $\sup_n \|T_n\| < \infty$.

Man sieht sofort aus (1.10), daß

$$D_n(t) = 0 \qquad \Longleftrightarrow \qquad t \in \left\{ \pm \frac{k\pi}{n+\frac{1}{2}} \ : \ k = 1, \ldots, n \right\},$$

was es nahelegt, Funktionen $f_h \in C(\mathbb{T})$, $h > 0$, zu definieren, die die Forderungen $\|f_h\| = 1$ und

$$f_h(t) = \operatorname{sgn} D_n(t), \quad t \in \left[\frac{k\,\pi}{n+\frac{1}{2}} + h, \frac{(k+1)\,\pi}{n+\frac{1}{2}} - h \right], \qquad k = -n, \ldots, n-1,$$

erfüllt. Für diese Funktionen ist dann

$$|\sigma_n(f_h)(0)| = \int_{-\pi}^{\pi} |D_n(t)| \, dt + g(h), \qquad |g(h)| \le 4\,(2n+1)\,h\,\|D_n\|,$$

woraus also

$$\|\sigma_n\| \ge \lim_{h\to 0} \frac{\|\sigma_n(f_h)\|}{\|f_h\|} = \lim_{h\to 0} \|\sigma_n(f_h)\| \ge \lim_{h\to 0} |\sigma_n(f_h)(0)| = \int_{-\pi}^{\pi} |D_n(t)| \, dt \tag{1.11}$$

folgt; es gilt sogar Gleichheit, aber das soll uns hier nicht interessieren. Unter Verwendung der Identität $|\sin t| < t$, $t \in (0, 2\pi]$, erhalten wir jetzt nämlich nach Einsetzen von (1.10) in (1.11), daß

$$\begin{aligned}
\|\sigma_n\| &\ge \int_{-\pi}^{\pi} |D_n(t)\,dt| = \int_0^{2\pi} \left| \frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{1}{2}t} \right| dt > 2 \int_0^{2\pi} \left| \frac{\sin\left(n+\frac{1}{2}\right)t}{t} \right| dt \\
&= 2 \int_0^{(2n+1)\pi} \frac{|\sin v|}{v}\, dv = 2 \sum_{k=0}^{2n} \int_{k\pi}^{(k+1)\pi} \frac{|\sin v|}{v}\, dv \ge 2 \sum_{k=0}^{2n} \frac{1}{(k+1)\pi} \underbrace{\int_{k\pi}^{(k+1)\pi} |\sin v| \, dv}_{=|\cos(k+1)\pi - \cos k\pi| = 2} \\
&= \frac{4}{\pi} \sum_{k=0}^{2n} \frac{1}{k+1}
\end{aligned}$$

und das divergiert natürlich für $n \to \infty$, womit (1.9) und damit Satz 1.5 bewiesen ist. □

Übung 1.2 Beweisen Sie die Formel (1.10). ♢

Für den Beweis von Satz 1.7 greifen wir sogar auf eine *konstruktive* Idee zurück, die auf Fejér[10] [21] zurückgeht, der übrigens auch in [22] Beispiele für stetige Funktionen mit divergenter Fourierreihe gab. Denn man kann aus den Partialsummen auf relativ einfache Art einen

[10] Lipót Fejér, 1880–1959, geboren als Leopold Weiss, hungarisierte seinen Namen um 1900. Nach dem Studium der Mathematik in Berlin und Budapest leistete er wesentliche Beiträge zur Funktionalanalysis und Fourieranalysis.

konvergenten Approximationsprozess “basteln”, indem man die *Fejérschen Mittel*

$$\varphi_n(f) := \frac{1}{n+1}\sum_{k=0}^{n}\sigma_k(f), \qquad n \in \mathbb{N}_0,$$

als *Mittelwert* der Partialsummen einführt; den zugehörigen *Fejér–Kern* bezeichnen wir mit F_n. Diese trigonometrischen Polynome konvergieren nun gleichmäßig gegen f und wir erhalten das folgende Resultat, aus dem Satz 1.7 unmittelbar folgt.

Satz 1.9 *(Konvergenz der Fejérschen Mittel)*
Für $f \in C(\mathbb{T})$ ist

$$\lim_{n\to\infty} \|f - \varphi_n(f)\| = 0. \tag{1.12}$$

Der Beweis von Satz 1.9 ist wieder eine Konsequenz aus einem allgemeineren Prinzip, das wir gleich kennenlernen werden. Zuerst bemerken wir, daß auch die Fejérkerne eine explizite Darstellung[11] haben, nämlich

$$F_n(t) = \frac{1}{2\pi}\frac{1}{n+1}\left(\frac{\sin\frac{n+1}{2}t}{\sin\frac{1}{2}t}\right)^2. \tag{1.13}$$

Übung 1.3 Beweisen Sie die Formel (1.13), möglicherweise unter Verwendung von Übung 1.2. ♢

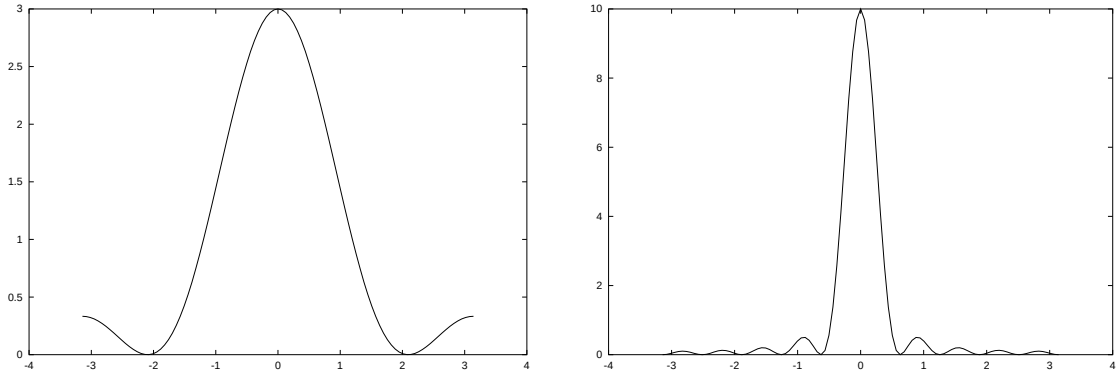


Abbildung 1.2: Die beiden Fejérkerne F_3 und F_{10}. Diese Kerne klingen sehr viel schöner ab als die Dirichletkerne aus Abb. 1.1, haben aber ansonsten auch ihr Maximum an der Stelle $t = 0$.

[11] Die Theorie der speziellen Funktionen und, spezieller, die *Summationstheorie* beschäftigt sich oftmals mit der Herleitung von geschlossenen Formen solcher Kerne, nur treten dabei oftmals nicht nur trigonometrische, sondern auch hypergeometrische Funktionen auf.

Eine wichtige und fast nicht zu überschätzende Eigenschaft von F_n, die aus (1.13) folgt, ist die *Positivität* des Kerns F_n und damit auch die Positivitäte des Operators φ_n, das heißt,

$$f \geq 0 \qquad \Longrightarrow \qquad \varphi_n(f) \geq 0.$$

Bemerkung 1.10 *Ein kurzer Vergleich der Kerne D_n und F_n aus Abb. 1.1 und Abb. 1.2 ist ganz interessant: Beide Kerne haben ihr Maximum an $t = 0$, beide Kerne klingen zum Rand hin ab, beide Kerne haben die Eigenschaft, daß ihr Integral den Wert 1 hat, aber trotzdem oszilliert D_n so heftig, daß*

$$\int_{-\pi}^{\pi} |D_n(t)| \, dt \to \infty \qquad \textit{während} \qquad \int_{-\pi}^{\pi} |F_n(t)| \, dt = \int_{-\pi}^{\pi} F_n(t) \, dt = 1$$

ist. Was Positivität doch so alles ausmachen kann.

Eine besonders wichtige Klasse positiver Kerne wird durch die folgende Definition gegeben.

Definition 1.11 *Eine Folge $K_n \in C(\mathbb{T})$, $n \in \mathbb{N}_0$, von Kernen heißt* approximative Identität, *wenn sie die folgenden Bedingungen erfüllt:*

1. *(Positivität) $K_n \geq 0$, $n \in \mathbb{N}_0$,*
2. *(Normiertheit)*
$$\int_{-\pi}^{\pi} K_n(t) \, dt = 1, \qquad n \in \mathbb{N}_0,$$
3. *(Lokalität) Für jedes $\delta > 0$ ist*
$$\lim_{n\to\infty} \max_{|t|>\delta} K_n(t) = 0.$$

Und in der Tat sind approximative Identitäten approximative Identitäten in dem Sinn, daß sie die Identität in der Operatornorm approximieren, daß also für die zugehörigen Operatoren κ_n, definiert durch $\kappa_n(f) = f * K_n$, die Grenzertaussage

$$\lim_{n\to\infty} \kappa_n = I$$

in der Operatornorm gilt, was wir auch wie folgt formulieren können.

Satz 1.12 *Sei K_n, $n \in \mathbb{N}_0$, eine approximative Indentität. Dann ist, für jedes $f \in C(\mathbb{T})$,*

$$\lim_{n\to\infty} \|\kappa_n(f) - f\| = 0. \tag{1.14}$$

Beweis: Jede Funktion $f \in C(\mathbb{T})$ kann mit einer Funktion $f \in C\,[-\pi,\pi]$ mit der zusätzlichen Eigenschaft $f(-\pi) = f(\pi)$ identifiziert werden und ist daher nicht nur stetig, sondern *gleichmäßig stetig*.
Sei nun $x \in [-\pi,\pi]$ vorgegeben. Dann ist, für jedes $n \in \mathbb{N}_0$ und jedes $\delta > 0$

$$
\begin{aligned}
|\kappa_n(f)(x) - f(x)| &= \Big| \underbrace{\int_{-\pi}^{\pi} f(x-t)\,K_n(t)\,dt}_{=f*K_n} - f(x) \underbrace{\int_{-\pi}^{\pi} K_n(t)\,dt}_{=1} \Big| \\
&= \left| \int_{-\pi}^{\pi} (f(x-t) - f(x))\; K_n(t)\,dt \right| \le \int_{-\pi}^{\pi} |f(x-t) - f(x)|\; K_n(t)\,dt \\
&= \int_{|t|\le\delta} |f(x-t) - f(x)|\; K_n(t)\,dt + \int_{|t|\ge\delta} |f(x-t) - f(x)|\; K_n(t)\,dt \\
&\le \sup_{|t|\le\delta} |f(x-t) - f(x)| \underbrace{\int_{-\pi}^{\pi} K_n(t)\,dt}_{=1} + \max_{|t|\ge\delta} K_n(t) \underbrace{\int_{-\pi}^{\pi} |f(x-t) - f(x)|\; dt}_{\le 4\pi\,\|f\|} \\
&\le \sup_{|t|\le\delta} |f(x-t) - f(x)| + 4\pi\,\|f\| \max_{|t|\ge\delta} K_n(t).
\end{aligned}
$$

Für ein vorgegebenes $\varepsilon > 0$ können wir nun, wegen der *gleichmäßigen* Stetigkeit von f ein $\delta > 0$ wählen, so daß $|f(x-t) - f(x)| \le \frac{\varepsilon}{2}$, und zwar *unabhängig von* x und dann wählen wir, wegen der Lokalität von K_n, den Index n *in Abhängigkeit von* δ so groß, daß

$$\max_{|t|\ge\delta} K_n(t) < \frac{\varepsilon}{8\pi\|f\|}.$$

Damit gibt es zu jedem $\varepsilon > 0$ einen Index $n_0 \in \mathbb{N}_0$, so daß

$$\|\kappa_n(f) - f\| < \varepsilon, \qquad n \ge n_0,$$

womit (1.14) bewiesen ist. □

Beweis von Satz 1.9: Mit Satz 1.12 wird die Sache nun ziemlich einfach: Wir müssen nur noch nachweisen, daß die Fejérkerne eine approximative Identität bilden. Und in der Tat ist die Positivität offensichtlich aus (1.13), die Normiertheit folgt aus $\sigma_k(1) = 1$, $k \in \mathbb{N}_0$, denn deswegen gilt

$$1 = \sigma_n(1) = \frac{1}{n+1} \sum_{k=0}^{n} \sigma_k(1) = \varphi_n(1) = 1 * K_n = \int_{-\pi}^{\pi} K_n(t)\,dt.$$

Da die Funktion $\sin t$ auf dem Intervall $\left[0, \frac{\pi}{2}\right]$ monoton steigend ist, ist

$$\min_{t\in[-\pi,-\delta]\cup[\delta,\pi]} \left| \sin\frac{1}{2}t \right| = \sin\frac{\delta}{2}$$

und somit

$$F_n(t) \leq \frac{1}{n}\left(\frac{1}{\sin\frac{\delta}{2}}\right)^2, \qquad |t| \geq \delta,$$

was für $n \to \infty$ gegen 0 konvergiert, womit auch die Lokalität erledigt ist. □

1.3 Fazit

Das letzte Kapitel hat uns schon einiges über grundlegende Ansätze und Fragestellungen der Approximationstheorie verraten und uns

- Dichtheitsaussagen
- die Konstruktion eines Approximationsprozesses
- eine Klasse von "guten" Approximationsprozessen

geliefert, wenn auch nicht für algebraische, sondern für trigonometrische Polynome. Und tatsächlich (und das werden wir noch merken) sind solche Aussagen für trigonometrische Polynome oftmals einfacher und zwar aus einem einfachen Grund:

> Der Torus $\mathbb{T}$ hat, im Gegensatz zu Intervallen wie $[-\pi, \pi]$, keinen Rand!

Doch dazu später noch mehr.

1.4 Approximation von Funktionen und Normen

Die Unterräume, mit denen man approximiert, hängen natürlich stark von dem Definitionsbereich des zugrundeliegenden Funktionenraums ab, der obendrein ein *Banachraum*, das heißt, vollständig unter der Norm sein sollte.

Übung 1.4 Zeigen Sie: Der Raum $C^1(\mathbb{T}) \subset C(\mathbb{T})$ aller stetig differenzierbaren Funktionen auf $\mathbb{T}$ ist *kein* Banachraum, wenn man die Supremumsnorm verwendet. ♢

Typische Banachräume, in denen man Approximationsprobleme betrachten kann sind

- die Räume
$$L_p(X) := \left\{ f \,:\, X \to \mathbb{R} \,:\, \int_X |f|^p \, dt < \infty \right\}$$
unter Verwendung der Normen
$$\|f\|_p := \left(\int_X |f(t)|^p \, dt \right)^{1/p}, \qquad 1 \leq p < \infty$$

- der Raum $C(X)$ der stetigen oder $C_u(X)$, der Raum der *gleichmäßig* stetigen Funktionen[12], die gleichmäßig beschränkt sind[13] mit der Norm

$$\|f\|_\infty = \sup_{x\in X} |f(x)| .$$

Typische (univariate) Beispiele für X und die zugehörigen Approximationsräume sind

- $X = \mathbb{T}$, dann verwendet man zumeist *trigonometrische* Polynome eines bestimmten Grades.
- $X = I$, wobei $I \subset \mathbb{R}$ ein kompaktes Intervall ist, dann verwendet man zumeist *algebraische* Polynome.
- $X = \mathbb{R}$, dann muß man sich Exponentialfunktionen der Form $f_\lambda(x) = e^{-\lambda x}$ zuwenden.

Etwas überspitzt kann man sagen, daß es eigentlich nur drei "relevante" Werte von p für die L_p–Räume gibt, nämlich

$p = 1$**:** Hier spricht man von L_1–Approximation, die benötigten Methoden sind zumeist der Optimierung entliehen.

$p = 2$**:** Bei der L_2–Approximation hat man es mit der "wohlbekannten" Hilbertraumtheorie zu tun (die Norm wird durch ein inneres Produkt gegeben) und man kann mit Orthogonalreihen arbeiten. Beispielsweise ist die Bestapproximation zu $f \in L_2(\mathbb{T})$ mit trigonometrischen Polynomen vom Grad $\leq n$ immer die Partialsumme $\sigma_n(f)$.

$p = \infty$**:** Die *Tschebyscheff*[14]*–Approximation*, mit der wir uns zuerst einmal befassen wollen, ist vielleicht das "eigenständigste" Produkt der Approximationstheorie. Wir werden sehen daß hier, nicht ohne Grund, die beiden Fälle $X = \mathbb{T}$ und $X = I$ eine wesentliche Rolle spielen.

Übung 1.5 Beweisen oder widerlegen Sie die folgenden Aussagen:

1. Die trigonometrischen Polynome sind dicht in $L_\infty(\mathbb{T})$.
2. Die trigonometrischen Polynome sind dicht in $C(I)$, $I \subseteq (-\pi, \pi)$ kompaktes Intervall. Wie sieht es mit $I = [-\pi, \pi]$ aus?
3. Die algebraischen Polynome sind dicht in $C_u(\mathbb{R})$.

$\diamondsuit$

[12] Achtung! Dies ist etwas anderes als der Raum $L_\infty(X)$ aller wesentlich beschränkten Funktionen, der etwas pathologische Eigenschaften hat.

[13] Dies wird im Falle $X = \mathbb{R}$ wichtig.

[14] Pafutny Tschebyscheff (oder Chebyshev oder Chebycheff oder Chebychov oder ..., je nach Transkription), 1821–1894, reiner (Primzahlsätze) und angewandter (Mechanik) Mathematiker in St. Petersburg.

Sunt qui quicquid in libris scriptum domi habent, noscere sibi videantur; cumque ulla de re mentio incidit, "hic liber", inquiunt, "in armario meo est" ...
Manche halten sich für Kenner all dessen, was in den Büchern ihrer Hausbibliothek steht; wird irgendeine Sache erwähnt, sagen sie: "Dieses Buch steht in meinem Schrank" ...

Petrarca, *De librorum copia – Von der Bücherfülle*, 1366

2 Polynomapproximation – Dichtheitsaussagen

In diesem Kapitel befassen wir uns mit Approximation durch *algebraische* Polynome auf einem *kompakten* Intervall, genauer, mit *Tschebyscheffapproximation*, also die Approximation von stetigen Funktionen bezüglich der Supremumsnorm.

2.1 Der Satz von Weierstraß

Jetzt also zu "dem" Satz der Approximationstheorie, dem Satz von Weierstraß für algebraische Polynome. Dazu erst mal ein klein wenig Notation.

Definition 2.1 *Mit $\Pi = \mathbb{R}[x]$ bezeichnen wir die Algebra der Polynome mit reellen Koeffizienten und mit*

$$\Pi_n = \operatorname{span}_{\mathbb{R}} \left\{ x^k \ : \ k = 0, \ldots, n \right\}, \qquad n \in \mathbb{N}_0,$$

den Untervektorraum der Polynome vom Grad $\leq n$.

Satz 2.2 (Satz von Weierstraß für algebraische Polynome) *Sei $I \subset \mathbb{R}$ ein kompaktes Intervall. Zu jeder Funktion $f \in C(I)$ und jedem $\varepsilon > 0$ gibt es ein Polynom $p \in \Pi$, so daß $\|f - p\| < \varepsilon$.*

Zu diesem Satz gibt es jede Menge Beweise, mehr oder weniger lange und komplizierte, elementare und nicht–elementare. Ein nettes Beispiel ist [44], wo der Satz von Weierstraß auf die Approximation von *Treppenfunktionen* durch Polynome zurückgeführt wird, siehe Übung 2.5. Hier allerdings gönnen wir uns den "Klassiker", nämlich den ersten *konstruktiven* Beweis des

Satzes von Weierstraß für algebraische Polynome[15], der 1912 von Bernstein[16] [8] gegeben wurde.

Definition 2.3 *Für $n \in \mathbb{N}_0$ und $0 \leq j \leq n$ ist das j–te* Bernstein–(Bézier[17])–Basispolynom *definiert als*

$$B_j^n(x) = \binom{n}{j} x^j (1-x)^{n-j}$$

und zu einer Funktion $f \in C[0,1]$ das n–te Bernsteinpolynom *als*

$$B_n f = \sum_{j=0}^{n} f\left(\frac{j}{n}\right) B_j^n. \tag{2.1}$$

Beweis von Satz 2.2: Wir können zuerst einmal ohne Einschränkung annehmen, daß $I = [0,1]$ ist, denn jedes andere Intervall läßt sich mit einer einfachen affinen Transformation auf diese Gestalt bringen und die Supremumsnorm läßt sich von solchen Transformationen ohnehin nicht beeindrucken. Es wird nicht weiter überraschen, daß wir beweisen, daß für jede stetige Funktion $f \in C(I)$ die Bernsteipolynome $B_n f$ für $n \to \infty$ gleichmäßig gegen f konvergieren.
Dazu bemerken wir zuerst, daß

$$\sum_{j=0}^{n} B_j^n(x) = \sum_{j=0}^{n} \binom{n}{j} x^j (1-x)^{n-j} = (x+1-x)^n = 1, \tag{2.2}$$

sowie

$$\begin{aligned}
\sum_{j=0}^{n} \frac{j}{n} B_j^n(x) &= \sum_{j=0}^{n} \frac{j}{n} \frac{n!}{j!(n-j)!} x^j (1-x)^{n-j} = \sum_{j=1}^{n} \frac{(n-1)!}{(j-1)!(n-j)!} x^j (1-x)^{n-j} \\
&= x \sum_{j=1}^{n} \frac{(n-1)!}{(j-1)!(n-j)!} x^{j-1} (1-x)^{n-j} = x \sum_{j=0}^{n-1} B_j^{n-1}(x) = x
\end{aligned}$$

und

$$\begin{aligned}
\sum_{j=0}^{n} \frac{j(j-1)}{n^2} B_j^n(x) &= \sum_{j=0}^{n} \frac{j(j-1)}{n^2} \frac{n!}{j!(n-j)!} x^j (1-x)^{n-j} \\
&= \frac{n-1}{n} \sum_{j=2}^{n} \frac{(n-2)!}{(j-2)!(n-j)!} x^j (1-x)^{n-j} = \frac{n-1}{n} x^2 \sum_{j=0}^{n-2} B_j^{n-2}(x) \\
&= \frac{n-1}{n} x^2.
\end{aligned}$$

[15] Und in diesem Fall hat man leider keine Fourierreihen zur Verfügung.

[16] Sergi Bernstein, 1880–1968, promovierte an der Sorbonne in Paris, Beiträge zur Numerischen Mathematik wie auch zur Stochastik (auch Anwendungen in der Genetik!).

[17] Pierre Bézier, 1910-1999, französicher Ingenieur, bei Renault für die Entwicklung von numerischen Systemen zur Modellierung und Behandlung von Flächen zuständig. Überraschenderweise weist Bźier darauf hin, daß er ziemlich genau 30 Jahre nach Bernstein an der "SupElec" (Ecole Supérieure d'Electricité) studierte und daß Bernstein seine Bernsteinpolynome entwickelte, um für eine Versicherungsgesellschaft Wahrscheinlichkeitsverteilungen zu plotten – dies ist allerdings eine unbewiesene Anekdote. Weitere Details in [59].

Mit diesen drei Identitäten erhalten wir, daß für jedes $x \in [0,1]$

$$\begin{aligned}\sum_{j=0}^{n}\left(\frac{j}{n}-x\right)^2 B_j^n(x) &= \sum_{j=0}^{n}\left(\frac{j^2}{n^2}-2\frac{j}{n}x+x^2\right)B_j^n(x)\\ &= \sum_{j=0}^{n}\left(\frac{j(j-1)}{n^2}+\frac{j}{n^2}-2\frac{j}{n}x+x^2\right)B_j^n(x)\\ &= \frac{n-1}{n}x^2+\frac{1}{n}x-2x^2+x^2=\frac{1}{n}\,x(1-x)\end{aligned}$$

Diese Formel erlaubt es uns nun zu zeigen, daß die Basispolynome recht gut lokalisiert sind: für $\delta > 0$ und $x \in [0,1]$ ist

$$\sum_{\left|\frac{j}{n}-x\right|\geq\delta} B_j^n(x) \leq \frac{1}{\delta^2}\sum_{j=0}^{n}\left(\frac{j}{n}-x\right)^2 B_j^n(x) = \frac{x(1-x)}{n\delta^2} \leq \frac{1}{4n\delta^2}. \tag{2.3}$$

Das war's dann auch schon fast! Sei nun $f \in C[0,1]$, dann ist f ja nicht nur stetig, sondern sogar *gleichmäßig stetig*, das heißt, für alle $\varepsilon > 0$ gibt es ein $\delta > 0$ so daß

$$|x-y| < \delta \qquad \Longrightarrow \qquad |f(x)-f(y)| < \frac{\varepsilon}{2}.$$

Sei also $\varepsilon > 0$ vorgegeben und $\delta > 0$ passend gewählt. Dann ist, für beliebiges $x \in [0,1]$,

$$\begin{aligned}|f(x)-B_nf(x)| &= \left|\sum_{j=0}^{n}(f(x)-f(j/n))\ B_j^n(x)\right| \leq \sum_{j=0}^{n}|f(x)-f(j/n)|\ B_j^n(x)\\ &= \sum_{\left|\frac{j}{n}-x\right|<\delta}\underbrace{|f(x)-f(j/n)|}_{\leq\varepsilon/2}\ B_j^n(x) + \sum_{\left|\frac{j}{n}-x\right|\geq\delta}\underbrace{|f(x)-f(j/n)|}_{\leq 2\|f\|_I}\ B_j^n(x)\\ &\leq \frac{\varepsilon}{2}\sum_{j=0}^{n}B_j^n(x) + \frac{\|f\|_I}{2n\delta^2} = \frac{\varepsilon}{2}+\frac{\|f\|_I}{2n\delta^2}.\end{aligned}$$

Die rechte Seite ist nun unabhängig von x und $< \varepsilon$ sobald $n > \varepsilon^{-1}\ \delta^{-2}\ \|f\|_I$. □

Bemerkung 2.4 *1. Eigentlich sind ja die Bernsteinpolynome auch wieder eine approximative Identität, nur eben eine "diskrete". Positivität haben wir nach wie vor, die Normiertheit wird von einer Integralbedingung zur Summenbedingung (2.2) und die Lokalität finden wir in (2.3).*

2. Man sieht, daß die Bersteinpolynome umso schneller konvergieren, je "stetiger" die Funktion f ist. Schließlich braucht man ja zu einem $\varepsilon > 0$ zuerst einmal den "Lokalisierungsparameter" δ und n bestimmt man dann so, daß $n > \frac{\|f\|}{\varepsilon\delta^2}$.

2.2 Der Satz von Stone

Nun haben wir also zwei Sätze, nämlich Satz 1.7 und Satz 2.2, die etwas über die Dichtheit von Polynomen in dem Banachraum der stetigen Funktionen aussagen, und die Beweise waren zwar beide konstruktiv aber doch irgendwie ziemlich unterschiedlich. Gibt es vielleicht ein "gemeinsames" Konzept, eine Verallgemeinerung, aus der beide Sätze folgern? Die Antwort ist "ja", wenn man ein bißchen mehr von der Struktur der stetigen Funktionen ausnutzt, nämlich die unschuldige Beobachtung, daß mit $f, g \in C(X)$ auch ihr Produkt $f \cdot g$ zu $C(X)$ gehört, die stetigen Funktionen also nicht nur einen Vektorraum, sondern sogar eine *Algebra* bilden.

Definition 2.5 *1. Eine* Algebra *ist eine Menge, auf der Addition und Multiplikation definiert sind, und die unter diesen Operationen abgeschlossen ist.*

2. Eine Subalgebra *oder* Unteralgebra $\mathscr{X} \subset \mathscr{A}$ *einer Algebra $\mathscr{A}$ ist eine Teilmenge von $\mathscr{A}$, die unter Addition und Multiplikation abgeschlossen ist.*

3. Ein Hausdorff[18]–Raum X *ist ein topologischer Raum, bei dem je zwei verschiedene Punkte $x, x' \in X$ durch offene Mengen getrennt werden können, das heißt, es gibt offene Mengen $U, U' \in X$ mit*

$$U \cap U' = \emptyset \qquad \textit{und} \qquad x \in U, \quad x' \in U'.$$

Übung 2.1 Ist in dem kompakten metrischen Raum $X = [0, 1]$ mit der kanonischen Metrik die Menge $\left[0, \frac{1}{2}\right)$ offen, abgeschlossen oder keines von beiden? ♢

Beispiel 2.6 *Die Polynome Π bilden eine Subalgebra von $C(I)$, denn Summe und Produkt zweier Polynome sind wieder ein Polynom. Π_n hingegen bildet genau dann eine Subalgebra, wenn $n = 0$ ist – denn der Körper $\mathbb{R}$ ist ebenfalls eine Algebra.*

Übung 2.2 Zeigen Sie:

1. Ist X ein kompakter metrischer Raum, dann ist $C(X)$ eine Algebra.
2. Die trigonometrischen Polynome bilden eine Subalgebra von $C(\mathbb{T})$.

♢

Die Erweiterung des Satzes von Weierstraß, die Stone[19] 1948 in [79] gab, hat nun folgende Gestalt.

[18] Felix Hausdorff, 1868–1942, einer der "Gründerväter" der Topologie und Mengenlehre, Erfinder der "Hausdorff–Dimension" (keine Überraschung) und des Bergiffs "metrischer Raum".

[19] Marshall Harvey Stone, 1903–1989, befasste sich mit Spektraltheorie, gruppentheoretischen Aspekten der Quantenmechanik und booleschen Algebren, ist aber wohl am bekanntesten durch seine Verallgemeinerung des Approximationssatzes von Weierstraß.

Satz 2.7 (Satz von Stone–Weierstraß) *Sei X ein kompakter metrischer Raum und $\mathscr{A} \subset C(X)$ eine Unteralgebra*[20] *von $C(X)$ mit den Eigenschaften*

1. *$\mathscr{A}$ ist* punktetrennend*, das heißt, für alle $x \neq x' \in X$ gibt es eine Funktion $f \in \mathscr{A}$, so daß $f(x) \neq f(x')$.*
2. *$1 \in \mathscr{A}$.*

Dann ist $\mathscr{A}$ dicht in $C(X)$, das heißt, für alle $f \in C(X)$ und alle $\varepsilon > 0$ gibt es ein $a \in \mathscr{A}$ so daß $\|f - a\|_\infty < \varepsilon$.

Bemerkung 2.8 *Die Forderung der Punktetrennung ist nicht nur hinreichend für die Dichtheit, sondern im wesentlichen auch* notwendig*: Ist nämlich $\mathscr{A}$* nicht *punktetrennend, das heißt, gibt es zwei "magische" Punkte $x, x' \in X$, so daß $a(x) = a(x')$ für alle $a \in \mathscr{A}$, dann gilt für jede Funktion*[21] *$f \in C(X)$ mit $f(x) \neq f(x')$ daß für jedes $a \in \mathscr{A}$*

$$\|f - a\| \geq \max\{|f(x) - a(x)|, |f(x') - a(x')|\} \geq \frac{1}{2}|f(x) - f(x')| > 0,$$

was auch für das Infimum über alle a gilt, weswegen Dichtheit unmöglich ist.

Beginnen wir mit einer einfachten Beobachtung, nämlich, daß Punktetrennung zur Interpolation an zwei beliebigen Punkten äquivalent ist.

Lemma 2.9 *Ein $\mathbb{R}$–Vektorraum $V \subset C(X)$ mit $1 \in V$ ist genau dann punktetrennend, wenn es zu beliebigen Punkten $x \neq x' \in X$ und $y, y' \in \mathbb{R}$ eine Funktion $f \in V$ gibt, so daß*

$$f(x) = y, \qquad f(x') = y'. \tag{2.4}$$

Beweis: Daß (2.4) Punktetrennung impliziert, ist klar. Sei also nun V punktetrennend und seien $x \neq x'$ und y, y' vorgegeben. Nun wählen wir uns eine Funktion $g \in V$, so daß $g(x) \neq g(x')$ und setzen

$$f = y' \frac{g - g(x) \cdot 1}{g(x') - g(x)} + y \frac{g - g(x') \cdot 1}{g(x) - g(x')}$$

□

Bevor wir uns ansehen, worin der Nutzen der Algebren besteht, erinnern wir uns kurz an den Begriff des *Abschlusses*, der aus der (Funktional-) Analyis bekannt sein sollte.

Definition 2.10 *Der* Abschluß $\overline{Y}$ *einer Menge $Y \subset C(X)$ besteht aus Y und den Grenzwerten aller Cauchyfolgen in Y bezüglich der Norm von $C(X)$.*
$Y \subset C(X)$ heißt abgeschlossen*, wenn $\overline{Y} = Y$.*

[20] Als Unteralgebra des Vektorraums $C(X)$ ist $\mathscr{A}$ damit auch automatisch ein Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$, je nachdem, ob man reell- oder komplexwertige Funktionen betrachtet.

[21] Und sowas gibt es auf jedem kompakten Hausdorff–Raum X.

Bemerkung 2.11 *Mit dieser Terminologie kann man die Aussage von Satz 2.7 auch kurz und knapp als*

$$\overline{\mathscr{A}} = C(X) \tag{2.5}$$

formulieren.

Proposition 2.12 *Sei $\mathscr{A} \subset C(X)$ eine Unteralgebra von $C(X)$ und seien $f, g \in \mathscr{A}$. Dann gilt:*

1. *Die Funktion $|f|$ gehört zu $\overline{\mathscr{A}}$.*
2. *Die Funktionen $f \vee g$ und $f \wedge g$, definiert durch*

   $$(f \vee g)(x) := \max\{f(x), g(x)\}, \qquad (f \wedge g)(x) := \min\{f(x), g(x)\}, \qquad x \in X,$$

   gehören ebefalls zu $\overline{\mathscr{A}}$.
3. *Für jede endliche Menge $\mathscr{F} \subset \mathscr{A}$ gilt*

   $$\bigvee_{f \in \mathscr{F}} f \in \overline{\mathscr{A}} \qquad \textit{und} \qquad \bigwedge_{f \in \mathscr{F}} f \in \overline{\mathscr{A}}.$$

Beweis: Zum Beweis von 1. verwenden wir Satz 2.2! Und zwar sei $p_n \in \Pi$ eine Folge von Polynomen, die auf $I = [-\|f\|, \|f\|\,]$ gleichmäßig gegen die Betragsfunktion $\lambda(x) = |x|$ konvergiert. Da mit $f \in \mathscr{A}$ auch

$$p_n(f) = \sum_{j=0}^{n} p_{n,j}\, f^j \in \mathscr{A}, \qquad p_n(x) = \sum_{j=0}^{n} p_{n,j}\, x^j,$$

ist, haben wir, daß

$$\|\, p_n(f) - |f|\,\|_X \le \|p_n - \lambda\|_I \to 0 \qquad \text{für } n \to \infty,$$

also ist $|f| \in \overline{\mathscr{A}}$. Für 2. verwenden wir Teil 1. und die Beobachtung, daß

$$f \vee g = \frac{1}{2}\,(f + g + |f - g|) \qquad \text{und} \qquad f \wedge g = \frac{1}{2}\,(f + g - |f - g|),$$

sowie die Tatsache, daß mit $\mathscr{A}$ auch $\overline{\mathscr{A}}$ eine Algebra ist. Die Aussage 3. folgt schließlich unmittelbar aus 2. □

Übung 2.3 Ist $\mathscr{A} \subset C(X)$ eine Algebra, dann ist auch $\overline{\mathscr{A}}$ eine Algebra. ♢

Nach den (notwendigen) Vorarbeiten jetzt aber endlich zum Beweis des Satzes von Stone–Weierstraß.

Beweis von Satz 2.7: Wir nehmen an, es wäre $\overline{\mathscr{A}} \subset C(X)$ eine *echte* Teilmenge von $C(X)$ und damit gibt es ein $\varepsilon > 0$ und eine Funktion $f^* \in C(X)$, so daß

$$\min_{f \in \overline{\mathscr{A}}} \|f - f^*\| > \varepsilon \tag{2.6}$$

und werden mit Hilfe von Lemma 2.9 und Proposition 2.12 eine Funktion $a \in \overline{\mathscr{A}}$ konstruieren, die $\|a - f^*\| < \varepsilon$ erfüllt und uns damit einen Widerspruch liefert.
Zu $x, x' \in X$ sei $g_{x,x'} \in \mathscr{A} \subset C(X)$ eine Funktion, so daß

$$g_{x,x'}(x) = f^*(x) \qquad \text{und} \qquad g_{x,x'}(x') = f^*(x');$$

so eine Funktion existiert nach Lemma 2.9. Außerdem definieren wir die offenen[22] Mengen

$$X \supset \Omega_{x,x'} := \left\{ y \in X \ : \ g_{x,x'}(y) < f^*(y) + \frac{\varepsilon}{2} \right\} \supset \{x, x'\}.$$

Daher ist für jedes $x \in X$

$$X = \bigcup_{x' \in X} \Omega_{x,x'} = \bigcup_{x' \in J_x} \Omega_{x,x'}, \qquad J_x \subset X, \, \#J_x < \infty, \tag{2.7}$$

wobei die Indexmenge J *endlich* gewählt werden kann, weil der erste Ausdruck in (2.7) eine offene Überdeckung einer kompakten Menge ist. Damit gehören die Funktionen

$$g_x := \bigwedge_{x' \in J_x} g_{x,x'}, \qquad x \in X,$$

jeweils zu $\overline{\mathscr{A}}$ und hat die Eigenschaft, daß

$$g_x(y) \leq f^*(y) + \frac{\varepsilon}{2}, \qquad x, y \in X.$$

Und weil das mit der Kompaktheit so erfolgreich war, machen wir's gleich nochmal: Jetzt definieren wir

$$X \supset \Omega_x := \left\{ y \in X \ : \ g_x(y) > f^*(y) - \frac{\varepsilon}{2} \right\} \supset \{x\}$$

und erhalten, daß

$$X = \bigcup_{x \in X} \Omega_x = \bigcup_{x \in J} \Omega_x, \qquad J \subset X, \, \#J < \infty,$$

weswegen die Funktion

$$a := \bigvee_{x \in J} g_x \qquad \in \overline{\mathscr{A}}$$

die Eigenschaft

$$f^*(x) - \frac{\varepsilon}{2} < a(x) < f^*(x) + \frac{\varepsilon}{2}, \qquad x \in X,$$

oder eben $\|a - f^*\| < \varepsilon$ hat, im offensichtlichen Widerspruch zu (2.6). □

Übung 2.4 Leiten Sie Satz 1.7 und Satz 2.2 aus Satz 2.7 ab. ♢

[22] Als Urbilder offener Mengen unter stetigen Funktionen sind diese Mengen wieder offen! Das ist übrigens die topologische Definition der Stetigket: "Urbilder offener Mengen sind offen".

2.3 Der Satz von Bishop

So, einen haben wir noch, es gibt nämlich *noch* eine Verallgemeinerung des Satzes von Stone bzw. Stone–Weierstraß, die von Bishop[23] [9] stammt und die sich ebenfalls mit Funktionenalgebren und deren Subalgebren beschäftigt. Dazu brauchen wir noch ein klein wenig Terminologie.

Definition 2.13 *Sei X ein kompakter Hausdorff–Raum und $\mathscr{A} \subset C(X)$ eine Subalgebra mit*[24] $1 \in \mathscr{A}$.

1. *Eine Menge $Y \subset X$ heißt $\mathscr{A}$–*antisymmetrisch*, wenn jede Funktion $f \in \mathscr{A}$ mit*[25] $f|_Y \subseteq \mathbb{R}$ *auf Y konstant ist.*

2. *Für $Y \subset X$ definieren wir die (Halb–)Norm*

$$\|f\|_Y := \|f|_Y\| = \sup_{y \in Y} |f(y)|\,, \qquad f \in C(X),$$

und den Abstand zu $\mathscr{A}$ bezüglich Y als

$$d_Y\left(f, \mathscr{A}\right) = \inf_{a \in \mathscr{A}} \|f - a\|_Y = \inf_{a \in \mathscr{A}} \sup_{y \in Y} |f(y)|\,, \qquad f \in C(X).$$

Satz 2.14 *(Bishop–Stone–Weierstraß)*
Sei $\mathscr{A}$ eine abgeschlossene *Unteralgebra von $C(X)$ und $1 \in \mathscr{A}$. Wenn für $f \in C(X)$ und jede $\mathscr{A}$–antisymmetrische Teilmenge Y von X ein $a \in \mathscr{A}$ existiert, so daß $f|_Y = a|_Y$, dann ist $\mathscr{A} = C(X)$.*

Bevor wir uns an den (kurzen und eleganten) Beweis machen, der von Ransford [62] stammt, schauen wir uns erst einmal an, warum das eine Verallgemeinerung von Satz 2.7 ist. Ist nämlich $\mathscr{A}$ punktetrennend, dann gibt es für je zwei Punkte $x, x' \in X$ immer mindestens eine Funktion $a \in \mathscr{A}$, so daß $a(x) \neq a(x')$; also kann es keine aus mehr als zwei Punkten bestehende Teilmenge von X geben, auf denen alle Funktionen aus $\mathscr{A}$ konstant sind. Da aber auf *einpunktigen* Mengen alle Funktionen trivialerweise konstant sind, haben wir sofort das folgende Resultat.

Lemma 2.15 *Ist $\mathscr{A}$ eine* punktetrennende *Algebra, dann ist sind die $\mathscr{A}$–antisymmetrischen Teilmengen von X genau von der Form $\{x\}$, $x \in X$.*

Und da mit $\mathscr{A}$ auch $\overline{\mathscr{A}}$ erst recht punktetrennend ist, liefern Lemma 2.15 und Ersetzen von $\mathscr{A}$ durch $\overline{\mathscr{A}}$ in Satz 2.14 auf ziemlich unmittelbare Art und Weise Satz 2.7. Aber es wird noch besser! Anstatt den Satz von Bishop direkt zu beweisen, zeigen wir sogar ein noch etwas allgemeineres Resultat, das auf Machado [50] zurückgeht.

[23] Mehr Information als "E. Bishop" habe ich hier leider nicht.

[24] Im Englischen wird diese Eigenschaft als *unital* bezeichnet, den deutschen Terminus Technicus konnte ich aber bisher nicht auftreiben.

[25] Das heißt, diese Definition ist ursächlich für die Algebra der *komplexwertigen* stetigen Funktionen auf X gedacht; beschränken wir uns auf reellwertige Funktionen, dann ist diese Bedingung halt eben redundant.

Satz 2.16 *Zu jedem $f \in C(X)$ gibt es eine* abgeschlossene *$\mathscr{A}$–antisymmetrische Teilmenge $Y \subset X$, so daß*

$$d_Y(f, \mathscr{A}) = d_X(f, \mathscr{A}) =: d(f, \mathscr{A}).$$

Tatsächlich ist diese Aussage eine sehr interessante *qualitative* Beschreibung des Approximationsverhaltens durch eine Algebra $\mathscr{A}$: Wir wissen jetzt also, daß der Abstand von $f \in C(X)$ zu der Algebra $\mathscr{A}$, was nichts anderes als der Abstand von f zu $\overline{\mathscr{A}}$ ist, immer auf einer $\mathscr{A}$–antisymmetrischen Menge gemessen werden kann und die ist vom "individuellen" $a \in \mathscr{A}$ *unabhängig*. Und selbstverständlich folgt daraus unmittelbar Satz 2.14. Denn ist $\|f - a\|_Y = 0$ für alle $a \in \mathscr{A}$ und alle $\mathscr{A}$–antisymmetrischen Teilmengen $Y \subset X$, dann gilt das insbesondere für das "magische" Y aus Satz 2.16 und dann ist $d(f, \mathscr{A}) = 0$ – was gerade die Dichtheit liefert. Jetzt aber an die Arbeit ...

Beweis: Wir halten $f \in C(X)$ fest und definieren die Familie

$$\mathscr{F} := \{F \subset X \ : \ d_F(f, \mathscr{A}) = d(f, \mathscr{A})\}.$$

Sei nun $\mathscr{C} \subset \mathscr{F}$ eine *Kette* in $\mathscr{F}$, d.h., eine Teilfamilie, die durch Inklusion *total* geordnet wird:

$$C, C' \in \mathscr{C} \qquad \Longrightarrow \qquad C \subset C' \quad \text{oder} \quad C' \subset C.$$

Für jedes solche $\mathscr{C}$ setzen wir

$$F_{\mathscr{C}} := \bigcap_{C \in \mathscr{C}} C \quad \subset X$$

und zeigen, daß

$$\emptyset \neq F_{\mathscr{C}} \in \mathscr{F}. \tag{2.8}$$

Da nämlich für vorgegebenes $a \in \mathscr{A}$ und jedes $C \in \mathscr{C} \subset \mathscr{F}$ die Menge

$$\{x \in C \ : \ |f(x) - a(x)| \geq d(f, \mathscr{A})\}$$

1. als abgeschlossene Teilmenge einer kompakten Menge kompakt ist,
2. nichtleer ist, ja sogar C enthält: $C \in \mathscr{F}$ bedeutet, daß

$$\|f - a\|_C \geq \inf_{g \in \mathscr{A}} \|f - g\|_C = d_C(f, \mathscr{A}) = d(f, \mathscr{A})$$

ist auch

$$\{x \in F_{\mathscr{C}} \ : \ |f(x) - a(x)| \geq d(f, \mathscr{A})\} = \bigcap_{C \in \mathscr{C}} \{x \in C \ : \ |f(x) - a(x)| \geq d(f, \mathscr{A})\} \tag{2.9}$$

nichtleer und kompakt; nun ist ja a in (2.9) beliebig, wir können also dasjenige[26] a wählen, so daß $d(f, \mathscr{A}) = \|f - a\|$ und für dieses a ist dann

$$d(f, \mathscr{A}) = \max_{x \in X} |f(x) - a(x)| \geq \max_{x \in F_{\mathscr{C}}} |f(x) - a(x)| \geq d(f, \mathscr{A}),$$

[26] Hier verwenden wir, daß $\mathscr{A}$ abgeschlossen ist, ansonsten müsste man halt mit Folgen und ε argumentieren.

weswegen $F_{\mathscr{C}} \in \mathscr{F}$ und damit insbesondere $\neq \emptyset$ ist – das liefert uns dann auch (2.8). Somit hat also jede Kette $\mathscr{C}$, also jede bezüglich "$\subset$" total geordnete Menge, ein minimales Element, nämlich $F_{\mathscr{C}}$, und nach dem Zornschen Lemma muß es daher ein minimales Element $F^* \neq 0$ in $\mathscr{F}$ geben, so daß für jede echte Teilmenge $Y \subset F^*$ die Ungleichung

$$d_Y(f, \mathscr{A}) < d_{F^*}(f, \mathscr{A}) = d(f, \mathscr{A}) \tag{2.10}$$

gilt.
Dieses F^* ist nun der Kandidat für unser Y aus dem Satz; was wir also noch tun müssen, ist zu beweisen, daß F^* antisymmetrisch ist, was wir per Widerspruch tun wollen. Wäre nämlich F^* nicht antisymmetrisch, dann gibt es ein $a^* \in \mathscr{A}$, das auf F^* nicht konstant ist, so daß also

$$a_- := \min_{x \in F^*} a^*(x) < \max_{x \in F^*} a^*(x) =: a_+.$$

Indem wir, wenn nötig, a^* durch $(a^* - a_-) / (a_+ - a_-)$ ersetzen, können wir also annehmen, daß $a_- = 0$ und $a_+ = 1$ ist. Nun definieren wir die beiden kompakten Mengen

$$Y_- = \left\{ x \in F^* \,:\, 0 \leq a^*(x) \leq \frac{2}{3} \right\} \qquad \text{und} \qquad Y_+ = \left\{ x \in F^* \,:\, \frac{1}{3} \leq a^*(x) \leq 1 \right\},$$

die die Eigenschaft $Y_-, Y_+ \subset F^*$ haben und, wegen der Minimalität von F^*, gibt es Funktionen $g_-, g_+ \in \mathscr{A}$, so daß

$$\|f - g_-\|_{Y_-} < d(f, \mathscr{A}) \qquad \text{und} \qquad \|f - g_+\|_{Y_+} < d(f, \mathscr{A})$$

Diese Funktionen müssen nun geeignet kombiniert werden, und dazu definieren wir, für $n \in \mathbb{N}_0$, die Funktionen

$$\phi_n(x) := (1 - x^n)^{2^n}, \qquad x \in [0, 1],$$

die die beiden Ungleichungen

$$0 \leq x < \frac{1}{3} \qquad \Longrightarrow \qquad \phi_n(x) \geq 1 - 2^n x^n \geq 1 - \left(\frac{2}{3}\right)^n \tag{2.11}$$

$$\frac{2}{3} < x \leq 1 \qquad \Longrightarrow \qquad \phi_n(x) \leq (1 + x^n)^{-2^n} \leq \frac{1}{2^n\, x^n} \leq \left(\frac{3}{4}\right)^n \tag{2.12}$$

erfüllen. Setzen wir nun

$$g_n := \phi_n(a^*)\; g_- + (1 - \phi_n(a^*))\; g_+, \qquad n \in \mathbb{N},$$

und

$$g^* = \lim_{n \to \infty} g_n,$$

mit $g_n, g^* \in \mathscr{A}$ wegen der Abgeschlossenheit von $\mathscr{A}$, dann ist

$$\lim_{n \to \infty} \|g_n - g_-\|_{Y_-} = \lim_{n \to \infty} \|g_n - g_+\|_{Y_+} = 0$$

und da

$$\begin{aligned}\|f-g_n\|_{F^*} &\leq \|f-g_n\|_X \, le \max\left\{\|f-g_n\|_{X\setminus Y_+}, \|f-g_n\|_{X\setminus Y_-}, \|f-g_n\|_{Y_-\cap Y_+}\right\}\\ &\leq \max\Big\{\underbrace{\|f-g_-\|_{X\setminus Y_+}}_{\leq\|f-g_-\|_{Y_-}} + \underbrace{\|g_n-g_-\|_{X\setminus Y_+}}_{\to 0}, \underbrace{\|f-g_+\|_{X\setminus Y_-}}_{\leq\|f-g_+\|_{Y_+}} + \underbrace{\|g_n-g_+\|_{X\setminus Y_-}}_{\to 0},\\ &\qquad \underbrace{\|f-g_n\|_{Y_-\cap Y_+}}_{<\max\left\{\|f-g_-\|_{Y_-},\|f-g_+\|_{Y_+}\right\}}\Big\}\end{aligned}$$

ist, folgt

$$\|f-g^*\|_{F^*} = \lim_{n\to\infty}\|f-g_n\|_{F^*} \leq \max\left\{\|f-g_-\|_{Y_-}, \|f-g_+\|_{Y_+}\right\} < d(f,\mathscr{A}),$$

was im Widerspruch zu $F^* \in \mathscr{F}$ steht. □

Bemerkung 2.17 *1. Die Verwendung des Zornschen Lemmas ist weder besonders konstruktiv noch konstruktivistisch. Laut [62, Remark (i)] ist das auch gar nicht notwendig: Ist nämlich X* metrisierbar*, dann kann man eine direkte Konstruktion von F^* angeben, da X in diesem Fall eine abzählbare Basis von offenen Mengen hat.*

2. Die Approximation von Treppenfunktionen durch Polynome der Form $(1-x^n)^{2^n}$ ist auch die Idee des Beweises von Kuhn [44], siehe auch [10].

Übung 2.5 Beweisen Sie die Ungleichung (2.11) und (2.12) und verwenden Sie sie, um den Satz von Weierstraß zu beweisen. Hinweis: Approximieren Sie erst durch Treppenfunktionen und approximieren Sie dann die Treppenfunktionen durch Polynome. ♢

2.4 Müntz–Sätze

Nun zu einer Dichtheitssaussage ganz anderer Natur, nämlich zu einer anderen Verallgemeinerung der *algebraischen* Polynome, bei der wir wieder das Intervall $I = [0,1]$ betrachten. Nur approximieren wir nicht mehr mit Polynomen, sondern mit Räumen, die von den Funktionen

$$x^{\alpha_0}, x^{\alpha_1}, x^{\alpha_2}, \ldots, \qquad 0 \leq \alpha_0 < \alpha_1 < \alpha_2, \ldots \in \mathbb{R},$$

aufgespannt werden – daß die α_j aufsteigend angeordnet sind, ist natürlich, was die entsprechenden Räume angeht, völlig irrelevant. Um die Notation ein bißchen einfacher zu machen, definieren wir zu einer Folge $\alpha = (\alpha_j \,:\, j \in \mathbb{N}_0)$ die Menge $x^\alpha = \{x^{\alpha_j} \,:\, j \in \mathbb{N}_0\}$ und den "polynomialen" Approximationsraum $\Pi(p)$ als

$$\Pi(\alpha) := \operatorname{span}_{\mathbb{R}} x^\alpha := \operatorname{span}_{\mathbb{R}}\{x^{\alpha_j} \,:\, j \in \mathbb{N}_0\}.$$

Die Frage, die man sich nun stellt, ist natürlich, für welche Folgen p der Raum $\Pi(\alpha)$ *dicht* in $C(I)$ ist. Dieses Problem wurde 1912 zuerst von Bernstein [7] behandelt[27] und 1914 von Müntz[28] [57] vollständig gelöst, weswegen man Aussagen dieser Art *Müntz–Sätze* bezeichnet. Sehen wir uns das erst einmal anhand einiger Beispiele an.

Beispiel 2.18 *1. Ist $\alpha = \mathbb{N}_0$, das heißt, ist $\alpha_j = j$, $j \in \mathbb{N}_0$, dann ist $\Pi(\alpha) = \Pi$ und dafür kennen wir ja inzwischen genug Dichtheitsaussagen.*

2. *Ist $\alpha = 1$, also $\alpha_j = 1$, $j \in \mathbb{N}_0$, dann brauchen wir natürlich nicht mit Dichtheit zu rechnen; generell ist klar, daß die Folge α nicht nur endlich viele verschiedene Werte enthalten darf, weil wir dann immer nur einen endlichdimensionalen Raum zur Verfügung haben und der reicht einfach nicht aus.*

3. *Ist $\alpha = 2\mathbb{N}_0$, dann liegt $\Pi(\alpha) \subset \Pi$ dicht in $C[0,1]$, aber* nicht *in $C[-1,1]$ – das Intervall, auf dem man approximieren will, spielt also eine bedeutende Rolle. Der Grund ist einfach: $\Pi(\alpha)$ ist eine Algebra, die in $[0,1]$ punktetrennend ist, denn jedes $p \in \Pi(\alpha)$ kann man als $q(x^2)$ schreiben und Π ist ja bekanntlich punktetrennend. Sind also $x, x' \in [0,1]$, dann wählt man ein Polynom $q \in \Pi$, so daß $q(\sqrt{x}) \neq q(\sqrt{x'})$ und das Polynom $p(x) = q(x^2) \in \Pi(\alpha)$ erfüllt dann $p(x) \neq p(x')$. Damit folgt das positive Resultat aus dem Satz von Stone, Satz 2.7.*

   Daß $\Pi(\alpha)$ auf $[-1,1]$ nicht punktetrennend ist, folgt aus der Tatsache, daß alle Polynome $q \in \Pi(a)$ gerade *sind, daß also $q(x) = q(-x)$ und damit kann keine Funktion $f \in C[-1,1]$ mit $f(-\xi) \neq f(\xi)$ für ein $x \in [-1,1]$ beliebig gut durch $\Pi(\alpha)$ approximiert werden, siehe auch Bemerkung 2.8.*

4. *Natürlich müssen wir es nicht unbedingt mit Polynomen zu tun haben. Wählt man $\alpha = \beta\mathbb{N}_0$, $\beta \in \mathbb{R}_+$, dann hat $\Pi(\alpha)$ nichts mehr mit Polynomen zu tun, wenn $\beta \notin \mathbb{N}$. Trotzdem haben wir sofort wieder Dichtheit, weil wir nun jedes $q \in \Pi(\alpha)$ als $q(x) = p(x^\beta)$ schreiben können.*

5. *Man kann aber auch Nicht–Algebren von Polynomen betrachten: Wählt man α als Folge aller* Primzahlen[29]*, dann ist $\Pi(\alpha)$ natürlich* keine *Algebra mehr und Bishop–Stone helfen uns gar nichts mehr. Trotzdem werden die Müntz–Sätze uns verraten, daß diese Menge dicht in $C(X)$ liegt.*

6. *Treiben wir's noch einen Schritt weiter: Wir können α als Folge der Primzahlen und β als* beliebige *Folge mit Werten in $\left[\frac{1}{2}, 1\right]$ wählen. Liegt dann $\Pi(\alpha \cdot \beta)$ dicht in $C[0,1]$, wobei*

[27] Originalton [57]: *"Das in Frage stehende allgemeine Problem ist in einer Preisschrift von Herrn S. Bernstein, welche viele andere Probleme der Approximation durch Polynome zu einer vollen Erledigung bringt, insofern unvollständig beantwortet worden, als dort teils nur notwendige, teils nur hinreichende Kriterien [...] angegeben werden [...]*

[28] Herman (Chaim) Müntz, 1884–1956, Studium der Mathematik und Philosophie (veröffentlichte auch philosophische Bücher im Stil von Nietzsche) war als (Privat-) Lehrer, aber auch als Assistent von Einstein tätig, bevor er eine Professur in Leningrad erhielt. Zu Beginn des zweiten Weltkriegs Emigration nach Schweden, wo er bis zu seinem Tod lebte.

[29] Mit $\alpha_0 = 1$ aus offensichtlichen Gründen.

"·" das komponentenweise Produkt zweier Vektoren bezeichnet? Die Antwort ist übrigens "ja".

Übung 2.6 Zeigen Sie: Ist $P \subset \Pi$ ein *endlichdimensionaler* Raum von Polynomen, dann gibt es ein $\varepsilon > 0$ und eine stetige Funktion $f \in C(I)$, so daß

$$\inf_{p \in P} \|f - p\| > \varepsilon.$$

Hinweis: Zeigen und verwenden Sie, daß $\overline{P} \subseteq \Pi_n$ mit $n = \max\{\deg p \ : \ p \in P\}$ $\diamondsuit$

Das Ziel dieses Abschnitts ist der Beweis des folgenden Resultats, das uns Auskunft über die Dichtheit des (Vektor–) Raums $\Pi(\alpha)$ in Abhängkeit von α gibt.

Satz 2.19 *(Müntz–Satz für $C[0,1]$)*
Sei α eine strikt steigende Folge von nichtnegativen Zahlen. Dann ist $\Pi(\alpha)$ dicht in $C[0,1]$, wenn $\alpha_0 = 0$ und

$$\sum_{j=1}^{\infty} \frac{1}{\alpha_j} = \infty. \tag{2.13}$$

Der Beweis dieser Aussage, der im wesentlichen aus [17] stammt, verwendet erstaunlicherweise einen anderen Müntz–Satz, der eine Dichtheitsaussage der bezüglich der Norm

$$\|f\|_2 := \sqrt{\int_0^1 |f(x)|^2 \, dx}$$

angibt, die natürlich für jede stetige Funktion $f \in C(I)$ wohldefiniert ist, schließlich ist $\|f\|_2 \leq \|f\|_\infty$. Diese Aussage lautet wie folgt.

Satz 2.20 *(Müntz–Satz für $L_2[0,1]$)*
Sei $\alpha \subset \mathbb{R} \setminus \{0\}$ eine strikt steigende Folge von Zahlen $> -\frac{1}{2}$. Dann ist $\Pi(\alpha)$ genau dann dicht in $L_2(I)$, wenn

$$\sum_{j=1}^{\infty} \frac{1}{|\alpha_j|} = \infty. \tag{2.14}$$

Zuerst sehen wir uns an, wie wir aus der Dichtheit in $L_2(I)$ die Dichtheit in $C(I)$ ableiten können – da $C(I)$ seinerseits bezüglich der 2–Norm in $L_2(I)$ dicht ist, könnten wir Satz 2.20 ja auch als Dichtheit von $\Pi(\alpha)$ in $C(I)$ bezüglich der 2–Norm formulieren. Schematisch sieht das wie folgt aus:

$$\begin{array}{ccccc}
\Pi(\alpha) & \subset & C(I) & \subset & L_2(I) \\
\overline{\Pi(\alpha)}^{\|\cdot\|_2} & = & \overline{C(I)}^{\|\cdot\|_2} & = & \overline{L_2(I)}^{\|\cdot\|_2} \\
& \text{Satz 2.20} & & & \\
\overline{\Pi(\alpha)}^{\|\cdot\|_\infty} & = & \overline{C(I)}^{\|\cdot\|_\infty} & & \\
& \text{Satz 2.19} & & &
\end{array}$$

Beweis von Satz 2.19: Wir betrachten, für $N \in \mathbb{N}_0$, einen Koeffizientenvektor $a = [a_0, \ldots, a_N] \in \mathbb{R}^{N+1}$, sowie für $x \in [0,1]$ und $n \in \mathbb{N}_0$, den Ausdruck

$$\begin{aligned}\left| x^n - \sum_{j=0}^{N} a_j x^{\alpha_j} \right| &= \frac{1}{n} \left| \int_0^x t^{n-1} - \sum_{j=1}^{N} a_j \frac{n}{\alpha_j} t^{\alpha_j - 1} \, dt \right| \\ &\leq \frac{1}{n} \int_0^x \left| t^{n-1} - \sum_{j=1}^{N} a_j \frac{n}{\alpha_j} t^{\alpha_j - 1} \right| \, dt \leq \frac{1}{n} \int_0^1 \left| t^{n-1} - \sum_{j=0}^{N} a_j \frac{n}{\alpha_j} t^{\alpha_j - 1} \right| \, dt \\ &\leq \frac{1}{n} \left(\int_0^1 \left| t^{n-1} - \sum_{j=0}^{N} \frac{a_j n}{\alpha_j} t^{\alpha_j - 1} \right|^2 dt \right)^{1/2} =: \frac{1}{n} \left\| t^{n-1} - \sum_{j=0}^{N} \widetilde{a}_j \, t^{\alpha_j - 1} \right\|_2 ,\end{aligned}$$

wobei wir beim letzten Schritt die *Höldersche Ungleichung*

$$\|fg\|_1 \leq \|f\|_p \, \|g\|_q , \qquad \frac{1}{p} + \frac{1}{q} = 1,$$

mit $g = 1$ verwendet haben. Da diese Rechnung also besagt, daß

$$\left\| x^n - \sum_{j=0}^{N} a_j x^{\alpha_j} \right\| \leq \frac{1}{n} \left\| t^{n-1} - \sum_{j=0}^{N} \widetilde{a}_j \, t^{\alpha_j - 1} \right\|_2 , \qquad \widetilde{a}_j = \frac{a_j n}{\alpha_j}, \quad j = 0, \ldots, N, \tag{2.15}$$

wäre die Dichtheit von $\Pi(\alpha - 1)$ in $L_2(I)$ hinreichend für die Dichtheit von $\Pi(\alpha)$ in $C(I)$ – siehe auch Übung 2.7. Nun unterscheiden wir drei Fälle[30]:

1. $\lim_{j\to\infty} \alpha_j = \infty$. Es gibt einen Index $N \in \mathbb{N}$, so daß $\alpha_j > 2$, $j \geq N$, und somit ist

$$\sum_{j=1}^{\infty} \frac{1}{|\alpha_j - 1|} = \sum_{j=1}^{N-1} \frac{1}{|\alpha_j - 1|} + \sum_{j=N}^{\infty} \frac{1}{|\alpha_j - 1|} \geq \frac{1}{2} \sum_{j=N}^{\infty} \frac{1}{\alpha_j} = \infty,$$

   wir können also Satz 2.20 anwenden und erhalten Dichtheit.

2. $-\frac{1}{2} < \lim_{j\to\infty} \alpha_j^{-1}$. Hier ist die Folge $\frac{1}{|\alpha_j - 1|}$ noch nicht einmal eine Nullfolge, die Reihe divergiert also auf alle Fälle und wir können den Rest dem Satz 2.20 überlassen.

3. $-1 < \lim_{j\to\infty} \alpha_j^{-1} \leq -\frac{1}{2}$. Es gibt eine Konstante $0 < c < 1$, so daß $\lim_{j\to\infty} c\alpha_j^{-1} > -\frac{1}{2}$ ist und nach Punkt 2. ist $\Pi(c\alpha)$ dicht in $C(I)$. Nun ist aber, für beliebiges $p \in \Pi(c\alpha)$ und $f \in q$

$$\|f - p\| = \max_{x\in I} |f(x) - p(x)| = \max_{x\in I} |f(x^c) - p(x^c)| = \left\| \widetilde{f} - \widetilde{q} \right\| ,$$

   wobei $q \in \Pi(\alpha)$ – aber diese Reparametrisierung läßt den Abstand zwischen der Funktion und den "Polynomen" unverändert.

[30] Wobei wir eine eventuelles $\alpha_j = 1$ aus der Folge α streichen; für die Dichtheit ist das irrelevant und wir wollen ja nicht, daß die Reihen divergieren, nur weil wir dummerweise an irgendeiner Stelle durch 0 geteilt haben ...

□

Übung 2.7 Zeigen Sie: Π ist dicht in $L_2(I)$. ♢

Was bleibt, ist der Beweis von Satz 2.20, der aber auch nicht so schlimm ist. Außerdem lernen wir dabei auch gleich ein paar der grundlegenden Ideen, die man bei der Bestapproximation *im quadratischen Mittel* so verwendet, bei dem die Norm ja durch das *innere Produkt*

$$\langle f, g\rangle := \int_0^1 f(t)g(t)\,dt \qquad \Longrightarrow \qquad \|f\|_2 = \sqrt{\langle f, f\rangle}$$

bestimmt ist.

Definition 2.21 *Für Funktionen $f_1, \ldots, f_n \in L_2(I)$ definieren wir die* Gramsche[31] Matrix

$$G(f_1, \ldots, f_n) = [\langle f_j, f_k\rangle \;:\; j, k = 1, \ldots, n] = \begin{bmatrix} \langle f_1, f_1\rangle & \ldots & \langle f_1, f_n\rangle \\ \vdots & \ddots & \vdots \\ \langle f_n, f_1\rangle & \ldots & \langle f_n, f_n\rangle \end{bmatrix} \tag{2.16}$$

und die Gramsche Determinante

$$g(f_1, \ldots, f_n) = \det G(f_1, \ldots, f_n) = \begin{vmatrix} \langle f_1, f_1\rangle & \ldots & \langle f_1, f_n\rangle \\ \vdots & \ddots & \vdots \\ \langle f_n, f_1\rangle & \ldots & \langle f_n, f_n\rangle \end{vmatrix}.$$

Lemma 2.22 *Seien $\phi_1, \ldots, \phi_n \in L_2(I)$ linear unabhängig und $\Phi := \operatorname{span}_{\mathbb{R}}\{\phi_1, \ldots, \phi_n\}$.*

1. *Für $f \in L_2(I)$ und $\phi^* \in \Phi$ gilt*

$$\|f - \phi^*\|_2 = \min_{\phi\in\Phi} \|f - \phi\|_2 \qquad \Longleftrightarrow \qquad \langle f - \phi^*, \Phi\rangle = 0. \tag{2.17}$$

2. *Für $f \in L_2(I)$ ist*

$$d_2(f, \Phi) = \min_{\phi\in\Phi} \|f - \phi\|_2 = \sqrt{\frac{g(f, \phi_1, \ldots, \phi_n)}{g(\phi_1, \ldots, \phi_n)}}. \tag{2.18}$$

Beweis: Beginnen wir mit (2.17). Sei $f \in L_2(I)$ und $\psi_1, \ldots, \psi_n$ eine Orthonormalbasis von Φ. Dann ist

$$\left\| f - \sum_{j=1}^n \langle f, \psi_j\rangle\, \psi_j \right\|_2^2 = \left\langle f - \sum_{j=1}^n \langle f, \psi_j\rangle\, \psi_j, f - \sum_{j=1}^n \langle f, \psi_j\rangle\, \psi_j \right\rangle$$

[31] Jorgen Pedersen Gram, 1850–1916, dänischer Mathematiker, erarbeitete für die Versicherungsgesellschaft "Hafnia" ein mathematisches Modell zur Forstverwaltung und gelangte über Fragen der Stochastik und Numerik später auch zur Zahlentheorie.

$$
\begin{aligned}
&= \langle f,f\rangle - 2\sum_{j=1}^{n}\langle f,\psi_j\rangle\langle f,\psi_j\rangle + \sum_{j,k=1}^{n}\langle f,\psi_j\rangle\langle f,\psi_k\rangle \underbrace{\langle\psi_j,\psi_k\rangle}_{\delta_{jk}}\\
&= \|f\|_2^2 - \sum_{j=1}^{n}\langle f,\psi_j\rangle^2 \le \|f\|_2^2 - \sum_{j=1}^{n}\langle f,\psi_j\rangle^2 + \sum_{j=1}^{n}\left(a_j - \langle f,\psi_j\rangle\right)^2\\
&= \|f\|_2^2 - 2\sum_{j=1}^{n} a_j\,\langle f,\psi_j\rangle + \sum_{j=1}^{n} a_j^2\langle\psi_j,\psi_j\rangle = \left\|f - \sum_{j=1}^{n} a_j\,\psi_j\right\|_2^2,
\end{aligned}
$$

weswegen die Bestapproximation gerade der Fall $a_j = \langle f,\psi_j\rangle$, $j = 1,\ldots,n$ ist. Und das ist für $k = 1,\ldots,n$ äquivalent zu

$$
\left\langle f - \sum_{j=1}^{n}\langle f,\psi_j\rangle\,\psi_j, \psi_k\right\rangle = \langle f,\psi_k\rangle - \sum_{j=1}^{n}\langle f,\psi_j\rangle\langle\psi_j,\psi_k\rangle = \langle f,\psi_k\rangle - \langle f,\psi_k\rangle = 0.
$$

Für (2.18) sei $\phi^* = a_1\phi_1 + \cdots + a_n\phi_n$ die[32] Bestapproximation von f in Φ; dann ist nach (2.17)

$$
d^2 = d_2^2(f,\Phi) = \|f-\phi^*\|_2^2 = \langle f-\phi^*, f-\phi^*\rangle = \langle f, f-\phi^*\rangle - \underbrace{\langle\phi^*, f-\phi^*\rangle}_{=0},
$$

also

$$
d^2 = \langle f,f\rangle - \sum_{j=1}^{n} a_j\,\langle f,\phi_j\rangle. \tag{2.19}
$$

Schreiben wir außerdem (2.17) bezüglich der Basis $\phi_1,\ldots,\phi_n$, dann heißt dies, daß für $j = 1,\ldots,n$

$$
0 = \langle f-\phi^*,\phi_j\rangle = \langle f,\phi_j\rangle - \sum_{k=1}^{n} a_k\,\langle\phi_j,\phi_k\rangle \tag{2.20}
$$

In Matrixform liefern (2.19) und (2.20), die sogenannte *Normalengleichungen*

$$
\begin{bmatrix}
\langle f,f\rangle & \langle f,\phi_1\rangle & \ldots & \langle f,\phi_n\rangle\\
\langle\phi_1,f\rangle & \langle\phi_1,\phi_1\rangle & \ldots & \langle\phi_1,\phi_n\rangle\\
\vdots & \vdots & \ddots & \vdots\\
\langle\phi_n,f\rangle & \langle\phi_n,\phi_1\rangle & \ldots & \langle\phi_n,\phi_n\rangle
\end{bmatrix}
\begin{bmatrix} 1\\ -a_1\\ \vdots\\ -a_n\end{bmatrix} =
\begin{bmatrix} d^2\\ 0\\ \vdots\\ 0\end{bmatrix},
$$

die wir auch als

$$
G(f,\phi_1,\ldots,\phi_n)\begin{bmatrix} 1\\ -a_1\\ \vdots\\ -a_n\end{bmatrix} = \begin{bmatrix} d_2\\ 0\\ \vdots\\ 0\end{bmatrix} \tag{2.21}
$$

[32] Nach dem, was wir gerade bewiesen haben, ist es wirklich **die** Bestapproximation – sie ist eindeutig.

schreiben können. Lösen wir nun das lineare Gleichungssystem (2.21) mit Hilfe der Cramerschen[33] Regel, dann erhalten wir für die erste Komponente der Lösung, daß

$$1 = \frac{\begin{vmatrix} d^2 & \langle f,\phi_1\rangle & \dots & \langle f,\phi_n\rangle \\ 0 & \langle \phi_1,\phi_1\rangle & \dots & \langle \phi_1,\phi_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \langle \phi_1,\phi_1\rangle & \dots & \langle \phi_1,\phi_n\rangle \end{vmatrix}}{\begin{vmatrix} \langle f,f\rangle & \langle f,\phi_1\rangle & \dots & \langle f,\phi_n\rangle \\ \langle \phi_1,f\rangle & \langle \phi_1,\phi_1\rangle & \dots & \langle \phi_1,\phi_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle \phi_1,f\rangle & \langle \phi_1,\phi_1\rangle & \dots & \langle \phi_1,\phi_n\rangle \end{vmatrix}} = d^2\,\frac{g\left(\phi_1,\dots,\phi_n\right)}{g\left(f,\phi_1,\dots,\phi_n\right)},$$

und das ist (2.18). □

Bemerkung 2.23 *Die Normalengleichungen (2.21) ermöglichen es uns, die Bestapproximation bezüglich der L_2–Norm durch "einfaches" Lösen eines linearen Gleichungssystems zu bestimmen und diese Lösung liefert uns als netten Nebeneffekt auch noch die Qualität der Approximation mit – schließlich müssen wir erst mal d^2 kennen, um die rechte Seite zu komplettieren, aber dafür haben wir ja die Formel (2.18).*
Für den Numeriker ist so eine Gram–Matrix übrigens nicht zu wild, denn sie ist symmetrisch und positiv definit und dafür gibt es vergleichsweise gute und stabile Lösungsverfahren.

Um den Quotienten der Gram–Determinanten aus (2.18) in den Griff zu bekommen, ja sogar *explizit* angeben zu können, verwendet man eine Determinantenformel, die auf Cauchy zurückgeht.

Lemma 2.24 *Seien $a, b \subset \mathbb{R}$ Zahlenfolgen und sei, für $n \in \mathbb{N}$,*

$$D_n = \left[\frac{1}{a_j + b_k} \;:\; j,k = 1,\dots,n\right] = \begin{bmatrix} \frac{1}{a_1+b_1} & \dots & \frac{1}{a_1+b_n} \\ \vdots & \ddots & \vdots \\ \frac{1}{a_n+b_1} & \dots & \frac{1}{a_n+b_n} \end{bmatrix}.$$

Dann ist

$$d_n := \det D_n = \frac{\prod\limits_{1\le j<k\le n} (a_j - a_k)\,(b_j - b_k)}{\prod\limits_{j,k=1}^{n} (a_j + b_k)}. \tag{2.22}$$

[33] Gabriel Cramer, 1704–1752, schweizerisch–französischer Mathematiker, hatte regen Kontakt zur Bernoulli–Familie. Amüsanterweise verwendete er die nach ihm benannte "Lösungsmethode" für lineare Gleichungssysteme (die er wohl nicht als erster angab) zur Untersuchung von algebraischen Kurven.

Beweis: Wir betrachten $a_1, \ldots, a_n$ und $b_1, \ldots, b_n$ als Variable, dann ist $d_n = \frac{p_n}{q_n}$ eine rationale Funktion in diesen $2n$ Variablen mit Nenner[34]

$$q_n = \prod_{j,k=1}^{n} (a_j + b_k)\,,$$

denn $q_n D_n$ ist eine Matrix, deren Einträge Polynome in den $2n$ Variablen sind. Da jeder Eintrag von D_n Grad -1 hat (siehe Übung 2.8) hat nach der Leibnitz–Regel d_n den Grad $\leq -n$, also hat das Zählerpolynom höchstens den Grad $n^2 - n$, da q_n ja Grad n^2 hat.
Ist nun $a_j = a_k$ für $j \neq k$, dann haben wir zwei identische Zeilen in D_n und damit ist $d_n = 0$, dasselbe gilt auch, wenn $b_j = b_k$ für $j \neq k$. Damit hat also der Zähler p_n von d_n die Form

$$p_n = c_n \left(\prod_{1 \leq j < k \leq n} (a_j - a_k) \right) \left(\prod_{1 \leq j < k \leq n} (b_j - b_k) \right).$$

Nun enthält aber jedes der beiden Produkte $\frac{n(n-1)}{2}$ Faktoren und damit sind die beiden Produkte zusammen schon ein Polynom vom Grad $n^2 - n$, weswegen c_n eine Konstante sein muß.
Bleibt zu zeigen, daß $c_n = 1$ ist. Zu diesem Zwecke multiplizieren wir die letzte Zeile von D_n mit a_n und erhalten mit $a_n \to \infty$, daß

$$\lim_{a_n \to \infty} a_n\, d_n = \lim_{a_n \to \infty} \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} \\ \vdots & \ddots & \vdots \\ \frac{1}{a_{n-1}+b_1} & \cdots & \frac{1}{a_{n-1}+b_n} \\ \frac{a_n}{a_n+b_1} & \cdots & \frac{a_n}{a_n+b_n} \end{vmatrix} = \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_n} \\ \vdots & \ddots & \vdots \\ \frac{1}{a_{n-1}+b_1} & \cdots & \frac{1}{a_{n-1}+b_n} \\ 1 & \cdots & 1 \end{vmatrix}$$

Und weil's so schön war lassen wir auch noch b_n wachsen und erhalten, daß

$$\lim_{b_n \to \infty} \lim_{a_n \to \infty} a_n\, d_n = \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_{n-1}} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \frac{1}{a_{n-1}+b_1} & \cdots & \frac{1}{a_{n-1}+b_{n-1}} & 0 \\ 1 & \cdots & 1 & 1 \end{vmatrix} = d_{n-1}. \tag{2.23}$$

Nach dem, was wir schon über d_n und d_{n-1} wissen, ist dann aber[35]

$$\frac{a_n\, d_n}{d_{n-1}} = a_n \frac{c_n}{c_{n-1}} \frac{\prod_{j=1}^{n-1} (a_j - a_n) \prod_{j=1}^{n-1} (b_j - b_n)}{\prod_{j=1}^{n} (a_n + b_j) \prod_{j=1}^{n-1} (a_j + b_n)}$$

[34] Nicht notwendigerweise gekürzt!

[35] Die Obergrenze "$n-1$" im zweiten Produkt des Nenners mag zuerst etwas verwirren, kommt aber daher, daß der Faktor $a_n + b_n$ nur *einmal* auftaucht und schon im ersten Produkt erledigt wird.

und somit

$$\frac{c_n}{c_{n-1}} = \lim_{b_n\to\infty}\lim_{a_n\to\infty} \underbrace{\frac{a_n\,d_n}{d_{n-1}}}_{\to 1}\underbrace{\frac{a_n+b_n}{a_n}}_{\to 1}\prod_{j=1}^{n-1}\underbrace{\frac{a_n+b_j}{a_j-a_n}}_{\to -1}\prod_{j=1}^{n-1}\underbrace{\frac{a_j+b_n}{b_j-b_n}}_{\to -1} = (-1)^{2n-2} = 1,$$

und da $c_1 = 1$ ist, folgt (2.22). □

Übung 2.8 Der *Grad* einer rationalen Funktion $f = \frac{p}{q}$ sei definiert als

$$\deg f = \deg p - \deg q.$$

Zeigen Sie, daß für rationale Funktionen f, g die von den Polynomen bekannten Gradaussagen

$$\deg(f+g) \le \max\{\deg f, \deg g\} \qquad \text{und} \qquad \deg fg = \deg f + \deg g$$

gelten. ♢

Jetzt aber zurück zu unserem Müntz–Satz.

Lemma 2.25 *Es sei $\alpha \subset \mathbb{R}\setminus\{0\}$ eine monoton steigende Folge von Zahlen $> -\frac{1}{2}$. Dann ist für $N \in \mathbb{N}_0$ und jedes $q \in \mathbb{R}$, $q > -\frac{1}{2}$*

$$d_N^2 := \min_{a_0,\ldots,a_N\in\mathbb{R}} \left\| x^q - \sum_{j=1}^{N} a_j x^{\alpha_j} \right\|_2^2 = \frac{1}{2q+1}\prod_{j=1}^{N}\left(\frac{\alpha_j - q}{\alpha_j + q + 1}\right)^2. \tag{2.24}$$

Bemerkung 2.26 *Gleichung (2.24) zeigt, daß $q = -\frac{1}{2}$ eine echte Grenze bei der Approximation darstellt und daß Approximation immer schwieriger wird, wenn $q \to -\frac{1}{2}$ geht. Das ist aber nicht so verwunderlich, denn die Funktion $f(x) = x^{-1/2}$ ist ja eben gerade nicht mehr quadratintegrierbar, gehört also nicht mehr zu $L_2(I)$!*

Beweis von Lemma 2.25: Wir verwenden natürlich jetzt Lemma 2.22, genauer (2.18), denn das liefert uns, daß

$$d_N^2 = \frac{g\left(x^q, x^{\alpha_0}, \ldots, x^{\alpha_N}\right)}{g\left(x^{\alpha_0}, \ldots, x^{\alpha_N}\right)}. \tag{2.25}$$

Nun sind aber Gram–Matrizen von Polynomen eher einfacher Natur: Da

$$\langle x^p, x^q\rangle = \int_0^1 t^{p+q}\,dt = \frac{1}{p+q+1}\; t^{p+q+1}\Big|_{t=0}^{1} = \frac{1}{p+q+1},$$

ist

$$G\left(x^q, x^{\alpha_0}, \ldots, x^{\alpha_N}\right) = \begin{bmatrix} \frac{1}{2q+1} & \frac{1}{q+\alpha_0+1} & \cdots & \frac{1}{q+\alpha_0+1} \\ \frac{1}{q+\alpha_0+1} & \frac{1}{2\alpha_0+1} & \cdots & \frac{1}{\alpha_0+\alpha_N+1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{q+\alpha_N+1} & \frac{1}{\alpha_0+\alpha_N+1} & \cdots & \frac{1}{2\alpha_N+1} \end{bmatrix}$$

und entsprechendes gilt für $G\left(x^{\alpha_0},\ldots,x^{\alpha_N}\right)$, so daß wir nun endlich Lemma 2.24 anwenden können, und zwar mit $a = b = \left[q+\frac{1}{2},\alpha_0+\frac{1}{2},\ldots,\alpha_N+\frac{1}{2}\right]$ im Zähler bzw. $a = b = \left[\alpha_0+\frac{1}{2},\ldots,\alpha_N+\frac{1}{2}\right]$ und so ist

$$\begin{aligned} d_N^2 &= \frac{g\left(x^q,x^{\alpha_0},\ldots,x^{\alpha_N}\right)}{g\left(x^{\alpha_0},\ldots,x^{\alpha_N}\right)} \\ &= \frac{\displaystyle\prod_{j=0}^{N}\left(\alpha_j+\frac{1}{2}-q-\frac{1}{2}\right)\left(\alpha_j+\frac{1}{2}-q-\frac{1}{2}\right)}{\left(q+\frac{1}{2}+q+\frac{1}{2}\right)\displaystyle\prod_{j=0}^{N}\left(\alpha_j+\frac{1}{2}+q+\frac{1}{2}\right)\left(\alpha_j+\frac{1}{2}+q+\frac{1}{2}\right)} \\ &= \frac{1}{2q+1}\frac{\displaystyle\prod_{j=0}^{N}\left(\alpha_j-q\right)^2}{\displaystyle\prod_{j=0}^{N}\left(\alpha_j+q+1\right)^2}, \end{aligned}$$

was (2.24) liefert. □

Das war's dann auch "schon" mit den Zutaten, es wird Zeit, den Beweis zu vervollständigen.
Beweis von Satz 2.20: Wegen der Dichtheit[36] der Polynome in $C(I)$ (siehe Satz 2.2) und damit auch in $L_2(I)$, genügt es, zu zeigen, daß

$$0 = \lim_{N\to\infty}\min_{a_0,\ldots,a_N\in\mathbb{R}}\left\|x^q-\sum_{j=0}^{N}a_jx^{\alpha_j}\right\|_2, \qquad q\in\mathbb{N}_0,$$

oder, nach Lemma 2.25, daß

$$\lim_{N\to\infty}\prod_{j=1}^{N}\left(\frac{\alpha_j-q}{\alpha_j+q+1}\right)^2 = 0, \qquad q\in\mathbb{N}_0. \tag{2.26}$$

Wir beginnen damit, daß wir zeigen, daß die Bedingung (2.14) *hinreichend* für die Dichtheit ist. Dazu unterscheiden wir zwei Fälle:

1. $p := \lim_{j\to\infty}\alpha_j < \infty$, d.h. (2.14) ist trivialerweise erfüllt. Da in diesem Fall

$$\left|\frac{\alpha_j-q}{\alpha_j+q+1}\right| = \Big|1-\underbrace{\frac{2q+1}{\alpha_j+1+q}}_{>0}\Big| < 1-\frac{q+\frac{1}{2}}{p+q+1},$$

für hinreichend großes j gilt, folgt (2.26) unmittelbar.

[36] Jeweils bezüglich der entsprechenden Norm.

2. $\lim_{j\to\infty}\alpha_j = \infty$. In diesem Fall schreiben wir das Produkt in (2.26) als

$$\prod_{j=0}^{N}\frac{\left(\frac{\alpha_j-q}{\alpha_j}\right)^2}{\left(\frac{\alpha_j+q+1}{\alpha_j}\right)^2}=\prod_{j=0}^{N}\frac{\left(1-\frac{q}{\alpha_j}\right)^2}{\left(1+\frac{q+1}{\alpha_j}\right)^2}.$$

Für hinreichend großes j sind alle Terme im Zähler kleiner als 1 und somit ist das Zählerprodukt

$$\prod_{j=0}^{N}\left(1-\frac{q}{\alpha_j}\right)$$

zumindest unabhängig von N beschränkt. Wählen wir M so, daß $\alpha_j > 0$, $j > M$, dann ist

$$\prod_{j=M}^{N}\left(1+\frac{q+1}{\alpha_j}\right)\geq 1+(q+1)\sum_{j=M}^{N}\frac{1}{\alpha_j}$$

und nachdem diese Summe und damit auch das Nennerprodukt divergiert, folgt ebenfalls (2.26).

Als "Bonus" sehen wir uns schließlich noch an, warum (2.14) auch *notwendig* für die Dichtheit ist. Wäre nämlich (2.14) verletzt, dann müsste auf jeden Fall $\lim_{j\to\infty}\alpha_j = \infty$ sein, aber eben[37]

$$\sum_{j=0}^{\infty}\frac{1}{\alpha_j}<\infty \qquad \Longrightarrow \qquad \sum_{j=0}^{\infty}\frac{1}{\alpha_j^p}<\infty, \quad p\geq 1.$$

Nun konvergieren aber die Zähler- und Nennerprodukte für jedes $q \notin \alpha$ aus der vorherigen Überlegung gegen einen strikt positiven Wert, und das heißt aber, daß

$$d_2\left(x^q, \Pi(\alpha)\right) > 0.$$

Und mindestens ein solches $q \in \mathbb{N}$ muß es nun schon geben, denn sonst wäre die Reihe $\sum \frac{1}{\alpha_j}$ ja schließlich divergent. □

Lemma 2.27 *Für $a_j \in (0,1)$, $j \in \mathbb{N}$, gilt*

$$\prod_{j=1}^{\infty}(1-a_j) > 0 \qquad \Leftrightarrow \qquad \sum_{j=1}^{\infty} a_j = \infty.$$

[37] Die Betragsstriche aus (2.14) können wir weglassen, denn die Konvergenz entscheidet sich ja jetzt "im Positiven", schließlich sind fast alle $\alpha_j > 0$.

Denkrunen lerne,
soll der Degen keiner
deinen Verstand bestehn!

Die Edda, "Die Runenlehren"

3 Approximation in linearen Räumen

In diesem Kapitel befassen wir uns jetzt mal mit ein bißchen Approximations"theorie", genauer, mit der Frage nach Existenz, Eindeutigkeit, Charakterisierung und Bestimmung der Bestapproximation in einem endlichdimensionalen (Funktionen-) Raum. Die Situation ist hierbei immer die folgende: Wir approximieren Funktionen in einem normierten linearen Raum, später natürlich vor allem die stetigen Funktionen[38] $C(X)$ mit der Norm $\|\cdot\|_\infty$, durch Elemente eines endlichdimensionalen linearen Teilraums. Typische Beispiele für lineare Approximation in $C(X)$ sind

- algebraische Polynome,
- trigonometrische Polynome,

typische Beispiele für *nichtlineare* Approximation (auch so was gibt es) sind

- *rationale Funktionenräume* der Form

$$\Pi_{n,m} := \left\{ f = \frac{p}{q} \ : \ p \in \Pi_n,\, q \in \Pi_m \right\}, \qquad n, m \in \mathbb{N}_0,$$

- *n–Term–Approximation*, wobei man für $N \in \mathbb{N}_0$ oder $N = \infty$ mit dem Raum

$$F_n(\Phi) := \left\{ \sum_{j=1}^{N} a_j\, \phi_j \ : \ \#\{j \ : \ a_j \neq 0\} \leq n \right\}, \qquad \Phi = \{\phi_1, \ldots, \phi_N\},$$

  approximiert. Solche Räume tauchen bei der Waveletapproximation, beispielsweise beim Entrauschen von Signalen, der Kantendetektion oder der Bildkompression auf.

Übung 3.1 Zeigen Sie, daß $\Pi_{n,m}$ und $F_n(\Phi)$ keine Vektorräume sind. ♢

[38] Hierbei darf X zuerst mal ein kompakter Hausdorffraum sein – allerdings immer mit unseren zwei "Musterbeispielen" $X = I$ und $X = \mathbb{T}$ im Hinterkopf! Und wir werden später sehen, daß dies oftmals die einzige Möglichkeit ist.

3.1 Approximation durch lineare Räume

Hier können wir mal richtig allgemein sein: Es sei F ein komplexer Vektorraum[39] mit einer Norm $\|\cdot\|$ und es seien $\phi_1, \ldots, \phi_n \in F$ linear unabhängig. Wir bezeichnen wieder mit Φ den Teilraum

$$\Phi = \operatorname{span}_{\mathbb{C}} \{\phi_j \ : \ j = 1, \ldots, n\}.$$

Definition 3.1 *Zu $f \in F$ bezeichnen wir mit*

$$d(f, \Phi) = \inf_{\phi \in \Phi} \|f - \phi\|$$

den Abstand von f zu Φ *und mit*

$$P_\Phi(f) := \{\phi^* \in \Phi \ : \ \|f - \phi^*\| = d(f, \Phi)\}$$

die Menge aller Bestapproximationen zu f in Φ. Die Menge P_Φ bezeichnet man auch als metrische Projektion von f auf Φ*, die man auch als* mengenwertige *Abbildung $P_\Phi \ : \ F \to \Phi$ auffassen kann*[40].

Proposition 3.2 *Für jedes $f \in F$ ist P_Φ eine* konvexe *Menge, das heißt,*

$$\phi, \phi' \in P_\Phi(f) \qquad \Longleftrightarrow \qquad \alpha\phi + (1-\alpha)\,\phi' \in P_\Phi(f), \quad \alpha \in [0,1].$$

Insbesondere ist also entweder[41] $\#P_\Phi(f) = 1$ *oder* $\#P_\Phi(f) = \infty$.

Beweis: Seien $\phi, \phi' \in P_\Phi(f)$ und $\alpha \in [0,1]$. Dann ist

$$\begin{aligned} d(f,\Phi) &\leq \|f - (\alpha\phi + (1-\alpha)\phi')\| = \|(\alpha + (1-\alpha))\, f - (\alpha\phi + (1-\alpha)\phi')\| \\ &\leq \alpha \underbrace{\|f - \phi\|}_{=d(f,\Phi)} + (1-\alpha) \underbrace{\|f - \phi'\|}_{=d(f,\Phi)} = d(f,\Phi). \end{aligned}$$

Der Rest ist trivial. □

Der letzte Satz im obigen Beweis war *nicht* richtig. Was wir nämlich bisher unterschlagen haben ist die *Existenz* einer Bestapproximation, oder, anders gesagt, daß $P_\Phi(f) \neq \emptyset$ für alle $f \in F$.

Satz 3.3 *Zu jedem $f \in F$ gibt es mindestens eine Bestapproximation $\phi^* \in \Phi$, das heißt, $P_\Phi(f) \neq \emptyset$, $f \in F$.*

[39] Also ein Vektorraum über $\mathbb{C}$.

[40] Und "Projektion" bedeutet nun gerade nichts anderes als, daß die Bestapproximation an eine Bestapproximation die Bestapproximation selbst ist, oder eben, daß für $f \in \Phi$ immer $P_\Phi f = f = \{f\}$ gilt.

[41] Daß beides gleichzeitig unmöglich ist, ist ja wohl klar.

Beweis: Was wir beweisen werden, ist eine *Kompaktheitsaussage*, die im wesentlichen auf der Endlichdimensionalität von Φ beruht.
Es gibt eine Folge $\psi_k \in \Phi$, $k \in \mathbb{N}_0$, so daß

$$d(f,\Phi) = \inf_{\phi\in\Phi} \|f-\phi\| = \lim_{k\to\infty} \|f-\psi_k\|,$$

und für diese Funktionen gilt, daß für jedes $k \in \mathbb{N}_0$

$$\|\psi_k\| \le \|f\| + \underbrace{\|f-\psi_k\|}_{\to d(f,\Phi)} \le \|f\| + \underbrace{\max_{k\in\mathbb{N}_0} \|f-\psi_k\|}_{<\infty} =: M < \infty.$$

Da ψ_k eine Folge in der *endlichdimensionalen und damit kompakten* Menge aller Vektoren mit Norm $\le M$ ist, können wir eine Teilfolge ψ_{k_j} extrahieren, die gegen einen Vektor

$$\phi^* = \lim_{j\to\infty} \psi_{k_j}$$

konvergiert und da die Norm stetig ist (siehe Übung 3.2), haben wir, daß

$$\|f-\phi^*\| = \left\| f - \lim_{j\to\infty} \psi_{k_j} \right\| = \lim_{j\to\infty} \left\| f-\psi_{k_j} \right\| = d(f,\Phi)$$

ist.
Bleibt also noch die Kompaktheit. Zu diesem Zweck seien a_{jk} diejenigen Koeffzienten, für die

$$\psi_j = \sum_{k=1}^{n} a_{jk}\,\phi_k, \qquad j \in \mathbb{N}_0,$$

gilt und aus denen wir die Koeffizienten der konvergenten Teilfolge extrahieren wollen. Wir setzen

$$M' := \sup_{j\in\mathbb{N}_0} \max_{k=1,\dots,n} |a_{jk}|$$

und unterscheiden zwei Fälle:

1. $M' < \infty$: Die individuellen Folgen $a_{\cdot,k}$, $k = 1,\dots,n$, sind beschränkt und wir können zuerst eine Teilfolge extrahieren, so daß $a_{j_\ell,1}$ konvergiert, aus dieser Teilfolge eine, so daß auch (noch) $a_{j_\ell,2}$ konvergiert und so weiter und damit erhalten wir nach endlich vielen Schritten eine Teilfolge $j_\ell \subset \mathbb{N}_0$, so daß

$$\phi^* = \lim_{\ell\to\infty} \psi_{j_\ell} = \sum_{k=1}^{n} a_{j_\ell,k}\phi_k.$$

2. $M' = \infty$: Dies wird einen Widerspruch zur linearen Unabhängigkeit von $\phi_1,\dots,\phi_n$ liefern. Hierzu nehmen wir an, daß

$$\max_{k=1,\dots,n} |a_{jk}| > 0, \qquad j \in \mathbb{N}_0,$$

ansonsten wählen wir uns eine passende Teilfolge. Dieses Maximum wird für mindestens ein k unendlich oft angenommen, sagen wir, für $k = 1$, andernfalls numerieren wir unsere Basisfunktionen $\phi_1, \ldots, \phi_n$ um. Also gibt es eine Teilindizierung j_ℓ, so daß

$$|a_{j_\ell,1}| \geq |a_{j_\ell,k}|\,, \qquad k = 2, \ldots, n.$$

Sei $\widetilde{\psi}_\ell = \psi_{j_\ell}/a_{j_\ell,1}$. Da $|a_{j_\ell,1}| \to \infty$, erhalten wir daß

$$\left\|\widetilde{\psi}_\ell\right\| \leq \frac{\|\psi_{j_\ell}\|}{|a_{j_\ell,1}|} \leq \frac{M}{|a_{j_\ell,1}|} \to 0$$

und da

$$\widetilde{\psi}_\ell = \sum_{k=1}^{n} \frac{a_{j_\ell,k}}{a_{j_\ell,1}}\,\phi_k = \phi_1 + \sum_{k=2}^{n} \underbrace{\frac{a_{j_\ell,k}}{a_{j_\ell,1}}}_{|\cdot|\leq 1}\,\phi_k \quad \to \quad \phi_1 + \sum_{k=2}^{n} b_k\,\phi_k,$$

nach Teil 1, erhalten wir, daß

$$0 = \left\|\phi_1 + \sum_{k=2}^{n} b_k\,\phi_k\right\|, \qquad \Longrightarrow \qquad \phi_1 = -\sum_{k=2}^{n} b_k\,\phi_k,$$

im Widerspruch zur linearen Unabhängigkeit.

□

Übung 3.2 Zeigen Sie, daß die Abbildung $\|\cdot\| \;:\; F \to \mathbb{R}$ stetig ist. ♢

Definition 3.4 *Die Norm $\|\cdot\|$ bzw. den normierten Raum $(F, \|\cdot\|)$ bezeichnet man als* strikt konvex*, wenn*

$$\|f + f'\| = \|f\| + \|f'\| \qquad \Longrightarrow \qquad f' = \lambda f, \quad 0 \leq \lambda \in \mathbb{R}.$$

Bemerkung 3.5 *1. Die Namensgebung "strikt konvex" kommt daher, daß die Norm immer eine* konvexe Funktion *ist, d.h.,*

$$\|\alpha\, f + (1-\alpha)\, f'\| \leq \alpha\, \|f\| + (1-\alpha)\, \|f\|\,, \qquad \alpha \in [0,1], \quad f, f' \in F; \tag{3.1}$$

Gilt obige Zusatzforderung, dann ist die Norm eine strikt konvexe *Funktion, wenn man den "Trivialfall" $f' = f$ ausschließt, bei dem die strikte Ungleichung wegen der positiven Homogenität der Norm nicht gelten kann.*

2. Man kann es aber auch noch anders sehen: Die Einheitskugel

$$B_1 := \{f \in F \;:\; \|f\| \leq 1\}$$

bezüglich der Norm $\|\cdot\|$ ist eine konvexe Menge – das erhält man wieder mittels (3.1); ist nun die Norm strikt konvex, dann ist auch die Einheitskugel strikt konvex, ihr Rand enthält also keine "Geradenstücke".

3. *Man kann leicht sehen, daß auf dem $\mathbb{R}^n$ die Normen $\|\cdot\|_1$ und $\|\cdot\|_\infty$ "nur" konvex, die Normen $\|\cdot\|_p$, $1 < p < \infty$, hingegen strikt konvex sind, siehe Abb 3.1.*

4. *Mit Blick auf Proposition 3.6 können wir also davon ausgehen, daß die Approximationstheorie bezüglich strikt konvexer Normen auch strukturell anders aussehen wird, als die bezüglich "nur" konvexer Normen.*

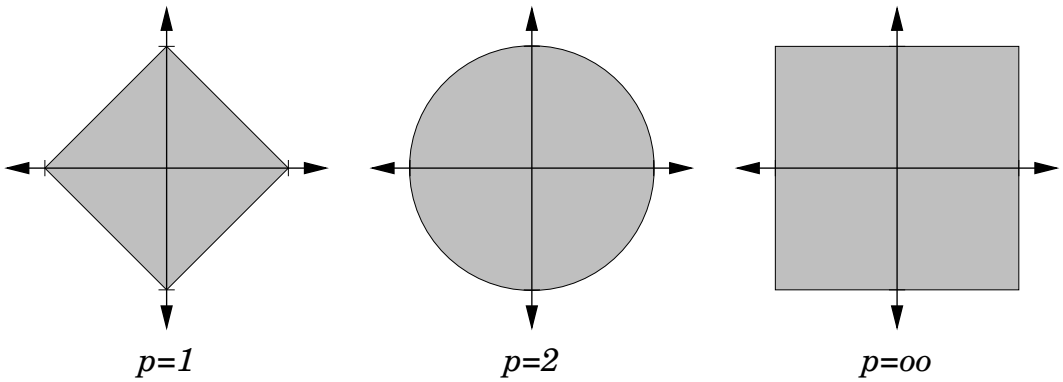


Abbildung 3.1: Die Einheitskugeln bezüglich der Normen $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$ im $\mathbb{R}^2$. Man sieht ganz gut, welche strikt konvex ist und welche es nicht sind.

Übung 3.3 Es sei $F = C(X)$, X ein kompakter Hausdorffraum. Zeigen Sie:

1. Die Norm
$$\|f\|_2 := \left(\int_X |f(t)|^2 \, dt \right)^{1/2}$$
ist strikt konvex.

2. Die Normen
$$\|f\|_1 := \int_X |f(t)| \, dt \qquad \text{und} \qquad \|f\|_\infty := \max_{x \in X} |f(x)|$$
sind nicht strikt konvex.

♢

Proposition 3.6 *Ist die Norm $\|\cdot\|$ strikt konvex, dann ist für jedes $f \in F$ die Bestapproximation eindeutig, also $\#P_\Phi(f) = 1$.*

Beweis: Nehmen wir an, zu $f \in F$ gäbe es zwei Bestapproximation $\phi \neq \phi' \in \Phi$. Wäre nun $f - \phi = \lambda\,(f - \phi')$ für ein $\lambda \in \mathbb{R}$, dann hieße dies, daß
$$f\,(1 - \lambda) = \phi - \lambda\phi' \qquad \Longrightarrow \qquad f \in \Phi \qquad \Longrightarrow \qquad \phi = \phi' = f,$$

was ein Widerspruch ist. Wegen der strikten Konvexität erhalten wir nun aber für jedes $\alpha \in (0,1)$, daß

$$d(f,\Phi) \le \|f - \alpha\,\phi - (1-\alpha)\,\phi'\| < \alpha \underbrace{\|f-\phi\|}_{=d(f,\Phi)} + (1-\alpha) \underbrace{\|f-\phi'\|}_{=d(f,\Phi)} = d(f,\Phi)\,,$$

und das ist natürlich ein Widerspruch. □

3.2 Das Kolmogoroff–Kriterium und extremale Signaturen

Jetzt werden wir wieder konkreter, und zwar sei nun X ein kompakter Hausdorffraum und $\Phi \subset C(X)$ ein endlichdimensionaler $\mathbb{C}$–Teilraum der komplexwertigen stetigen Funktionen auf X.

Das Kolmogoroff[42]–Kriterium aus [37] ist das stetige Gegenstück zum L_2–Bestapproximationskriterium (2.17) aus Lemma 2.22. Da allerdings nun die "Orthogonalität" nicht mehr bezüglich eines Integrals betrachtet wird, brauchen wir zuerst einen neuen Begriff.

Definition 3.7 *Für $f \in C(X)$ und $\phi \in \Phi$ bezeichnen wir mit*

$$E(f,\phi) := \{x \in X \;:\; |f(x) - \phi(x)| = \|f-\phi\|\}$$

die Extremalpunktmenge *von f und ϕ.*

Satz 3.8 *$\phi^* \in \Phi$ ist genau dann Bestapproximation an $f \in C(X)$ in Φ, wenn*

$$\max_{x \in E(f,\phi^*)} \Re\left((f(x) - \phi^*(x))\,\overline{\phi(x)}\right) \ge 0, \qquad \phi \in \Phi. \tag{3.2}$$

Bevor wir diesen Satz beweisen, sollten wir erst einmal versuchen, zu verstehen, was er uns sagt.

Bemerkung 3.9 *1. Indem wir alle ϕ durch $-\phi$ ersetzen, erhalten wir die zu (3.2) äquivalente*[43] *Bedingung*

$$\min_{x \in E(f,\phi^*)} \Re\left((f(x) - \phi^*(x))\,\overline{\phi(x)}\right) \le 0, \qquad \phi \in \Phi. \tag{3.3}$$

2. In der reellen Version bedeutet ja (3.2), daß

$$\max_{x \in E(f,\phi^*)} (f(x) - \phi^*(x))\;\phi(x) \ge 0, \qquad \phi \in \Phi. \tag{3.4}$$

[42] Andrey Nikolaevich Kolmogoroff (oder "Kolmogorov"), 1903–1987, trug wesentlich zu den Grundlagen der Wahrscheinlichkeitstheorie, aber auch zu Approximationstheorie, Topologie, Funktionalanalysis, Geometrie und so einigem mehr bei. Mit anderen Worten: einer der ganz, ganz großen Mathematiker des 20. Jahrhunderts!

[43] Optimierer würden fast schon "duale" sagen ...

3. *Dies liefert auch eine "Minimalforderung" an die Größe der Extremalpunktmenge und an das Verhalten der "Fehlerfunktion" $f - \phi^*$, denn diese muß auf der Extremalpunktmenge so viele Vorzeichenwechsel haben haben, daß es keine Funktion $\phi \in \Phi$ mehr gibt, die imstande ist, all diese Vorzeichenwechsel mitzumachen. Denn jedes $\phi \in \Phi$, das es schafft,*

$$\operatorname{sgn} \phi(x) = -\operatorname{sgn}\,(f(x) - \phi^*(x))\,, \qquad x \in E\,(f, \phi^*)\,,$$

zu erfüllen, "zerstört" ja auch (3.4).

Übung 3.4 Zeigen Sie: Ist $\Phi = \Pi_{n-1}$, dann besteht jede Extremalpunktmenge aus mindestens $n + 1$ Punkten. ♢

Beweis von Satz 3.8: Sei ϕ^* eine Bestapproximation und nehmen wir an, daß (3.2) nicht gilt, das heißt, es gibt ein $\varepsilon > 0$ und ein $\psi \in \Phi$, so daß

$$\max_{x \in E(f,\phi^*)} \Re\left((f(x) - \phi^*(x))\,\overline{\psi(x)}\right) = -2\varepsilon < 0$$

ist; wir werden einen Widerspruch generieren, indem wir mit Hilfe von ψ eine *bessere* Approximation an f als ϕ^* konstruieren, und zwar, indem wir die Familie

$$\phi_\lambda := \phi^* - \lambda\,\psi, \qquad \lambda \in \mathbb{R}_+,$$

betrachten. Da die Funktion $g := \Re\,(f(x) - \phi^*(x))\;\overline{\psi(x)}$ stetig ist, gibt es eine offene Menge G, $E\,(f, \phi^*) \subset G \subset X$, so daß $g(x) < -\varepsilon$, $x \in G$. Für $x \in G$ ist nun

$$\begin{aligned}
|f(x) - \phi_\lambda(x)|^2 &= |(f - \phi^*) + \lambda\psi|^2 \\
&= ((f(x) - \phi^*(x)) + \lambda\psi(x))\;\overline{((f(x) - \phi^*(x)) + \lambda\psi(x))} \\
&= (f(x) - \phi^*(x))\;\overline{(f(x) - \phi^*(x))} + \lambda^2\psi(x)\overline{\psi(x)} \\
&\qquad + \underbrace{\lambda\;(f(x) - \phi^*(x))\;\overline{\psi(x)} + \lambda\,\overline{(f(x) - \phi^*(x))}\,\psi(x)}_{=2\lambda\Re\left((f(x)-\phi^*(x))\,\overline{\psi(x)}\right)} \\
&= |f(x) - \phi^*(x)|^2 + \lambda^2\,|\psi(x)|^2 + 2\lambda\Re\left((f(x) - \phi^*(x))\;\overline{\psi(x)}\right) \\
&< |f(x) - \phi^*(x)|^2 + \lambda^2\,\|\psi\|^2 - 2\lambda\varepsilon
\end{aligned}$$

und wir erhalten für

$$\lambda < \frac{\varepsilon}{\|\psi\|^2} \qquad \Longrightarrow \qquad \lambda\left(\lambda\;\|\psi\|^2 - 2\varepsilon\right) < -\lambda\varepsilon$$

die Verbesserung

$$|f(x) - \phi_\lambda(x)|^2 < |f(x) - \phi^*(x)|^2 - \lambda\epsilon, \qquad x \in G. \tag{3.5}$$

Auf der abgeschlossenen (und somit kompakten) Menge $X \setminus G$ hingegen ist, da $E\,(f, \phi^*) \subset G$, der Fehler $f - \phi^*$ im Absolutbetrag immer *strikt kleiner* als $\|f - \phi^*\|$ und wegen der Kompaktheit gibt es ein $\delta > 0$, so daß

$$\max_{x \in X \setminus G} |f(x) - \phi^*(x)| \le \|f - \phi^*\| - \delta. \tag{3.6}$$

Ist nun zusätzlich[44] $\lambda < \frac{\delta}{2\|\psi\|}$, dann ist für jedes $x \in X \setminus G$

$$\begin{aligned} |f(x) - \phi_\lambda(x)| &= |f(x) - \phi^*(x) + \lambda\,\psi(x)| \le |f(x) - \phi^*(x)| + \lambda\,|\psi(x)| \\ &\le \|f - \phi^*\| - \delta + \lambda\,\|\psi\| < \|f - \phi^*\| - \delta + \frac{\delta}{2\|\psi\|}\|\psi\| = \|f - \phi^*\| - \frac{\delta}{2}, \end{aligned}$$

und somit erhalten wir, daß

$$0 < \lambda < \min\left\{\frac{\varepsilon}{\|\psi\|^2}, \frac{\delta}{2\|\psi\|}\right\} \qquad \Longrightarrow \qquad \|f - \phi_\lambda\| < \|f - \phi^*\|,$$

was den gewünschten Widerspruch liefert.
Für die Umkehrung nehmen wir an, daß (3.2) erfüllt ist und erhalten mit der nun schon bekannten Rechnung, daß für jedes $\phi \in \Phi$ und $x \in E(f, \phi^*)$

$$\begin{aligned} |f(x) - \phi(x)|^2 &= |(f(x) - \phi^*(x)) + (\phi^*(x) - \phi(x))|^2 \\ &= |f(x) - \phi^*(x)|^2 + |\phi^*(x) - \phi(x)|^2 + 2\underbrace{\Re\Big((f(x) - \phi^*(x))\overbrace{\overline{(\phi^*(x) - \phi(x))}}^{\in\Phi}\Big)}_{\ge 0} \\ &\ge |f(x) - \phi^*(x)|^2 = \|f - \phi^*\|^2, \end{aligned}$$

weswegen erst recht $\|f - \phi\| \ge \|f - \phi^*\|$ und somit ϕ^* eine Bestapproximation ist. □

Wir haben jetzt also eine Charakterisierung der Bestapproximation über das Vorzeichenverhalten auf der Extremalpunktmenge $E(f, \phi^*)$. Anders gesagt: Für jedes $\phi \in \Phi$ brauchen wir uns "nur" das Vorzeichenverhalten auf dieser Menge anzusehen und zu versuchen, ein $\psi \in \Phi$ zu finden das dort dasselbe Vorzeichenverhalten hat – gibt es so etwas, dann hat $-\psi$ das entgegengesetzte Vorzeichenverhalten und unser ϕ ist keine Bestapproximation, gibt es sowas nicht, dann ist ϕ die gesuchte Bestapproximation.

Leider gibt es da aber noch ein kleines Problem: Wir haben keine Ahnung, wie diese Extremalpunktmenge aussieht, ob sie endlich oder unendlich, abzählbar oder überabzählbar ist. Doch mit ein bißchen mehr Aufwand[45] kann man hier einiges mehr an Information erhalten.

Definition 3.10 *1. Eine* Signatur[46] *σ der Länge $r + 1 \in \mathbb{N}_0$ besteht aus Punkten $X_\sigma = \{x_0, \ldots, x_r\} \subset X$ und komplexen Zahlen*[47] *$\sigma_0, \ldots, \sigma_r \in \mathbb{C}$ mit der Eigenschaft, daß $|\sigma_j| = 1$, $j = 0, \ldots, r$.*

2. Eine Signatur σ heißt extremal *für $\Phi \subset C(X)$, wenn es zusätzlich* positive[48] *Zahlen $\mu_0, \ldots, \mu_r \in \mathbb{R}_+$ gibt, so daß*

$$\sum_{j=0}^{r} \mu_j\,\sigma_j\,\overline{\phi(x_j)} = 0, \qquad \phi \in \Phi. \tag{3.7}$$

[44] Wir bekommen zwei Bedingungen an λ und wählen einfach die kleinere der beiden.

[45] Den wir nun auch betreiben wollen!

[46] Der Name kommt nicht von "Unterschrift", sondern vom englischen Wort "sign" und bedeutet im wesentlich "Vorzeichenvorgabe".

[47] Den "Vorzeichen"; im rellen Fall ist $\sigma_j = \pm 1$, im komplexen ist $\sigma_j = e^{i\theta_j}$, $\theta_j \in \mathbb{T}$.

[48] Und damit reelle! Schließlich ist $\mathbb{C}$ ja kein archimedisch geordneter Körper mehr.

3. *Da jede komplexe Zahl via Real- und Imaginärteil als relle Zahl aufgefasst werden kann, können wir jeden komplexen Vektorraum der Dimension n auch als rellen Vektorraum der Dimension $2n$ auffassen. Die* reelle Dimension *eines solchen Vektorraums bezeichnen wir als*

$$\dim_{\mathbb{R}} V = \begin{cases} n, & V \simeq \mathbb{R}^n, \\ 2n, & V \simeq \mathbb{C}^n. \end{cases}$$

Diese Notation überträgt sich auch auf endlichdimensionale Funktionenräume, denn schließlich ist ja entweder $\Phi \simeq \mathbb{R}^n$ oder $\Phi \simeq \mathbb{C}^n$.

Solche extremalen Signaturen sind nun die Antwort auf unsere Frage nach den Extremalmengen, die im folgenden Satz von Rivlin und Shapiro [63] gegeben wurde – insbesondere, was die *Endlichkeit* der zu betrachtenden Extremalmengen angeht.

Satz 3.11 (Satz von Rivlin und Shapiro) *Sei $f \in C(X)$ und $f \neq \phi^* \in \Phi$. Dann ist ϕ^* genau dann eine Bestapproximation an f, wenn es eine extremale Signatur σ der Länge $r+1$, $r \leq \dim_{\mathbb{R}} \Phi$ für Φ gibt mit $X_\sigma \subset E(f, \phi^*)$ und*

$$\operatorname{sgn}\left(f(x_j) - \phi^*(x_j)\right) = \sigma_j, \qquad j = 0, \ldots, r. \tag{3.8}$$

Wieder einmal wird der Beweis ein klein wenig mehr Arbeit machen. Diesmal brauchen wir Hilfsmittel aus der Konvexitätstheorie – das ist übrigens kein Zufall, wenn man sich den Titel der Arbeit [63] ansieht, denn schließlich hat lineare Optimierung eine ganze Menge mit Konvexität zu tun!

Definition 3.12 *Seien $x_0, \ldots, x_r \in X$.*

1. *Ein Element $x \in X$ heißt* Konvexkombination *von $x_0, \ldots, x_r$, wenn*

$$x = \sum_{j=0}^{r} \alpha_j x_j, \qquad \alpha_j \geq 0,\ j = 0, \ldots, r, \quad \sum_{j=0}^{r} \alpha_j = 1. \tag{3.9}$$

2. *Mit Δ_r bezeichnen wir das r–dimensionale* Einheitssimplex

$$\Delta_r := \left\{\alpha = (\alpha_0, \ldots, \alpha_r) \in \mathbb{R}^{r+1} \ :\ \alpha_j \geq 0,\ \alpha_0 + \cdots + \alpha_r = 1\right\}$$

und sein Inneres, Δ_r°, mit

$$\Delta_r^\circ := \left\{\alpha \in \Delta_r \ :\ \alpha_j > 0,\ j = 0, \ldots, r.\right\}.$$

3. *Die* konvexe Hülle *von $Y \subset X$ ist definiert als*

$$[Y] := \left\{\alpha\, y + (1-\alpha)\, y' \ :\ \alpha \in [0,1],\ y, y' \in Y\right\}$$

4. *$Y \subset X$ heißt* konvex, *wenn $Y = [Y]$.*

Übung 3.5 Zeigen Sie, daß

$$[Y] = \left\{ \sum_{j=0}^{r} \alpha_r \, y_r \ : \ y_j \in Y, \alpha \in \Delta_r, \, r \in \mathbb{N}_0. \right\}$$

♢

Als erstes eine kleine Hilfsaussage über die Darstellung konvexer Hüllen im $\mathbb{R}^n$.

Lemma 3.13 *Für $Y \subset \mathbb{R}^n$ ist*

$$[Y] = \left\{ \sum_{j=0}^{r} \alpha_j \, y_j \ : \ y_j \in Y, \, \alpha \in \Delta_r^\circ, \, r \leq n \right\} \tag{3.10}$$

Beweis: Zwei Dinge sind in diesem Lemma entscheidend und zwar die Darstellung als *strikte* Konvexkombination, d.h., $\alpha \in \Delta_r^\circ$, und die obere Schranke für die Anzahl der Punkte, die konvex kombiniert werden sollen[49]. Da die rechte Seite von (3.10) eine konvexe Menge ist, die Y enthält, ist $\subseteq$ klar.
Sei also $y \in [Y]$; nach Übung 3.5 gibt es (mindestens) eine Darstellung von y der Form (3.10), zumindest, nachdem man alle Nullkoeffizienten eliminiert hat. Wenn es mehrere Darstellungen gibt, dann wählen wir eine, bei der r minimal wird. Nehmen wir nun an, daß $r > n$ ist, dann sind die Vektoren $y_j - y_0$, $j = 1, \ldots, r$, linear abhängig und daher gibt es nichttriviale Koeffizienten $\beta_1, \ldots, \beta_r$, so daß

$$0 = \sum_{j=1}^{r} \beta_j \ (y_j - y_0) = -\left(\sum_{j=1}^{r} \beta_j \right) \ y_0 + \sum_{j=1}^{r} \beta_j \, y_j =: \sum_{j=0}^{r} \gamma_j \, y_j, \qquad \sum_{j=0}^{r} \gamma_j = 0,$$

und mindestens ein γ_j ist von Null verschieden, sagen wir γ_1. Nun ist für jedes $\lambda \in \mathbb{R}$

$$y = y + 0 = \sum_{j=0}^{r} \alpha_j \, y_j + \lambda \sum_{j=0}^{r} \gamma_j \, y_j = \sum_{j=0}^{r} \underbrace{(\alpha_j + \lambda \gamma_j)}_{=:\widetilde{\alpha}_j} \, y_j, \qquad \widetilde{\alpha} \in \Delta_r,$$

aber durch $\lambda = -\frac{\alpha_1}{\gamma_1}$ erhalten wir, daß $\widetilde{\alpha}_1 = 0$ ist, was der Minimalität von r widerspricht. □

Als nächstes ein paar Umformulierungen der Kolmogoroff–Kriteriums (3.2).

Proposition 3.14 *Für $\phi^* \in \Phi$ und $f \in C(X)$ sind äquivalent:*

1. *ϕ^* ist Bestapproximation an f in Φ.*
2. *Wir haben, daß*

$$0 \in \left[\bigcup_{x \in E(f, \phi^*)} \left[(f(x) - \phi^*(x)) \, \overline{\phi_j(x)} \ : \ j = 1, \ldots, n \right] \right]. \tag{3.11}$$

[49] Was geometrisch nichts anderes bedeutet, als daß man redundante Punkte wegwirft.

3. *Es gibt Punkte* $x_0, \ldots, x_r \in E(f, \phi^*)$ *und* $\alpha \in \Delta_r^\circ$, $r \leq \dim_{\mathbb{R}} \Phi$, *so daß*

$$\sum_{j=0}^{r} \alpha_j \left(f(x_j) - \phi^*(x_j)\right) \overline{\phi_k(x_j)} = 0, \qquad k = 1, \ldots, n. \tag{3.12}$$

4. *Es gibt Punkte* $x_0, \ldots, x_r \in E(f, \phi^*)$ *und* $\alpha \in \Delta_r^\circ$, $r \leq \dim_{\mathbb{R}} \Phi$, *so daß*

$$\sum_{j=0}^{r} \alpha_j \left(f(x_j) - \phi^*(x_j)\right) \overline{\phi(x_j)} = 0, \qquad \phi \in \Phi. \tag{3.13}$$

Bemerkung 3.15 *Die Bedingung (3.13) zeigt wohl die meiste Ähnlichkeit mit der L_2–Charakterisierung (2.17), nur hat man hier ein inneres Produkt*[50]

$$\langle f, g \rangle = \sum_{j=0}^{r} \alpha_j f(x_j) \overline{g(x_j)}, \qquad f, g \in C(X)$$

mit Summation über Punktauswertungen ("diskrete Maße"), die allerdings von ϕ^* und f abhängen.

Proposition 3.16 (Trennhyperebenensatz) *Ist $\Omega \subset \mathbb{R}^n$ abgeschlossen und konvex und $y \notin \Omega$, dann gibt es $a \in \mathbb{R}^n$ und $c \in \mathbb{R}$, so daß*

$$a^T y + c < 0 \leq a^T \Omega + c := \left\{a^T w + c \ : \ x \in \Omega\right\}. \tag{3.14}$$

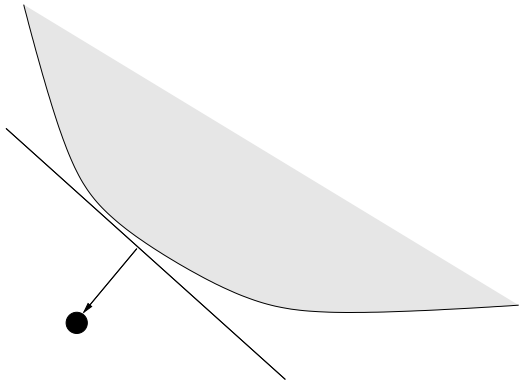


Abbildung 3.2: Die Idee der Trennhyperebene: Die konvexe Menge liegt immer auf *einer* Seite jeder Hyperebene, die mit dieser Menge nichtleeren Schnitt hat. Gehört nun ein Punkt nicht dazu, dann kann man immer eine Hyperebene finden, so daß dieser Punkt auf einer Seite dieser Hyperebene liegt und die konvexe Menge auf der anderen.

[50] Allerdings mit der Einschränkung, daß $\langle f, f \rangle = 0$ sein kann für manche $f \in C(X)$.

Beweis: Da die euklidische Norm strikt konvex ist, liefert uns Proposition 3.6 die Existenz[51] einer *eindeutigen* Bestapproximation w^* aus Ω an y. Diese zeichnet sich dann durch

$$\begin{aligned} 0 &< \|y-w\|_2^2 - \|y-w^*\|_2^2 = \|w\|_2^2 - \|w^*\|_2^2 + 2y^T(w^*-w) \\ &= (w+w^*)^T(w-w^*)^T - 2y^T(w-w^*) = ((w-y)+(w-y))(w-w^*). \end{aligned}$$

Ersetzen wir w durch $\alpha w + (1-\alpha)w^*$, $\alpha \in [0,1]$, dann erhalten wir

$$0 < (\alpha(w-y) + (2-\alpha)(w^*-y))\,\alpha(w-w^*),$$

was wir durch 2α dividieren können, um dann durch Grenzübergang $\alpha \to 0$ "zu Fuß" das entsprechende Kolmogoroff–Kriterium[52]

$$0 \le (w^*-y)^T(w-w^*) = a^T w + c, \qquad a := w^*-y, \quad c := -a^T w^*,$$

zu erhalten, was uns sagt, daß $a^T\Omega \ge 0$ ist. Andererseits ist, mit $w = y$,

$$a^T y + c = (w^*-y)^T(y-w^*) = -\|y-w^*\|_2^2 < 0.$$

□

Beweis von Proposition 3.14: Wir beginnen mit (3.11) und bemerken zuerst, daß die Abbildung

$$g : X \to Z = \begin{cases} \mathbb{R}^n, \\ \mathbb{C}^n, \end{cases} \qquad g(x) = \left[(f(x)-\phi^*(x))\,\overline{\phi_j(x)} \;:\; j=1,\ldots,n \right],$$

stetig ist und die kompakte Menge $E(f,\phi^*)$ auf eine kompakte Menge $Y = g(E(f,\phi^*)) \subset Z$ abbildet. Also ist $[Y]$, die Menge, für die wir uns interessieren, eine konvexe und kompakte Menge im $\mathbb{R}^n$ oder im $\mathbb{C}^n$, je nachdem, ob wir den rellen oder den komplexen Fall betrachten.

1. *Der relle Fall:* Da $[Y]$ der Durchschnitt aller Halbräume ist, die $[Y]$ enthalten, siehe Übung 3.6, ist $0 \notin [Y]$ dazu äquivalent, daß es eine *Trennhyperebene* H gibt, so daß $0 \notin H$ und $Y \subset [Y] \subset H$, siehe Abb. 3.2.

   Das ist aber äquivalent dazu, daß es $c \in \mathbb{R}$ und $a \in \mathbb{R}^n$ gibt, so daß $c \ge \langle y, a\rangle$, $y \in Y$, also

$$0 > c \ge \max_{y\in Y}\langle y,a\rangle = \max_{y\in Y} a^T y = \max_{x\in E(f,\phi^*)} (f(x)-\phi^*(x)) \underbrace{\left(\sum_{j=1}^n \phi_j(x)\,a_j\right)}_{=:\phi(x)},$$

   was genau die Negation von (3.4) ist, also äquivalent dazu ist, daß ϕ^* keine Bestapproximation ist.

[51] Der Beweis nutzt nur Konvexität, nicht aber Approximation durch einen linearen Unterraum.

[52] Man beachte: Eigentlich ist der Trennhyperebenensatz eine **direkte** Konsequenz des Kolmogoroff–Kriteriums im $\mathbb{R}^n$.

2. *Der komplexe Fall* geht praktisch analog, nur werden jetzt die Halbräume durch das bilineare reellwertige Skalarprodukt $\Re\langle \cdot, \cdot \rangle$ bestimmt:

$$H(y,a) = \left\{ z \in \mathbb{C}^n \ : \ \Re\left(\sum_{j=1}^{n} \overline{z_j}\, y_j \right) \le a \right\}, \qquad y \in \mathbb{C}^n, \quad a \in \mathbb{R}.$$

Die Bedingung (3.12) folgt nun aus einer komponentenweisen Betrachtung von (3.11) und Lemma 3.13, und (3.13) ist offensichtlich äquivalent zu (3.12). □

Übung 3.6 Als *Halbräume* im $\mathbb{R}^r$ bezeichnet man Mengen der Form

$$H = H(y,a) = \left\{ x \in \mathbb{R}^r \ : \ y^T x \le a \right\}, \qquad y \in \mathbb{R}^r, \quad a \in \mathbb{R}.$$

Zeigen Sie: Ist $Y \subset \mathbb{R}^r$ kompakt, dann ist

$$[Y] = \bigcap_{Y \subset H} H.$$

♢

Jetzt aber endlich, zum Abschluß dieses Kapitels, der Beweis, der noch fehlt.
Beweis von Satz 3.11: Sei ϕ^* ein Bestapproximant. Mit den Punkten $x_0, \ldots, x_r$ aus Proposition 3.14 setzen wir $X_\sigma = \{x_0, \ldots, x_r\}$, schreiben

$$\alpha_j \left(f(x_j) - \phi^*(x_j) \right) = \underbrace{\alpha_j \ |f(x_j) - \phi^*(x_j)|}_{=:\mu_j} \underbrace{e^{i\theta_j}}_{=:\sigma_j}, \qquad \mu_j > 0,\ |\sigma_j| = 1, \qquad j = 0, \ldots, r,$$

und nach (3.13) ist dies die gewünschte extremale Signatur.
Sei nun σ eine extremale Signatur der Länge $r+1$, $r \le \dim_{\mathbb{R}} \Phi$, dann ist für jedes $\phi \in \Phi$

$$0 = \sum_{j=0}^{r} \mu_j\, \sigma_j\, \overline{(\phi^*(x_j) - \phi(x_j))}$$

und somit, nach unseren Annahmen an die Signatur

$$\begin{aligned} \|f - \phi\| \sum_{j=0}^{r} |\mu_j| \ &\ge \ \sum_{j=0}^{r} |f(x_j) - \phi(x_j)| \ \underbrace{|\sigma_j|}_{=1} \ |\mu_j| \ge \left| \sum_{j=0}^{r} \overline{(f(x_j) - \phi(x_j))} \mu_j\, \sigma_j \right| \\ &= \ \Big| \sum_{j=0}^{r} \overline{(f(x_j) - \phi^*(x_j))}\, \mu_j\, \sigma_j + \underbrace{\sum_{j=0}^{r} \overline{(\phi^*(x_j) - \phi(x_j))}\, \mu_j\, \sigma_j}_{=0} \Big| \\ &= \ \sum_{j=0}^{r} \underbrace{|f(x_j) - \phi^*(x_j)|}_{=\|f-\phi^*\|} \mu_j \ |\sigma_j| = \|f - \phi^*\| \sum_{j=0}^{r} |\mu_j| \end{aligned}$$

und da die μ_j alle *strikt* positiv waren, muß ϕ^* eine Bestapproximation sein. □

3.3 Haar–Räume und Alternanten

Als nächstes befassen wir uns mit der folgenden Frage:

> *Wie müssen Teilräume $\Phi \subset C(X)$ beschaffen sein, damit man für* jedes $f \in C(X)$ genau ein *Element bester Approximation hat?*

Definition 3.17 *Ein Raum $\Phi \subset C(X)$ heißt* Haar[53]–Raum *oder* Tschebyscheff–Raum*, wenn $\#P_\Phi(f) = 1$ für jedes $f \in C(X)$. Eine Basis von Φ heißt* Tschebyscheff–System*, wenn Φ ein Haar–Raum ist.*

Haar–Räume haben also die schöne und gewünschte Eigenschaft, daß die Bestapproximation stets eindeutig ist, nur haben wir mit Definition 3.17 noch nicht allzuviel gewonnen, sondern nur der Eigenschaft, die uns interessiert, einen anderen Namen gegeben. Allerdings gibt es eine weitere Beschreibung der Haar–Räume, die zumindest teilweise auf Haar selbst [27] zurückgeht.

Satz 3.18 *Sei X ein kompakter Hausdorffraum, $\#X \geq n+1$, und sei Φ ein n–dimensionaler Teilraum von $C(X)$. Dann sind äquivalent:*

1. *Φ ist ein Haar–Raum.*
2. *Eindeutige Interpolation mit Φ ist immer möglich: Zu je n Punkten $x_1, \ldots, x_n \in X$ und Werten $y_1, \ldots, y_n \in \mathbb{C}$ gibt es genau ein $\phi \in \Phi$, so daß*

$$\phi(x_j) = y_j, \qquad j = 1, \ldots, n.$$

3. *jedes $0 \neq \phi \in \Phi$ besitzt höchstens $n-1$ Nullstellen, d.h.,*

$$\#Z(\phi) = \#\{x \in X \,:\, \phi(x) = 0\} < n, \qquad \phi \in \Phi. \tag{3.15}$$

Beweis: Beginnen wir mit der (einfachen) Äquivalenz von 2) und 3). In der Tat ist (3.15) dazu äquivalent, daß für[54] $x_1, \ldots, x_n \in X$ die einzige Lösung des homogenen Interpolationsproblems $\phi(x_j) = 0$ die Funktion $\phi = 0$ sein kann. Das ist aber wiederum dazu äquivalent, daß für eine[55] Basis $\phi_1, \ldots, \phi_n$ von Φ

$$0 \neq \det \begin{bmatrix} \phi_1(x_1) & \ldots & \phi_1(x_n) \\ \vdots & \ddots & \vdots \\ \phi_n(x_1) & \ldots & \phi_n(x_n) \end{bmatrix},$$

[53] Alfréd Haar, 1885–1933, ungarischer Mathematiker, studierte in Göttingen bei Hilbert und ist neben dem Haar–Raum auch durch *Haar–Maße* (invariante Maße auf Gruppen) und die *Haar–Wavelets* bekannt.

[54] Natürlich *disjunkte* Punkte

[55] Und damit für jede! Warum?

was wiederum äquivalent dazu ist, daß das Interpolationsproblem eindeutig lösbar ist.
Nehmen wir nun an, daß 2) oder, was dasselbe ist, 3) erfüllt ist und sei für $f \in C(X)$ wieder $\phi^* \neq f$ eine Bestapproximation an f. Dann muß $\#E(f,\phi) \geq n+1$ sein, denn sonst gäbe es ein Interpolationspolynom ϕ mit der Eigenschaft[56]

$$\phi(x) = -(f(x) - \phi^*(x)), \qquad x \in E(f,\phi^*),$$

und dann ist

$$\max_{x\in E(f,\phi^*)} \Re\left((f(x)-\phi^*(x))\,\overline{\phi(x)}\right) = \max_{x\in E(f,\phi^*)} |f(x)-\phi^*(x)|^2 < 0,$$

im Widerspruch zum Kolmogoroff–Kriterium aus Satz 3.8. Um zu zeigen, daß ϕ^* tatsächlich eindeutig ist, nehmen wir an, ψ^* wäre eine weitere Bestapproximation an f. Dann ist auch $\psi = \frac{1}{2}(\phi^* + \psi^*)$ eine Bestapproximation an f und natürlich ist

$$d := \|f-\phi^*\| = \|f-\psi^*\| = \|f-\psi\|.$$

Für jedes $x \in E(f,\psi)$ erhalten wir nun aber, daß

$$d = |f(x)-\psi(x)| \leq \frac{1}{2}\underbrace{|f(x)-\phi^*(x)|}_{\leq d} + \frac{1}{2}\underbrace{|f(x)-\psi^*(x)|}_{\leq d} \leq d,$$

also ist

$$(\phi^*-\psi^*)(x) = 0, \qquad x \in E(f,\psi),$$

und da $\#E(f,\psi) \geq n+1$ ist, muß $\phi^* = \psi^*$ sein.
Sei nun Φ ein Haar–Raum; nun kommt der "Haarige" Teil[57], nämlich der Nachweis der Interpolationseigenschaft. Nehmen wir also an, 2) wäre nicht erfüllt, das heißt, es gäbe Punkte $x_1,\ldots,x_n \in X$ und ein $\psi \in \Phi$, so daß $\psi(x_j) = 0$, $j = 1,\ldots,n$. Für eine Basis $\phi_1,\ldots,\phi_n$ von Φ erhalten wir dieses ψ als Lösung eines linearen Gleichungssystems:

$$0 = a^T V = [a_1 \ldots a_n] \begin{bmatrix} \phi_1(x_1) & \ldots & \phi_1(x_n) \\ \vdots & \ddots & \vdots \\ \phi_n(x_1) & \ldots & \phi_n(x_n) \end{bmatrix} \qquad \Longrightarrow \qquad \psi := \sum_{j=1}^n a_j\,\phi_j.$$

Da dieses System eine nichttriviale Lösung $a \neq 0$ besitzt, muß es auch ein $c \neq 0$ geben[58], so daß $Vc = 0$ ist, d.h.,

$$\sum_{j=1}^n c_j \phi_k(x_j) = 0, \quad k = 1,\ldots,n \qquad \Longleftrightarrow \qquad \sum_{j=1}^n c_j\,\phi(x_j) = 0, \quad \phi \in \Phi. \tag{3.16}$$

[56] Sollte $\#(f,\phi) < n$ sein, dann suchen wir uns einfach noch ein paar beliebige Punkte und schreiben dort beliebige Werte vor.

[57] Nicht weil der Beweis übermäßig kompliziert wäre, sondern weil das wohl der Teil ist, der in [27] bewiesen wurde – allerdings ist dies nur Hörensagen aus [47].

[58] Die Matrix ist quadratisch und wenn sie "von links" einen Rangdefekt hat, dann auch "von rechts". Der Standardspruch aus der linearen Algebra lautet: *Zeilenrang = Spaltenrang.*

Sei nun $f^* \in C(X)$ eine beliebige stetige Funktion, die die beiden Eigenschaften

$$f^*(x_j) = \frac{\overline{c_j}}{|c_j|}, \quad j = 1, \ldots, n, \qquad \text{und} \qquad \|f^*\| = 1 \tag{3.17}$$

besitzt – beispielsweise der stückweise lineare Interpolant. Nun setzen wir mittels unserer "Nulllösung" ψ

$$g := \left(1 - \frac{|\psi|}{\|\psi\|}\right) f^*,$$

und erhalten, daß

$$g(x_k) = f^*(x_k) = \frac{\overline{c_j}}{|c_j|} \qquad \text{und} \qquad \|g\| \leq \left\|1 - \frac{|\psi|}{\|\psi\|}\right\| \|f^*\| \leq 1, \tag{3.18}$$

also ist nach Lemma 3.19, das wir gleich beweisen werden, $d(g, \Phi) = 1$. Da für jedes $0 < \lambda < \|\psi\|^{-1}$ und jedes $x \in X$

$$\begin{aligned} |g(x) - \lambda\,\psi(x)| &\leq |g(x)| + \lambda\,|\psi(x)| \leq \left|1 - \frac{|\psi(x)|}{\|\psi\|}\right| \underbrace{|f^*(x)|}_{\leq \|f\| \leq 1} + \lambda\,|\psi(x)| \\ &\leq 1 - \underbrace{\left(\frac{1}{\|\psi\|} - \lambda\right)}_{>0} |\psi(x)| \leq 1 = d(g, \Phi), \end{aligned}$$

ist, wäre $\lambda\psi$ eine Bestapproximation für alle hinreichend kleinen Werte von λ, was nicht so ganz verträglich ist mit der Existenz einer *eindeutigen* Bestapproximation wäre, weswegen 2) eben doch erfüllt sein muß. □

Um den Beweis von Satz 3.18 zu vervollständigen, fehlt also nur noch das folgende Resultat.

Lemma 3.19 *Für jede stetige Funktion $f^* \in C(X)$, die (3.18) erfüllt, ist $d(f^*, \Phi) = 1$.*

Beweis: Da $1 = \|f^*\| = \|f^* - 0\|$ ist offensichtlich $d(f^*, \Phi) \leq 1$. Gäbe es nun ein $\phi \in \Phi$, so daß $\|f - \phi\| < 1$, dann ist für $k = 1, \ldots, n$

$$1 > |f(x_k) - \phi(x_k)|^2 = \underbrace{|f(x_k)|^2}_{=1} + |\phi(x_k)|^2 - 2\Re\left(\phi(x_k)\,\overline{f(x_k)}\right),$$

weswegen für $k = 1, \ldots, n$

$$0 < \Re\left(\phi(x_k)\,\overline{f(x_k)}\right) = \Re\left(\frac{c_k}{|c_k|}\,\phi(x_k)\right) \qquad \Longrightarrow \qquad \Re\left(c_k\,\phi(x_k)\right) > 0$$

sein muss. Schreiben wir andererseits ϕ als $\phi = \sum_j a_j \phi_j$, dann liefert uns (3.16), daß

$$\sum_{k=1}^{n} c_k\,\phi(x_k) = \sum_{j,k=1}^{n} a_j c_k\,\phi_j(x_k) = \sum_{j=1}^{n} a_j \underbrace{\sum_{k=1}^{n} c_k \phi_j(x_k)}_{=0} = 0,$$

was natürlich ein Widerspruch zu

$$0 < \sum_{k=1}^{n} \Re\left(c_k\,\phi\left(x_k\right)\right) = \Re\left(\sum_{k=1}^{n} c_k\,\phi\left(x_k\right)\right), \qquad j = 1, \ldots, n,$$

ist. □

Wie man sieht – Haar–Räume sind etwas feines, sie liefern nicht nur eindeutige Interpolation, sondern sind, so ganz "nebenbei" auch gleich noch gerade diejenigen Räume, die eindeutig interpolieren können. Damit können wir ein paar Beispiele für Haar–Räume angeben.

Beispiel 3.20 *1. Die Polynomräume Π_n, $n \in \mathbb{N}$, sind Haar–Räume für jedes kompakte $X \subset \mathbb{R}$.*

2. Die trigonometrischen Polynome der Ordnung n, also

$$\{1, \sin x, \cos x, \ldots, \sin nx, \cos nx\}$$

sind ein Tschebyscheff–System auf $\mathbb{T}$.

3. Die Räume

$$\operatorname{span}_{\mathbb{R}}\{\sin x, \ldots, \sin nx\} \qquad \textit{und} \qquad \operatorname{span}_{\mathbb{R}}\{1, \cos x, \ldots, \cos nx\}$$

sind Haar–Räume für jedes kompakte $X \subset [0, \pi]$.

4. Für je n verschiedene Werte $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ sind die Räume

$$\Lambda_n := \operatorname{span}_{\mathbb{R}}\left\{e^{\lambda_j \cdot} \;:\; j = 1, \ldots, n\right\}$$

Haar–Räume für jedes kompakte $X \subset \mathbb{R}$.

Beweis: Eigenschaft 1) ist die wohlbekannte Interpolationsfähigkeit der Polynome, für 2) identifizieren wir $\mathbb{T}$ mit $\mathbb{T}^* = \{z \in \mathbb{C} \;:\; |z| = 1\}$ via $z = e^{ix}$, $x \in \mathbb{T}$, $z \in \mathbb{T}^*$. Da

$$\cos x = \frac{1}{2}\left(e^{ix} + e^{-ix}\right) = \frac{1}{2}\left(z + z^{-1}\right) \qquad \sin x = \frac{1}{2i}\left(e^{ix} - e^{-ix}\right) = \frac{1}{2i}\left(z - z^{-1}\right),$$

können wir die trigonometrischen Polynome auf $\mathbb{T}$ mit den *Laurentpolynomen* auf $\mathbb{T}^*$ indentifizieren, das heißt, mit den "Polynomen" der Form

$$f(z) = \sum_{j=-n}^{n} a_j z^j = z^{-n} \sum_{j=0}^{2n} a_{j-n} z^j,$$

die auf $\mathbb{T}^*$ wohldefiniert sind und dort brav interpolieren können.

Eigenschaft 3) ist eine einfache[59] Übungsaufgabe, siehe Übung 3.7, nur 4) ist etwas interessanter[60] im Beweis. Hierbei zeigen wir durch Induktion über n, daß jede Funktion der Form

$$f_n(x) = \sum_{j=1}^{n} c_j \, e^{\lambda_j x}, \qquad x \in \mathbb{R}, \quad c_j \in \mathbb{R},\, j = 1, \ldots, n,$$

höchstens $n-1$ Nullstellen haben kann, was für $n = 1$ ziemlich offensichtlich ist. Hätte nun die Funktion

$$f_{n+1}(x) = \sum_{j=1}^{n+1} c_j \, e^{\lambda_j x}$$

mindestens $n+1$ Nullstellen, dann hätte auch

$$g_{n+1}(x) := e^{-\lambda_{n+1} x} \, f_{n+1}(x) = \sum_{j=1}^{n} c_j e^{(\lambda_j - \lambda_{n+1})x} + c_{n+1}, \qquad x \in \mathbb{R},$$

mindestens $n+1$ Nullstellen, also hätte, nach dem Satz von Rolle,

$$g'_{n+1}(x) = \sum_{j=1}^{n} (\lambda_j - \lambda_{n+1}) \; c_j \, e^{(\lambda_j - \lambda_{n+1})x}, \qquad x \in \mathbb{R},$$

mindestens n Nullstellen, im Widerspruch zur Indukionsannahme. □

Übung 3.7 Zeigen Sie, daß $1, \cos x, \ldots, \cos nx$, $n \in \mathbb{N}$, und $\sin x, \ldots, \sin nx$, Tschebyscheff–Systeme auf $[0, \pi]$ sind. ♢

Man sieht also, Haar–Räume gibt es tatsächlich! Allerdings bedeutet, was man zuerst mal nicht so einfach sieht, die Existenz von Haar–Räumen eine sehr starke Einschränkung an den zugrundeliegenden metrischen Raum X, nämlich, daß er entweder ein Intervall[61] oder aber der Torus sein muß.

Beispiel 3.21 *Enthält $X \subset \mathbb{R}^2$ eine offene Menge oder eine Verzweigung, dann gibt es keinen Haar–Raum über X.*

Beweis: Wir wählen zu $x_1, \ldots, x_n$ zwei stetige Funktion $u, v : [0,1] \to X$ so, daß

$$u(0) = v(1) = x_1, \qquad u(1) = v(0) = x_2, \qquad u(t) \neq v(t), \quad t \in [0,1],$$

und $u(t), v(t) \notin \{x_3, \ldots, x_n\}$. Mit anderen Worten: u und v vertauschen x_1 und x_2 ohne daß irgendwann ein doppelter oder mehrfacher Punkt in der Liste $u(t), v(t), x_3, \ldots, x_n$ auftaucht. Setzen wir nun

$$D(t) = \det \begin{vmatrix} \phi_1(u(t)) & \ldots & \phi_n(u(t)) \\ \phi_1(v(t)) & \ldots & \phi_n(v(t)) \\ \phi_1(x_3) & \ldots & \phi_n(x_3) \\ \vdots & \ddots & \vdots \\ \phi_1(x_n) & \ldots & \phi_n(x_n) \end{vmatrix},$$

[59] Hoffentlich!

[60] Dieser, wie ich finde, wirklich nette Beweis stammt aus [56]

[61] Oder eine Vereinigung von endlch vielen Intervallen.

dann ist $D(t)$ eine stetige Funktion in t, die die Eigenschaft $D(0) = -D(1)$ hat, weswegen es ein $t^* \in [0,1]$ geben muß, so daß $D(t^*) = 0$ und somit eindeutige Interpolation an $u(t^*), v(t^*), x_3, \ldots, x_n$ unmöglich ist. Dieses Vorgehen ist in Abb. 3.3 grafisch dargestellt. □

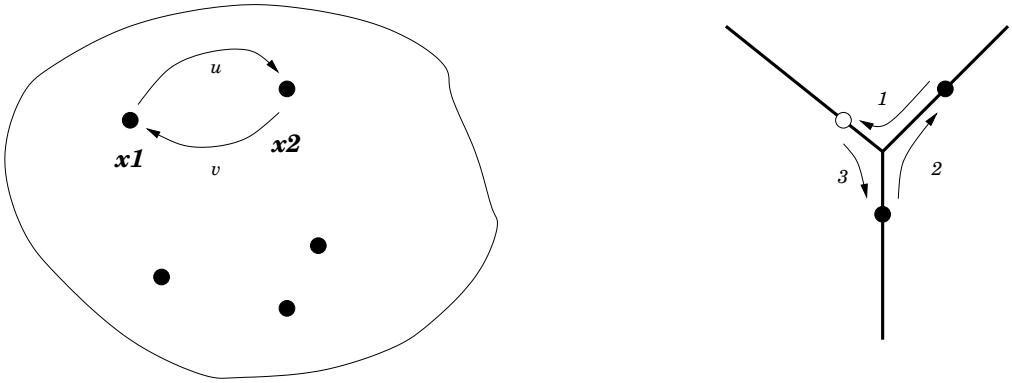


Abbildung 3.3: Die Vertauschung der Punkte x_1 und x_2. Irgendwo dazwischen muß was schiefgehen. Rechts die Art, wie man beim Vorliegen einer Verzweigung verfährt (siehe z.B. [47, S. 25]), man vertauscht "nach Art des Rangierbahnhofs", indem man zuerst x_1 aufs "Abstellgleis" verschiebt, dann x_2 dahin, wo er hingehört und schließlich x_1 an seine Zielposition bringt.

Diese Beobachtung schränkt natürlich unsere Möglichkeiten mit Haar–Räumen zu arbeiten, drastisch ein, denn im wesentlichen funktioniert die Interpolation nur, wenn X keine "Verzweigungen" besitzt, denn ansonsten können wir immer zwei Punkte durch Rangieren vertauschen. Offenbar haben Intervalle und der Torus diese Eigenschaften[62], aber daß dies wirklich die einzigen sind, erfordert ein klein wenig mehr an topologischen Überlegungen. Trotzdem gilt das folgende Resultat, das wir hier aber nicht beweisen wollen[63], der schon klassische Satz von Mairhuber[64] aus dem Jahre 1956.

Satz 3.22 (Satz von Mairhuber) *[52]*
Ein kompakter metrischer Raum X besitzt genau dann nichttriviale Haar–Räume in $C(X)$, wenn X homöomorph[65] zu einer kompakten Teilmenge von $\mathbb{T}$ ist.

Für den Rest dieses Kapitels gehen wir davon aus, daß wir es mit *reellwertigen* Funktionen und entsprechend auch mit *rellen* Haar–Räumen bzw. Tschebyscheff–Systemen zu tun haben, aber auch, daß unser metrischer Raum X mindestens $\dim \Phi + 1$ Punkte enthält.

[62] Weswegen die "Startersets" für Modelleisenbahner, die nur einen Schienenkreis und keine Weiche enthalten, auch so langweilig sind.

[63] Oder könnern ...

[64] John C. Mairhuber, von dem (mir) keine biografischen Daten bekannt sind, hat nach meinem Wissensstand nur zwei Arbeiten geschrieben, nämlich [52] und [53], aber zumindest die erste davon ist ein Klassiker geworden.

[65] Also topologisch äquivalent.

Definition 3.23 *Sei $f \in C(X)$, $X = I$ oder $X = \mathbb{T}$ und sei $\phi \in \Phi$. Eine geordnete Punktmenge*[66] *$x_0 < \cdots < x_n$ heißt* Alternante *der Länge n, zu f und ϕ, wenn*

$$f(x_j) - \phi(x_j) = \sigma_j \|f - \phi\|, \quad j = 0, \ldots, n \qquad \textit{und} \qquad \sigma_{j+1} = -\sigma_j, \quad j = 0, \ldots, n-1, \tag{3.19}$$

wobei $\sigma_j \in \{\pm 1\}$.

Der folgende Satz, der auf Tschebyscheff persönlich zurückgeht, der ihn 1857 für Polynome bewiesen hat[67], gibt uns eine einfachere Beschreibung der Bestaproximation vermittels Alternanten.

Satz 3.24 *(Tschebyscheffscher Alternantensatz)*
Sei Φ ein n–dimensionaler reeller Haar–Raum in $C(X)$. Dann ist $\phi^* \in \Phi$ genau dann eine beste Approximation an $f \in C(X)$ wenn es eine Alternante der Länge n zu f und ϕ^* gibt.

Übung 3.8 Zeigen Sie, daß es auf $\mathbb{T}$ nur Haar–Räume ungerader Dimension geben kann.
Hinweis: Diese Aufgabe steht nicht zufällig an dieser Stelle. ♢

Wir beweisen den Satz, was auch historisch korrekt ist, in zwei Teilen, indem wir zuerst den einfachen Beweis von Tschbyscheff angeben, der zeigt, daß die Existenz einer Alternanten *hinreichend* für eine Bestapproximation ist und dann zeigen, daß für Haar–Räume extremale Signaturen (in $E(f, \phi*)$, um genau zu sein) nichts anderes als Alternanten sind.

Proposition 3.25 *(Tschebyscheff)*
Sei Φ ein n–dimensionaler Haar–Raum[68]*. Existiert zu $\phi^* \in \Phi$ und $f \in C(X)$ eine Alternante der Länge n, dann ist ϕ^* die Bestapproximation an f aus Φ.*

Beweis: Sei $\psi \in \Phi$ die Bestapproximation aus Φ und nehmen wir an, daß $\psi \neq \phi^*$, das heißt $\|f - \psi\| < \|f - \phi^*\|$. Für unsere Alternantenpunkte $x_0, \ldots, x_n$ gilt nun, daß für $j = 0, \ldots, n$

$$\begin{aligned}(\psi - \phi^*)(x_j) &= (f - \phi^*)(x_j) - (f - \psi)(x_j) = \sigma_j \|f - \phi^*\| - (f - \psi)(x_j) \\ &= \sigma_j \underbrace{\left(\|f - \phi^*\| - \sigma_j (f - \psi)(x_j)\right)}_{>0} \quad \begin{cases} > 0, & \sigma_j > 0, \\ < 0, & \sigma_j < 0, \end{cases}\end{aligned}$$

weswegen $\psi - \phi^*$ dasselbe Vorzeichenwechselverhalten haben muß wie $f - \phi^*$. Damit muß aber $\psi - \phi^*$ zwischen den $n + 1$ Alternantenpunkten mindestens n Nullstellen haben – zu viel für ein von Null verschiedenes Element des Haar–Raums Φ. □

Aus "angewandter" Sicht ist Proposition 3.25 eigentlich vollkommen ausreichend, denn sie ermöglicht es uns, zu entscheiden, ob ein gegebenes ϕ^* Bestapproximation ist, indem man

[66] Im Fall $X = \mathbb{T}$ hängt das natürlich davon ab, durch welches Intervall der Länge 2π wir $\mathbb{T}$ darstellen, denn am "runden Tisch" sind ja zuerst einmal alle gleichberechtigt.

[67] Zumindest laut [47].

[68] Wie gesagt: Tschebyscheff hat dies für die algebraischen Polynome vom Grad $\leq n$ gemacht.

ermittelt, wo der maximale Fehler angenommen wird und ob die Vorzeichen sich schön abwechseln. Trotzdem wollen wir natürlich auch Satz 3.24 fertigbeweisen und das erfolgt, indem wir zeigen, daß jede extremale Signatur eine Alternante enthalten muß; da nach dem Satz von Rivlin&Shapiro, Satz 3.11, für jede Bestapproximation ϕ^* von $f \in C(X)$ eine extremale Signatur σ mit $X_\sigma \subseteq E(f, \phi^*)$ existiert, die das Vorzeichenverhalten des Approximationsfehlers beschreibt, existiert somit auch eine Alternante.

Proposition 3.26 *Sei Φ ein n–dimensionaler reeller Haar–Raum*

1. *jede extremale Signatur σ für Φ enthält eine Signatur σ' der Länge $n+1$, die man so anordnen kann, daß*

$$x'_0 < \cdots < x'_n \qquad \textit{und} \qquad \sigma'_{j-1} = -\sigma'_j, \quad j = 1, \ldots, n. \tag{3.20}$$

2. *Ist umgekehrt σ' eine Signatur, die (3.20) erfüllt, dann ist σ' eine extremale Signatur.*

Beweis: Für 1) sei σ eine extremale Signatur, sagen wir, der Länge $N+1$. Wir indizieren X_σ als $x_0 < x_1 < \cdots < x_N$ und bilden nun Blöcke X_j, $j = 0, \ldots, k$, so daß

$$X_j < X_{j+1} \qquad \text{und} \qquad x_\ell, x_m \in X_j \quad \Rightarrow \quad \sigma_\ell = \sigma_m. \tag{3.21}$$

Anschaulich bedeutet (3.21), daß man gerade alle aufeinderfolgenden Punkte mit gleicher Vorzeichenvorgabe in einem Block zusammenfaßt. Wäre die Proposition nun falsch, dann würde k, die Anzahl der Blöcke, $k < n$ erfüllen. Wir wählen nun $k-1$ Punkte $y_0, \ldots, y_{k-1}$ so, daß

$$X_0 < y_0 < X_1 < y_1 < \cdots < X_{k-1} < y_{k-1} < X_k$$

und wählen einen $k+1$–dimensionalen Teilraum Φ_{k+1} von Φ, der *auf dem nichttrivialen kompakten Intervall* $[a, b]$ mit $x_0 < a < y_0$ und $y_k < b < x_N$ ebenfalls ein Haar–Raum ist, siehe Übung 3.9. Die (eindeutige) Lösung ϕ des Interpolationsproblems

$$\phi(a) = \sigma_0, \qquad \phi(y_j) = 0, \quad j = 0, \ldots, k-1,$$

in Φ_{k+1} hat nun gerade die Maximalzahl von k Nullstellen[69] und erfüllt daher, daß $\operatorname{sgn} \phi(x_j) = \sigma_j$, $j = 0, \ldots, N$. Daß bedeutet aber, daß mit den $\mu_0, \ldots, \mu_N$ der extremalen Signatur

$$0 = \sum_{j=0}^{N} \mu_j \, \sigma_j \, \phi(x_j) = \sum_{j=0}^{N} \underbrace{\mu_j}_{>0} \, \underbrace{\sigma_j \, \sigma_j}_{=1} \, \underbrace{|\phi(x_j)|}_{>0} > 0$$

ist, was einen offensichtlichen Widerspruch darstellt.
Für 2) betrachten wir zu einer Basis $\phi_1, \ldots, \phi_n$ von Φ die $n \times n+1$–Matrix

$$A = \begin{bmatrix} \phi_1(x'_0) & \ldots & \phi_1(x'_n) \\ \vdots & \ddots & \vdots \\ \phi_n(x'_0) & \ldots & \phi_n(x'_n) \end{bmatrix},$$

[69] Und die müssen einfach sein, wer's nicht glaubt, soll's selbst beweisen, und zwar durch einmaliges Abdividieren der Nullstelle und Renormalisierung des Resultats, was zu einer weiteren Lösung kleineren Grades führt.

die den maximalen Rang n hat, weil Φ ein Haar–Raum ist. Also gibt es einen eindimensionalen Lösungsraum des Problems $Ay = 0$ und wir können eine Lösung y^* so normieren[70], daß $y_0^* = \sigma_0'$. Setzen wir nun $y_j^* = \mu_j \, \sigma_j$, $j = 0, \ldots, n$, dann ist

$$\sum_{j=0}^{n} \mu_j \, \sigma_j \, \phi_k \left(x_j' \right) = \sum_{j=0}^{n} y_j^* \, \phi_k \left(x_j' \right) = 0, \qquad k = 1, \ldots, n,$$

also ist das dadurch definierte σ eine extremale Signatur der Länge $n+1$ und wegen obiger Überlegung muß $\sigma_j = -\sigma_{j-1}$ sein, wegen $\sigma_0 = \sigma_0'$ und (3.21) also $\sigma_j = \sigma_j'$. Setzt man also $\mu_j' = \mu_j$, dann ist σ' tatsächlich eine extremale Signatur. □

Übung 3.9 Zeigen Sie: Ist Φ ein n–dimensionaler Haar–Raum auf dem kompakten Intervall I, dann gibt es für jedes kompakte Intervall $J \subset I$, eine Folge von Teilräumen

$$\Phi_1 \subset \Phi_2 \subset \cdots \subset \Phi_n = \Phi, \qquad \dim \Phi_j = j, \; j = 1, \ldots, n,$$

die Haar–Räume auf J sind. ♢

Beispiel 3.27 *Man kann den Alternantensatz 3.24 bereits nutzen, um Bestapproximationen zu "raten" und dann die Bestapproximationseigenschaft über die Existenz einer Alternante, das heißt, via Proposition 3.25, nachzuweisen. Der Einfachheit halber interessieren wir uns hier für* lineare *Approximation, das heißt für Approximation mit* Π_1.

1. *Sei* $X = [0, \pi]$ *und* $f(x) = \sin x$. *Die Bestapproximation hier ist offenbar* $\phi(x) = \frac{1}{2}$, *denn es ist*

$$\frac{1}{2} = \left\| \sin(\cdot) - \frac{1}{2} \right\| = -\left(\sin 0 - \frac{1}{2} \right) = \left(\sin \frac{\pi}{2} - \frac{1}{2} \right) = -\left(\sin \pi - \frac{1}{2} \right).$$

2. *Mit* $X = [-\pi, \pi]$ *und* $f(x) = \sin x$ *wird es ein bißchen interessanter. Wir können uns aber das Leben leichtmachen, indem wir zuerst einmal bemerken, daß die Bestapproximation eine* ungerade *Funktion sein muß, siehe Übung 3.10, dann können wir zuerst mal nur auf* $[0, \pi]$ *mit dem eindimensionalen Haar–Raum von solchen* $\phi \in \Pi_1$ *approximieren, die* $\phi(0) = 0$ *erfüllen und diese Bestapproximation ist offensichtlich* $\phi^*(x) = \frac{2}{3\pi} \, x$, *denn für diese Funktion ist*

$$\frac{2}{3} = \| \sin - \phi \|_{[0,\pi]} = (\sin - \phi^*) \left(\frac{\pi}{2} \right) = -(\sin - \phi^*)(\pi).$$

Die symmetrische Fortsetzung auf $[-\pi, \pi]$ *hat dann sogar die vier potentiellen Alternantenpunkte* $\pm\pi, \pm\frac{\pi}{2}$.

Die beiden Approximationen sind in Abb. 3.4 dargestellt.

Aus Punkt 2) von Beispiel 3.27 sehen wir auch, daß sowohl die Punkte $-\pi, -\frac{\pi}{2}, \frac{\pi}{2}$ als auch die Punkte $-\frac{\pi}{2}, \frac{\pi}{2}, \pi$ eine Alternante bilden, was wir auf alle Fälle mal festhalten wollen.

[70] Eventuell müssen wir zuerst einige Nulleinträge von y löschen, aber diese tauchen in der extremalen Signatur ja dann auch nicht auf.

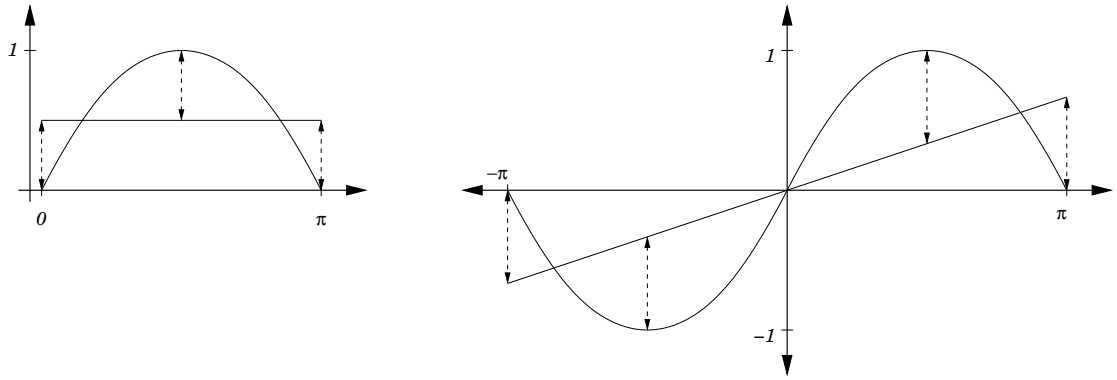


Abbildung 3.4: Die beiden linearen Bestapproximationen an $\sin x$ auf $X = [0, \pi]$ (links) und $X = [-\pi, \pi]$ (rechts). Im ersten Fall ist die Alternante eindeutig, im zweiten nicht.

Korollar 3.28 *Die Alternante aus Satz 3.24 muß nicht eindeutig sein.*

Übung 3.10 Zeigen Sie: Ist X symmetrisch um 0, das heißt, ist $X = -X$ und ist $f \in C(X)$ eine (un)gerade Funktion, das heißt, ist $f(x) = f(-x)$ bzw. $f(x) = -f(-x)$, dann gibt es auch eine (un)gerade Bestapproximation an f. Welche offensichtliche Konsequenz hat das für Haar–Räume? ♢

So, einen haben wir noch an dieser Stelle, eine eigentlich ganz einfache Abschätzung, die uns aber im nächsten Kapitel noch ganz gut weiterhelfen wird. Dieses Resultat geht auf de la Vallée–Poussin[71] zurück.

Satz 3.29 *Sei Φ ein n–dimensionaler Haar–Raum, $f \in C(X)$, $\phi \in \Phi$, und seien $x_0 < \cdots < x_n$ so, daß*

$$\text{sgn } (f - \phi)(x_j) = -\text{sgn } (f - \phi)(x_{j-1}), \qquad j = 1, \ldots, n.$$

Dann ist

$$d(f, \Phi) \geq \min_{j=0,\ldots,n} |(f - \phi)(x_j)|. \tag{3.22}$$

Beweis: Nach Proposition 3.26 gibt es zu der Signatur σ, definiert durch

$$X_\sigma = \{x_0, \ldots, x_n\} \qquad \text{und} \qquad \sigma_j = \text{sgn } (f - \phi)(x_j), \quad j = 0, \ldots, n,$$

Koeffizienten $\mu_j > 0$, $j = 0, \ldots, n$, so daß σ eine *extremale* Signatur ist. Mit derselbe Argumentation wie im Beweis von Satz 3.11 ist nun, unter Verwendung der Bestapproximation

[71] Charles Jean Gustave Nicolas Baron de la Vallée Poussin, 1866–1962, belgischer Mathematiker mit Beiträgen unter anderem zur Approximationstheorie, Potentialtheorie und Zahlentheorie.

$\phi^* \in \Phi$

$$\begin{aligned} d(f,\Phi)\sum_{j=0}^{n}\mu_j &= \|f-\phi^*\|\sum_{j=0}^{n}\mu_j \geq \left|\sum_{j=0}^{n}(f-\phi^*)(x_j)\ \sigma_j\,\mu_j\right| \\ &= \left|\sum_{j=0}^{n}(f-\phi)(x_j)\ \sigma_j\mu_j\right| = \sum_{j=0}^{n}|(f-\phi)(x_j)|\,\mu_j \\ &\geq \min_{j=0,\ldots,n}|(f-\phi)(x_j)|\sum_{j=0}^{n}\mu_j, \end{aligned}$$

woraus (3.22) unmittelbar folgt. □

3.4 Der Remez–Algorithmus

Der Remez[72]–Algorithmus ist ein Verfahren zur *iterativen* Bestimmung der Bestapproximation $\phi^* \in \Phi$ für einen Haar–Raum Φ. Das heißt, eigentlich wird es ein Algorithmus zur Bestimmung einer Alternante! Denn: Ist $A^* = \{x_0^*, \ldots, x_n^*\}$ zufällig eine Alternante für f und die Bestapproximation $\phi^* \in \Phi$, dann ist nach (3.22)

$$d(f,\Phi) = \|f-\phi^*\| = \max_{j=0,\ldots,n}\left|(f-\phi^*)\left(x_j^*\right)\right| = \min_{j=0,\ldots,n}\left|(f-\phi^*)\left(x_j^*\right)\right| \leq d(f,\Phi)$$

Da andererseits aber (3.22) auch für die Menge A^* gilt, ist obendrein

$$d(f,\Phi) \geq d_{A^*}(f,\Phi) \geq \min_{x\in A^*}|(f-\phi^*)(x)| = \max_{x\in A^*}|(f-\phi^*)(x)| = \|f-\phi^*\|_{A^*} = d(f,\Phi),$$

und wir erhalten die folgende Beobachtung.

Bemerkung 3.30 *Die Alternante A^* zur Bestapproximation $\phi^* \in \Phi$ an $f \in C(X)$ zeichnet sich dadurch aus, daß*

$$d_{A^*}(f,\Phi) = d_X(f,\Phi) = d(f,\Phi)$$

ist. Außerdem ist ϕ^* auch diskrete *Bestapproximation an f auf der A^*.*

Übung 3.11 Sei $A \subset X$ eine endliche Teilmenge von $X = I$ oder $X = \mathbb{T}$. Zeigen Sie, daß jede beschränkte Funktion auf A stetig ist. ♢

Damit wird die "Strategie" unseres Verfahrens zur Bestimmung der Bestapproximation aus einem n–dimensionalen Haar–Raum an eine stetige Funktion $f \in C(X)$ schon etwas weniger nebulös:

1. Beginne mit einer beliebigen $n+1$–elementigen Menge $A_0 = \{x_{00}, \ldots, x_{0n}\} \subseteq X$.

[72] Evgeny Yakovlevich Remez, 1896–1975, weißrussisch–ukrainischer (historoisch eher "sovjetischer") Mathematiker, neben dem Algroithmus zur Bestimmung der Bestapproximation beschäftigte er sich auch mit Näherungslösungen von Differentialgleichungen und Mathematikgeschichte.

2. Für $j = 0, 1, 2, \ldots$

   (a) Bestimme die *diskrete* Bestapproximation ϕ_j^* auf A_j an f:

   $$\left\|f - \phi_j^*\right\|_{A_j} = \inf_{\phi\in\Phi} \max_{x\in A_j} \left|f(x) - \phi_j^*(x)\right|.$$

   Nach Satz 3.24[73] ist dann A_j eine Alternante für f und ϕ_j^* und mit Satz 3.29 folgt, daß

   $$\left\|f - \phi_j^*\right\|_{A_j} = \max_{x\in A_j} \left|f(x) - \phi_j^*(x)\right| = \min_{x\in A_j} \left|f(x) - \phi_j^*(x)\right| \leq d\left(f, \Phi\right).$$

   (b) Ist $\left\|f - \phi_j^*\right\|_{A_j} = \left\|f - \phi_j^*\right\|$, dann ist A_j eine Alternante zur Bestapproximation, denn

   $$d\left(f, \Phi\right) \leq \left\|f - \phi_j^*\right\| = \left\|f - \phi_j^*\right\|_{A_j} \leq d\left(f, \Phi\right)$$

   liefert, daß ϕ_j^* die gesuchte Bestapproximation sein muß.

   (c) Ist $\left\|f - \phi_j^*\right\|_{A_j} < \left\|f - \phi_j^*\right\|$, dann bestimmen wir *durch Ersetzen eines Punktes in* A_j eine neue Menge A_{j+1}, so daß

   $$\text{sgn}\left(f - \phi_j^*\right)\left(x_{j+1,k}\right) = -\text{sgn}\left(f - \phi_j^*\right)\left(x_{j+1,k-1}\right), \qquad k = 1, \ldots, n, \tag{3.23}$$

   und

   $$\min_{x\in A_{j+1}} \left|f(x) - \phi_j^*(x)\right| = \left\|f - \phi_j^*\right\|_{A_j} \qquad \text{und} \qquad \left\|f - \phi_j^*\right\|_{A_{j+1}} = \left\|f - \phi_j^*\right\| \tag{3.24}$$

Es bleiben allerdings ein paar "kleine Detailfragen" zu klären:

1. Wie bestimmt man die diskrete Bestapproximation an $n + 1$ Punkten[74]?
2. Wie bestimmen wir das neue A_{j+1} konkret.
3. Warum funktioniert das Ganze eigentlich?

Beginnen wir mit dem letzten der drei Punkte, 3), denn diese Eigenschaft läßt sich leicht mit dem folgenden Resultat erklären.

Lemma 3.31 *Sei zu $f \in C(X)$ und $\phi \in \Phi$ eine $n+1$–elementige Menge $A = \{x_0, \ldots, x_n\}$ gegeben, so daß*

$$\text{sgn}\left(f - \phi\right)\left(x_j\right) = -\text{sgn}\left(f - \phi\right)\left(x_{j-1}\right), \qquad j = 1, \ldots, n \tag{3.25}$$

und sei

$$\delta = \min_{x\in A\setminus\{x_j\}} \left|f(x) - \phi(x)\right| = \max_{x\in A\setminus\{x_j\}} \left|f(x) - \phi(x)\right| < \left|f\left(x_j\right) - \phi\left(x_j\right)\right| \tag{3.26}$$

für ein $\delta > 0$ und $j \in \{0, \ldots, n\}$. Dann ist $d_A\left(f, \Phi\right) > \delta$.

[73] Mit A_j anstelle von X!

[74] Scherzfrage: Was ist die diskrete Bestapproximation an n Punkten? Und was ist der Approximationsfehler? Antwort: Interpolation!

Dieses Lemma[75] sagt uns also, daß die A_j aus dem obigen Algorithmusentwurf die Eigenschaft

$$d_{A_0}(f,\Phi) < d_{A_1}(f,\Phi) < \cdots \leq d(f,\Phi)$$

haben und daß deswegen die Folge der Abstände konvergieren muß. Da außerdem die Alternanten A_j als Elemente von X^{n+1} aufgefaßt werden können, konvergiert zumindest eine Teilfolge gegen eine "Alternante" A^* mit der Eigenschaft

$$d_{A_j}(f,\Phi) < d_{A^*}(f,\Phi) \leq d(f,\Phi),$$

und wäre dieses A^* nicht die Alternante zur Bestapproximation, dann könnten wir es mit einem weiteren Schritt des Algorithmus nach Lemma 3.31 nochmals verbessern. Die Idee dieses Verfahrens ist also durchaus schon mal sinnvoll ...

Beweis von Lemma 3.31: Wegen (3.25) und (3.26) liefert uns Satz 3.29 sofort, daß $d_A(f,\Phi) \geq \delta$. Sei nun ϕ^* die Bestapproximation an f auf A, dann ist A eine Alternante für f und ϕ^*. Wäre jetzt $d_A(f,\Phi) = \delta$, dann hat die Funktion $\phi - \phi^*$ mindestens n Nullstellen, nämlich $A \setminus \{x_j\}$, müßte also die Nullfunktion sein, was aber im Widerspruch zur Annahme steht, daß ϕ^* Bestapproximation ist – schließlich ist ja ϕ keine Bestapproximation, weil A keine Alternante für f und ϕ ist; der Absolutbetrag von $f - \phi$ an der Stelle x_j ist einfach zu groß. □

Der Übergang von einer "Alternate" zur nächsten, also unser Punkt 2) erfolgt somit nach dem folgenden Verfahren.

Algorithmus 3.32 *(Remez–Algorithmus, Austauschschritt)*
Gegeben: *$f \in C(X)$, Alternante $A \subset X$, $\#A = n+1$, ϕ Bestapproximation auf A.*

1. *Bestimme $x \in X$, so daß*
$$|f(x) - \phi(x)| = \|f - \phi\|.$$
2. *Ist $x \in A$, dann ist ϕ Bestapproximation an f auf X.*
3. *Andernfalls unterscheide drei Fälle:*
   (a) *$x < x_0$. Setze*
   $$A' = \begin{cases} \{x, x_1, \ldots, x_n\}, & \operatorname{sgn}(f-\phi)(x) = \operatorname{sgn}(f-\phi)(x_0) \\ \{x, x_0, x_1, \ldots, x_{n-1}\}, & \operatorname{sgn}(f-\phi)(x) = -\operatorname{sgn}(f-\phi)(x_0). \end{cases}$$
   (b) *$x_k < x < x_{k+1}$. Setze*
   $$A' = A \setminus \{x_k, x_{k+1}\} \cup \begin{cases} \{x, x_{k+1}\}, & \operatorname{sgn}(f-\phi)(x) = \operatorname{sgn}(f-\phi)(x_k), \\ \{x, x_k\}, & \operatorname{sgn}(f-\phi)(x) = \operatorname{sgn}(f-\phi)(x_{k+1}). \end{cases}$$
   (c) *$x > x_n$. Setze*
   $$A' = \begin{cases} \{x_0, \ldots, x_{n-1}, x\}, & \operatorname{sgn}(f-\phi)(x) = \operatorname{sgn}(f-\phi)(x_n) \\ \{x_1, x_2, \ldots, x_n, x\}, & \operatorname{sgn}(f-\phi)(x) = -\operatorname{sgn}(f-\phi)(x_n). \end{cases}$$

[75] Wir werden es gleich beweisen, aber zuerst sollten wir wieder mal "die Geschichte fertigerzählen".

Ergebnis: $A' \subset X$, $\#A' = n+1$, *mit*

$$\operatorname{sgn}\,(f-\phi)\left(x'_j\right) = -\operatorname{sgn}\,(f-\phi)\left(x'_{j-1}\right), \qquad j = 1,\ldots,n,$$

und es gibt $\delta > 0$ *und* $x \in A'$*, so daß*

$$|(f-\phi)(x)| = \|f-\phi\| \qquad \textit{und} \qquad |(f-\phi)(x')| = \delta \le \|f-\phi\|, \quad x' \in A \setminus \{x\}.$$

Dieser Austauschschritt ersetzt also, wenn möglich, genau einen Punkt in einer (diskreten) Alternante und zwar so, daß an diesem Punkt ein echt größerer Wert angenommen wird. Die Norm ist nicht unbedingt notwendig und sorgt lediglich dafür, daß die erreichte "Vergrößerung" so groß ist wie möglich. Ein "normaler" Austauschschritt ist in Abb 3.5 dargestellt.

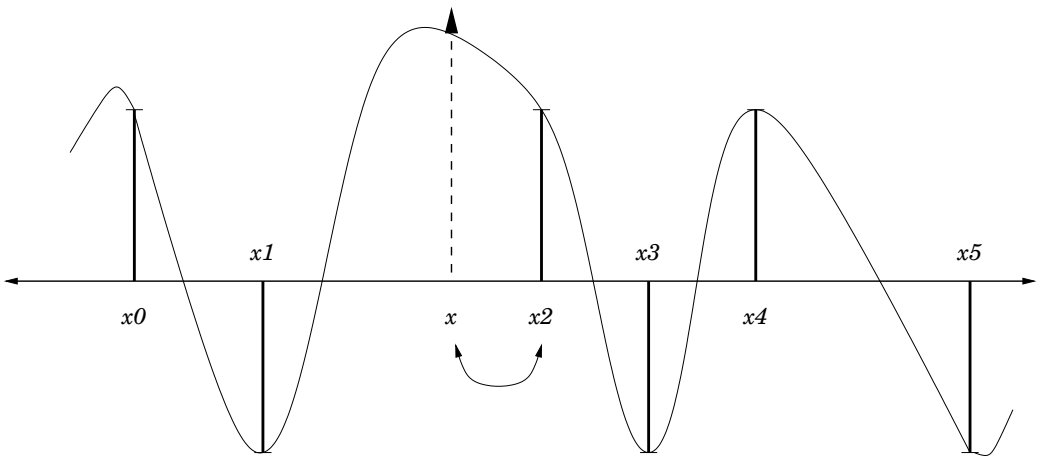


Abbildung 3.5: Ein Austauschschritt des Remez–Algorithmus, der benachbarte Punkt der diskreten Alternante mit *demselben* Vorzeichen wird durch die Abszisse des Maximums ersetzt.

Übung 3.12 Zeigen Sie: Ist $\Phi \subset C(X)$ ein n–dimensionaler Haar–Raum, dann ist Φ auch ein Haar–Raum auf A für jedes $A \subset X$, $\#A \ge n+1$, und umgekehrt. ♢

Bemerkung 3.33 *Ein Problem bleibt allerdings ungelöst in Algorithmus 3.32, ganz einfach deswegen, weil es für* beliebige *stetige Funktionen halt einfach nicht zu lösen ist, nämlich die Bestimmung des Maximum. Der naive Ansatz, die Funktion $f-\phi$ einfach "fein genug" abzutasten, scheitert daran, daß man über eine stetige Funktion kaum Aussagen machen kann, wenn man sie nur an endlich vielen Werten kennt.*
Anders sieht die Sache schon aus, wenn man beispielsweise weiß, daß die Funktion f und alle Funktionen aus Φ "kontrolliert" stetig sind[76].

[76] Beispielsweise Lipschitz–stetig.

Was also noch bleibt, ist die Bestimmung der *diskreten* Bestapproximation auf einer $n+1$–elementigen Menge, mit anderen Worten:

Zu $f \in C(X)$ und $A \subset X$, $\#A = n+1$, bestimme man $\phi_A \in \Phi$, so daß

$$\|f - \phi_A\|_A = d_A(f, \Phi) = \min_{\phi \in \Phi} \max_{x \in A} |(f - \phi)(x)|.$$

Zur Lösung verwenden wir einen Ansatz aus [60], das Resultat findet sich aber auch schon in [75]. Dazu schreiben wir Zuerst einmal $A = \{x_0, \ldots, x_n\}$ und setzen $\delta_j := (f - \phi)(x_j)$, $j = 0, \ldots, n$, denn dann besteht unser Minimierungsproblem darin, ϕ so zu wählen, daß $\max_{j=0,\ldots,n} |\delta_j|$ *minimiert* wird. Das aber ist ein Minimierungsproblem mit einer lineren Nebenbedingung, denn da[77] $\phi(x_j) = f(x_j) - \delta_j$, $j = 0, \ldots, n$, ist

$$\begin{aligned} 0 &= \det \begin{bmatrix} \phi(x_0) & \phi_1(x_0) & \ldots & \phi_n(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(x_n) & \phi_1(x_n) & \ldots & \phi_n(x_n) \end{bmatrix} = \det \begin{bmatrix} f(x_0) - \delta_0 & \phi_1(x_0) & \ldots & \phi_n(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_n) - \delta_n & \phi_1(x_n) & \ldots & \phi_n(x_n) \end{bmatrix} \\ &= \det \begin{bmatrix} f(x_0) & \phi_1(x_0) & \ldots & \phi_n(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_n) & \phi_1(x_n) & \ldots & \phi_n(x_n) \end{bmatrix} - \det \begin{bmatrix} \delta_0 & \phi_1(x_0) & \ldots & \phi_n(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \delta_n & \phi_1(x_n) & \ldots & \phi_n(x_n) \end{bmatrix}, \end{aligned} \tag{3.27}$$

und eine Leibnitz–Entwicklung der Matrix ganz rechts liefert Koeffizienten c_j, $j = 0, \ldots, n$, so daß

$$\sum_{j=0}^{n} c_j \, \delta_j = \det \begin{bmatrix} f(x_0) & \phi_1(x_0) & \ldots & \phi_n(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_n) & \phi_1(x_n) & \ldots & \phi_n(x_n) \end{bmatrix}.$$

Das sieht schon viel besser aus, denn nun haben wir das folgende Resultat aus [60], das uns etwas über die δ_j sagt.

Lemma 3.34 *Zu gegebenen $0 \neq a \in \mathbb{R}^n$ und $b \in \mathbb{R}$ ist*

$$\min_{a^T y = b} \max_{j=1,\ldots,n} |y_j| = \frac{|b|}{|a_1| + \cdots + |a_n|} =: \rho = \max_{j=1,\ldots,n} \left|y_j^*\right|, \tag{3.28}$$

wobei

$$y_j^* = \rho \operatorname{sgn}(a_j \, b), \qquad j = 1, \ldots, n. \tag{3.29}$$

Dieses Lemma[78] sagt uns also, daß $\delta_j = \sigma_j \, \rho$ für $j = 0, \ldots, n$, weil aber A für das optimale ϕ_A eine Alternante sein muß, ist demnach $\rho = \|f - \phi_A\|_A$. Setzen wir das in (3.27) ein, dann erhalten wir, daß

$$\det \begin{bmatrix} f(x_0) & \phi_1(x_0) & \ldots & \phi_n(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_n) & \phi_1(x_n) & \ldots & \phi_n(x_n) \end{bmatrix} = \rho \det \begin{bmatrix} \sigma_0 & \phi_1(x_0) & \ldots & \phi_n(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_n & \phi_1(x_n) & \ldots & \phi_n(x_n) \end{bmatrix},$$

[77] Tirivialerweise!

[78] Daß wir es erst später beweisen werden, das wird inzwischen wohl niemanden mehr überraschen ...

also

$$\rho = \left| \frac{\det \begin{bmatrix} f(x_0) & \phi_1(x_0) & \dots & \phi_n(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ f(x_n) & \phi_1(x_n) & \dots & \phi_n(x_n) \end{bmatrix}}{\det \begin{bmatrix} 1 & \phi_1(x_0) & \dots & \phi_n(x_0) \\ -1 & \phi_1(x_1) & \dots & \phi_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^n & \phi_1(x_n) & \dots & \phi_n(x_n) \end{bmatrix}} \right|, \tag{3.30}$$

und unser ϕ erhalten wir als Lösung des Interpolationsproblems

$$\phi(x_j) = \sigma(-1)^j \rho + f(x_j), \qquad j = 0, \dots, n,\ \sigma \in \{-1, 1\}, \tag{3.31}$$

wobei die freien Variablen die Koeffizienten von ϕ bezüglich einer Basis $\phi_1, \dots, \phi_n$ von Φ und das "Vorzeichen" σ sind. Anders gesagt, ist $[\sigma, a_1, \dots, a_n]^T$ die Lösung des Gleichungssystems

$$\begin{bmatrix} \rho & \phi_1(x_0) & \dots & \phi_n(x_0) \\ -\rho & \phi_1(x_1) & \dots & \phi_n(x_1) \\ \vdots & \vdots & \ddots & \vdots \\ (-1)^n \rho & \phi_1(x_n) & \dots & \phi_n(x_n) \end{bmatrix} \begin{bmatrix} \sigma \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f(x_0) \\ \vdots \\ f(x_n) \end{bmatrix}, \tag{3.32}$$

dann ist $\phi = a_1 \phi_1 + \dots + a_n \phi_n$.

Beweis von Lemma 3.34: Zuerst bemerken wir, daß mit (3.28) und (3.29)

$$a^T y^* = \sum_{j=1}^n a_j \rho \underbrace{\operatorname{sgn}(a_j b)}_{=\operatorname{sgn} a_j \operatorname{sgn} b} = \rho \operatorname{sgn} b \sum_{j=1}^n |a_j| = |b| \operatorname{sgn} b = b,$$

also ist y^* wirklich eine zulässige Lösung. Für jedes $y \in \mathbb{R}^n$, das $a^T y = b$ und $\|y\|_\infty \le \rho$ erfüllt, ist dann

$$y_j = \gamma_j y_j^* = \gamma_j \rho (\operatorname{sgn} a_j)(\operatorname{sgn} b), \qquad \gamma_j \in [-1, 1], \qquad j = 1, \dots, n,$$

und wegen

$$\begin{aligned} |b| &= (\operatorname{sgn} b)\, b = (\operatorname{sgn} b) \sum_{j=1}^n a_j y_j = (\operatorname{sgn} b) \sum_{j=1}^n a_j \gamma_j \rho (\operatorname{sgn} a_j)(\operatorname{sgn} b) \\ &= \rho \sum_{j=1}^n |a_j| \gamma_j = \frac{|b|}{|a_1| + \dots + |a_n|} \sum_{j=1}^n |a_j| \gamma_j, \end{aligned}$$

also

$$\sum_{j=1}^n |a_j| = \sum_{j=1}^n |a_j| \gamma_j \qquad \Longrightarrow \qquad \gamma_j = 1, \quad j = 1, \dots, n,$$

muß y^* sogar die *eindeutige* Minimallösung sein, wenn $a_j \neq 0$, $j = 1, \ldots, n$. □

Nun können wir, (3.32) sei Dank, also unseren zweiten Baustein für den Algorithmus zur Berechnung der Bestapproximation angeben, nämlich die Bestimmung der diskreten Bestapproximation, was letztlich auch Punkt 1) erledigt. Um alles etwas knapper formulieren zu können bezeichnen wir mit $\Phi(A)$ die (erweiterete) $n+1 \times n$–Matrix des Interpolationsproblems, mit $f(A)$ den $n+1$–Vektor der rechten Seite und mit $\sigma = [-1, 1, -1, \ldots, (-1)^{n+1}]^T$ den $n+1$–Vektor der mit wechselnden Vorzeichen.

Algorithmus 3.35 *(Remez–Algorithmus, diskrete Bestapproximation)*
Gegeben: $f \in C(X)$, $A \subset X$, $\#A = n+1$.

1. *Berechne*
$$\rho = \frac{\det[f(A), \Phi(A)]}{\det[\sigma, \Phi(A)]}$$
2. *Berechne den Vektor* $a = [a_0, \ldots, a_n]^T \in \mathbb{R}^{n+1}$ *als Lösung des linearen Gleichungssystems*
$$[\rho\,\sigma, \Phi(A)]\, a = f(A).$$

Ergebnis: *Diskrete Bestapproximation auf* A*:*

$$\phi = \sum_{j=1}^{n} a_j\, \phi_j \qquad \in \Phi.$$

Und damit können wir schließlich unseren Remez–Algorithmus zusammenbauen.

Algorithmus 3.36 *(Remez–Algorithmus)*
Gegeben: $f \in C(X)$, n*–dimensionaler Haar–Raum* $\Phi \subset C(X)$.

1. *Wähle* $A_0 \subset X$, $\#A_0 = n+1$, *beliebig.*
2. *Für* $j = 0, 1, 2, \ldots$
   (a) *Bestimme die diskrete Bestapproximation* ϕ_j^* *mit*
   $$\left\|f - \phi_j^*\right\|_{A_j} = d_{A_j}(f, \Phi)$$
   über Algorithmus 3.35.
   (b) *Bestimme* A_{j+1} *aus* A_j *und* ϕ_j^* *über den Austauschschritt aus Algorithmus 3.32.*
3. *Abbruchbedingung: Für Toleranz* u *ist*
$$\frac{\left\|f - \phi_j^*\right\|_{A_j}}{\left\|f - \phi_j^*\right\|_X} \geq 1 - u.$$

Ergebnis: *(näherungsweise) Bestapproximation ϕ^*_∞ und zugehörige Alternante A_∞.*

Beispiel 3.37 *Sehen wir uns doch einmal die Funktionsweise des Remez–Algorithmus an Punkt 2 aus Beispiel 3.27 an, nämlich die Approximation von $f(x) = \sin x$ durch $\Phi = \operatorname{span}\{1, x\}$ auf $X = [-\pi, \pi]$.*

1. *Da $\dim \Phi = 2$ beginnen wir doch einfach mal mit $2 + 1 = 3$* gleichverteilten *Punkten in X und zwar*[79] *$A_0 = \{-\pi, 0, \pi\}$. Die (symmetrische) Bestapproximation an diesen drei Punkten ist $\phi_0 = 0$ – die interpoliert sogar! Der Fehler $|\sin x - \phi_0(x)| = |\sin x|$ nimmt nun also sein Maximum an den beiden Stellen $\pm\frac{\pi}{2}$ an, von denen wir uns eine aussuchen können.*

2. *Wählen wir also $-\frac{\pi}{2}$, dann können wir wahlweise den Extremalpunkt $-\pi$ oder 0 durch $-\frac{\pi}{2}$ ersetzen, denn die "Alternante" hatte ja die besondere Eigenschaft, daß die diskrete Bestapproximation* interpolierte*, also alle $\sigma_j = 0$ waren.*

3. *Zur $A_1 = \left\{-\pi, -\frac{\pi}{2}, \pi\right\}$ ist nun die diskrete Bestapproximierende die Funktion $\phi_1 = -\frac{1}{2}$, denn schließlich ist*

$$\frac{1}{2} = \underbrace{\sin(-\pi)}_{=0} + \frac{1}{2} = -\left(\sin -\frac{\pi}{2} + \frac{1}{2}\right) = \sin \pi + \frac{1}{2}.$$

   Der maximale Fehler, nämlich $\frac{3}{2}$, wird nun an der Stelle $\frac{\pi}{2}$ angenommen, die wir für π eintauschen können.

4. *Mit $A_2 = \left\{-\pi, -\frac{\pi}{2}, \frac{\pi}{2}\right\}$ sind wir nun aber fertig, denn das ist ja, wie wir aus Beispiel 3.27 wissen, bereits eine Alternante.*

Der Remez–Algorithmus muss also nicht unbedingt konvergieren, er kann auch nach endlich vielen Schritten terminieren[80].

[79] Es ist nicht gerade unüblich, daß Extrema am Rand angenommen werden, weswegen nie eine schlechte Idee ist, die Randpunkte in die "Anfangsalternante" einzubeziehen.

[80] Was natürlich auch eine Form von Konvergenz ist, sogar eine besonders schnelle.

Bisweilen erweist sich das wahre Wissen als bedeutungslos, und dann kann man es auch erfinden.

Javier Marias, *Alle Seelen*

4 Mehr über Bernsteinpolynome

In diesem Kapitel wollen wir einige weitere Aspekte der Approximationstheorie kennenlernen, die man sehr schön am Beispiel der (univariaten) Bernsteinpolynome darstellen kann, nämlich

- Simultanapproximation
- "Shape preserving approximation",
- Saturation,

Eigenschaften, die die polynomiale Bestapproximation *nicht* hat, in manchen Fällen leider, in anderen eher glücklicherweise.

4.1 Ableitungen von Bernsteinpolynomen

Wir beginnen das Ganze mal ziemlich unschuldig, indem wir uns die Ableitungen des n*–ten Bernsteinpolynoms*

$$B_n f(x) = \sum_{j=0}^{n} f\left(\frac{j}{n}\right) B_j^n(x), \qquad B_j(x) = \binom{n}{j} x^j (1-x)^{n-j}, \tag{4.1}$$

zu einer Funktion $f \in C(I)$, $I = [0,1]$, ansehen. Dazu leitet man zweckmäßigerweise die *Basispolynome* B_j^n ab und erhält, daß für $j = 0, \ldots, n$

$$\begin{aligned}
\frac{d}{dx} B_j^n(x) &= \frac{n!}{j!(n-j)!} \left(j\, x^{j-1} (1-x)^{n-j} - (n-j)\, x^j (1-x)^{n-j-1} \right) \\
&= n \left(\frac{n-1}{(j-1)!(n-j)!} x^{j-1} (1-x)^{n-j} + \frac{n-1}{j!(n-j-1)!} x^j (1-x)^{n-j-1} \right) \\
&= n \left(B_{j-1}^{n-1}(x) - B_j^{n-1}(x) \right),
\end{aligned}$$

wobei $B_{-1}^{n-1} = B_n^{n-1} = 0$ ist. Also ist

$$\frac{d}{dx} B_n f(x) = \sum_{j=0}^{n} f\left(\frac{j}{n}\right) \frac{d}{dx} B_j^n(x) = \sum_{j=0}^{n} f\left(\frac{j}{n}\right) n \left(B_{j-1}^{n-1}(x) - B_j^{n-1}(x) \right)$$

$$
\begin{aligned}
&= n \sum_{j=1}^{n} f\left(\frac{j}{n}\right) B_{j-1}^{n-1}(x) - n \sum_{j=0}^{n-1} f\left(\frac{j}{n}\right) B_j^{n-1}(x) \\
&= n \sum_{j=0}^{n-1} \left(f\left(\frac{j+1}{n}\right) - f\left(\frac{j}{n}\right) \right) B_j^{n-1}(x).
\end{aligned}
$$

Definition 4.1 *Für $h > 0$ ist die* Vorwärtsdifferenz Δ_h *definiert als*

$$\Delta_h f(x) = f(x+h) - f(x), \qquad f \in C(I), \quad x, x+h \in I,$$

und für $k \geq 1$ die Iteration

$$\Delta_h^k f(x) = \Delta_h \, \Delta_h^{k-1} f(x) = \underbrace{\Delta_h \cdots \Delta_h}_{k} f(x), \qquad x, x+kh \in I.$$

Übung 4.1 Zeigen Sie: Für $k \geq 1$ ist

$$\Delta_h^k f(x) = \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} f(x+jh).$$

♢

Mit Definition 4.1 hat die Ableitungsformel von oben also die Form

$$\frac{d}{dx} B_n f(x) = n \sum_{j=0}^{n-1} \Delta_{1/n} f\left(\frac{j}{n}\right) B_j^{n-1}(x), \tag{4.2}$$

und wir erhalten praktisch unmittelbar das folgende Resultat.

Satz 4.2 *Für $f \in C(I)$, $n \in \mathbb{N}_0$ und $0 \leq k \leq n$ ist*[81]

$$\frac{d^k}{dx^k} B_n f = \frac{n!}{(n-k)!} \sum_{j=0}^{n-k} \Delta_{1/n}^k f\left(\frac{j}{n}\right) B_j^{n-k}. \tag{4.3}$$

Beweis: Für $k = 0$ ist (4.3) trivial und für $k = 1$ nichts anderes als (4.2). Allgemein ist

$$
\begin{aligned}
\frac{d^{k+1}}{dx^{k+1}} B_n f(x) &= \frac{d}{dx} \frac{d^k}{dx^k} B_n f(x) = \frac{n!}{(n-k)!} \sum_{j=0}^{n-k} \Delta_{1/n}^k f\left(\frac{j}{n}\right) \frac{d}{dx} B_j^{n-k} \\
&= \frac{n!}{(n-k)!} \sum_{j=0}^{n-k} \Delta_{1/n}^k f\left(\frac{j}{n}\right) (n-k) \left(B_{j-1}^{n-k-1}(x) - B_j^{n-k-1}(x) \right) \\
&= \frac{n!}{(n-k-1)!} \sum_{j=0}^{n-k-1} \left(\Delta_{1/n}^k f\left(\frac{j+1}{n}\right) - \Delta_{1/n}^k f\left(\frac{j}{n}\right) \right) B_j^{n-k-1}(x) \\
&= \frac{n!}{(n-k-1)!} \sum_{j=0}^{n-k} \Delta_{1/n}^{k+1} f\left(\frac{j}{n}\right) B_j^{n-k-1}(x),
\end{aligned}
$$

[81] Die letzte Einschränkung, $k \leq n$, ist mehr "für die Galerie": Ist nämlich $k > n$, dann ist die Ableitung sowieso Null, wir haben es bei B_n ja mit einem Polynom der Ordnung n zu tun.

was per Induktion die Behauptung ergibt. □

4.2 Simultanapproximation

Die Ableitungsformel aus Satz 4.2 erlaubt es uns, eine interessante Eigenschaft der Bernsteinpolynome festzuhalten: Sie approximieren nämlich nicht nur die Funktion sondern gleichzeitig auch eventuell vorhandene Ableitungen der Funktion. Genauer gilt der folgende Satz.

Satz 4.3 *Für $f \in C^k(I)$, $k \geq 0$, ist*

$$0 = \lim_{n\to\infty} \max_{0\leq j\leq k} \left\| \frac{d^j}{dx^j} (f - B_n f) \right\| =: \lim_{n\to\infty} \|f - B_n f\|_k . \tag{4.4}$$

Übung 4.2 Zeigen Sie, daß $C^k(I)$ bezüglich der Norm

$$\|f\|_k := \max_{0\leq j\leq k} \left\|f^{(j)}\right\|$$

ein Banachraum ist. ♢

Das entscheidende Hilfsmittel zum Beweis von Satz 4.3 ist eine Darstellung der Vorwärtsdifferenz über die entsprechende Ableitung.

Lemma 4.4 *Für $k \geq 1$, $f \in C^k(I)$ und $h > 0$ ist*

$$\Delta_h^k f(x) = \int_{[0,h]^k} f^{(k)} \left(x + \sum_{j=1}^k t_j \right) \, dt_1 \cdots dt_k. \tag{4.5}$$

Beweis: Für $k = 1$ ist

$$\Delta_h f(x) = f\,(x + h) - f(x) = \int_0^h f'(t)\, dt,$$

und generell

$$\begin{aligned}
\Delta_h^{k+1} f(x) &= \Delta_h\, \Delta_h^k f(x) = \Delta_h \int_{[0,h]^k} f^{(k)} \left(x + \sum_{j=1}^k t_j \right) \, dt_1 \cdots dt_k \\
&= \int_{[0,h]^k} f^{(k)} \left(x + h + \sum_{j=1}^k t_j \right) - f^{(k)} \left(x + \sum_{j=1}^k t_j \right) \, dt_1 \cdots dt_k \\
&= \int_{[0,h]^k} \int_0^h f^{(k+1)} \left(x + \sum_{j=1}^k t_j + t \right) \, dt_1 \cdots dt_k \, dt \\
&= \int_{[0,h]^{k+1}} f^{(k+1)} \left(x + \sum_{j=1}^{k+1} t_j \right) \, dt_1 \cdots dt_{k+1},
\end{aligned}$$

was per Induktion den Beweis komplettiert. □

Der folgende Begriff wird uns später noch mehr "quälen", im Moment benötigen wir ihn aber eher, um die Sachen einfach und knapp formulieren zu können.

Definition 4.5 *Zu $f \in C(I)$ und $\delta > 0$ ist der* Stetigkeitsmodul $\omega(f, \delta)$ *definiert als*

$$\omega(f, \delta) = \sup_{0<h<\delta} \sup_{x, x+h \in I} |f(x+h) - f(x)|. \tag{4.6}$$

Übung 4.3 Zeigen Sie: Für jedes $f \in C(I)$

1. ist die Funktion $\omega(f, \cdot)$ monoton steigend,
2. ist
$$\lim_{\delta \to 0} \omega(f, \delta) = 0.$$
3. ist
$$\lim_{\delta \to 0} \delta^{-1} \omega(f, \delta) = 0 \qquad \Longleftrightarrow \qquad f \equiv \text{const.}$$

♢

Bemerkung 4.6 *Der Stetigkeitsmodul, genauer, dessen "Abfall" für $\delta \to 0$ misst, wie stetig eine Funktion ist. Ist beispielsweise f Lipschitz–stetig, dann ist $\omega(f, \delta) \leq M\,\delta$ für eine Konstante $M > 0$, fällt also relativ schnell ab. Es kann aber auch Funktionen mit beliebig "schlechter" Stetigkeit geben, siehe Übung 4.4.*

Übung 4.4 Sei $a_1 \geq a_2 \geq \cdots > 0$ eine fallende positive Nullfolge[82]. Zeigen Sie: Es gibt eine Funktion $f \in C(I)$, so daß $\omega\left(f, \frac{1}{n}\right) \geq a_n$, $n \in \mathbb{N}$.
Hinweis: Man muß f eigentlich nur an "relativ wenigen" Punkten wirklich definieren. ♢

Jetzt wird's mal einen Moment lang wirklich technisch, aber wir werden gleich sehen, wofür diese Abschätzung gut ist.

Lemma 4.7 *Für $f \in C^k(I)$ und $n \in \mathbb{N}_0$ ist*

$$\left| n^k \Delta^k_{1/n} f\left(\frac{j}{n}\right) - f^{(k)}\left(\frac{j}{n-k}\right) \right| \leq \omega\left(f^{(k)}, \frac{k}{n}\right), \qquad j = 0, \ldots, n-k. \tag{4.7}$$

Beweis: Zuerst mal sehen wir, daß

$$\frac{j}{n-k} - \frac{j}{n} = \frac{nj - nj + kj}{n(n-k)} = \frac{k}{n} \underbrace{\frac{j}{n-k}}_{\leq 1} \leq \frac{k}{n}$$

und

$$\frac{j+k}{n} - \frac{j}{n-k} = \frac{nj + (n-k-j)k - nj}{n(n-k)} = \frac{k}{n} \frac{n-k-j}{n-k} \leq \frac{k}{n},$$

[82] Das heißt, daß $a_n > 0$ und $a_n \to 0$ für $n \to \infty$.

also ist für $j = 0, \dots, n-k$

$$\left| x - \frac{j}{n-k} \right| \le \frac{k}{n}, \qquad x \in \left[\frac{j}{n}, \frac{j+k}{n} \right]. \tag{4.8}$$

Somit erhalten wir unter Verwendung von Lemma 4.4

$$\begin{aligned}
&\left| n^k \Delta_{1/n}^k f\left(\frac{j}{n}\right) - f^{(k)}\left(\frac{j}{n-k}\right) \right| \\
&= n^k \left| \Delta_{1/n}^k f\left(\frac{j}{n}\right) - n^{-k} f^{(k)}\left(\frac{j}{n-k}\right) \right| \\
&= n^k \left| \int_{\left[0,\frac{1}{n}\right]^k} f^k \Bigg(\underbrace{\frac{j}{n} + \sum_{j=1}^k t_k}_{\in\left[\frac{j}{n}, \frac{j+k}{n}\right]} \Bigg) - f^{(k)}\left(\frac{j}{n-k}\right) \, dt_1 \cdots dt_k \right| \\
&\le n^k \int_{\left[0,\frac{1}{n}\right]^k} \underbrace{\left| f^k \Bigg(\frac{j}{n} + \sum_{j=1}^k t_k \Bigg) - f^{(k)}\left(\frac{j}{n-k}\right) \right|}_{\le \omega\left(f^{(k)}, \frac{k}{n}\right)} \, dt_1 \cdots dt_k \\
&\le \omega\left(f^{(k)}, \frac{k}{n}\right) \, n^k \underbrace{\int_{\left[0,\frac{1}{n}\right]^k} dt_1 \cdots dt_k}_{= n^{-k}} = \omega\left(f^{(k)}, \frac{k}{n}\right).
\end{aligned}$$

□

Beweis von Satz 4.3: Für $\ell = 0, \dots, k$ ist nach Satz 4.2 und Lemma 4.7

$$\begin{aligned}
\left\| B_n^{(\ell)} f - B_{n-\ell} f^{(\ell)} \right\| &= \left\| \frac{n!}{(n-\ell)!} \sum_{j=0}^{n-\ell} \Delta_{1/n}^\ell f\left(\frac{j}{n}\right) B_j^{n-\ell} - \sum_{j=0}^{n-\ell} f^{(\ell)}\left(\frac{j}{n-\ell}\right) B_j^{n-\ell} \right\| \\
&\le \left(n^\ell - \frac{n!}{(n-\ell)!} \right) \left\| \sum_{j=0}^{n-\ell} \Delta_{1/n}^\ell f\left(\frac{j}{n}\right) B_j^{n-\ell} \right\| \\
&\quad + \left\| \sum_{j=0}^{n-k} \left| n^\ell \, \Delta_{1/n}^\ell f\left(\frac{j}{n}\right) - f^{(\ell)}\left(\frac{j}{n-\ell}\right) \right| \, B_j^{n-\ell} \right\| \\
&\le \left(1 - \prod_{j=1}^{\ell-1} \frac{n-j}{n} \right) \left\| f^{(\ell)} \right\| + \omega\left(f^{(\ell)}, \frac{\ell}{n}\right),
\end{aligned}$$

siehe auch Übung 4.5, also ist

$$\begin{aligned}
\left\| B_n^{(\ell)} f - f^{(\ell)} \right\| &\le \left\| B_{n-\ell} f^{(\ell)} - f^{(\ell)} \right\| + \left\| B_n^{(\ell)} f - B_{n-\ell} f^{(\ell)} \right\| \\
&\le \underbrace{\left\| B_{n-\ell} f^{(\ell)} - f^{(\ell)} \right\|}_{\to 0} + \underbrace{\left(1 - \prod_{j=1}^{\ell-1} \left(1 - \frac{j}{n} \right) \right)}_{\to 0} \left\| f^{(\ell)} \right\| + \underbrace{\omega\left(f^{(\ell)}, \frac{\ell}{n}\right)}_{\to 0},
\end{aligned}$$

und solange man nur *endlich viele* Werte $\ell = 0, \ldots, k$ betrachtet, konvergiert das auch "gleichmäßig" in ℓ gegen Null, und das liefert (4.4). □

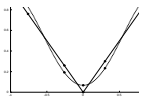
Übung 4.5 Zeigen Sie, daß für $f \in C^k(I)$

$$\left|\Delta_h^k f(x)\right| \leq \frac{1}{h^k} \left\|f^{(k)}\right\|, x, x + kh \in I.$$

♢

4.3 Shape preservation

Bestapproximation ist ja eine feine Sache: Unter allen zugelassenen Funktionen weicht die Bestapproximation von der Zielfunktion global am wenigsten ab. Anders wird es aber, wenn die Approximation die "Gestalt" der Funktion f widerspiegeln soll, denn da stören Oszillationen möglicherweise, siehe Abb. 4.1. Besonders schöne (und auch praktisch wichtige) "Shape pro-



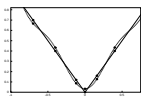


Abbildung 4.1: Bestapproximationen vom Grad 5 (links) und 9 (rechts) an $f(x) = |x|$ auf $I = [-1, 1]$. Wie man sieht, fordert der Alternantensatz seinen Preis – die beiden Approximationen an die konvexe Funktion f sind nicht mehr konvex.

perties" sind beispielsweise

- Posivität,
- Monotonie,
- Konvexität,

die sich für $f \in C^2(I)$ als $f \geq 0$, $f' \geq 0$, $f'' \geq 0$ beschreiben lassen. Diese Eigenschaften werden nun von Bernsteinpolynomen erhalten.

Satz 4.8 *Ist $f \in C^k(I)$, dann ist $f^{(k)} \geq 0$ genau dann wenn $B_n^{(k)} f \geq 0$ ist.*

Satz 4.8 folgt sofort aus dem folgenden Resultat.

Proposition 4.9 *Sei $k \geq 0$ und $f \in C(I)$*

1. *Ist $\Delta_h^k f \geq 0$, dann ist auch $B_n^{(k)} f \geq 0$.*
2. *Ist zusätzlich $f \in C^k(I)$, dann ist*

$$f^{(k)} \geq 0 \qquad \Longleftrightarrow \qquad \Delta_h^k f \geq 0, \quad h > 0.$$

Beweis: Da

$$B_n^{(k)} f = \frac{n!}{(n-k)!} \sum_{j=0}^{n-k} \Delta_{1/n}^k f \left(\frac{j}{n} \right) B_j^{n-k},$$

folgt 1) unmittelbar; andererseits ist 2) eine unmittelbare Folgerung aus der Identität

$$\Delta_h^k f = \int_{[0,h]^k} f^{(k)} \left(\cdot + \sum_{j=1}^{k} t_j \right) dt_1 \cdots dt_k$$

ist. □

Korollar 4.10 *Eine Funktion $f \in C(I)$ ist genau dann konvex, wenn alle ihre Bernsteinpolynome $B_n f$, $n \in \mathbb{N}$, konvex sind.*

Übung 4.6 Zeigen Sie: $f \in C(I)$ ist genau dann konvex, wenn $\Delta_h^2 f(x) \geq 0$ für alle $h > 0$ und alle x mit $x, x + 2h \in I$. ♢

Eine weitere "shape"–Eigenschaft, die wirklich große Bedeutung in der Praxis hat, ist die *Variationsverminderung*. Um auch zu wissen, was da wirklich vermindert wird, erst einmal eine Definition.

Definition 4.11 *Sei $f \in C(I)$.*

1. *Die* totale Variation *von f ist definiert als*

$$V(f) := \sup_{x_0 < \ldots < x_n \in I} \sum_{j=0}^{n-1} \left| f\left(x_{j+1}\right) - f\left(x_j\right) \right|. \tag{4.9}$$

2. *Die* Variation *eines Vektors $(f_j \ : \ j = 0, \ldots, n)$ ist definiert als als*

$$V\left(f_j \ : \ j = 0, \ldots, n\right) := \sum_{j=0}^{n-1} \left| f_j - f_{j-1} \right|. \tag{4.10}$$

Übung 4.7 (Geometrische Interpretation der Variation eines Vektors)
Es sei $f = (f_j : j = 0, \ldots, n) \in \mathbb{R}^{n+1}$ und sei ℓ_f die auf den Intervallen $\left[\frac{j}{n}, \frac{j+1}{n}\right]$, $j = 0, \ldots, n-1$, *stückweise lineare* Funktion, die außerdem

$$\ell_f\left(\frac{j}{n}\right) = f_j, \qquad j = 0, \ldots, n,$$

erfüllt. Zeigen Sie, daß $V(f_j : j = 0, \ldots, n) = V(\ell_f)$. ♢

Bemerkung 4.12 *Die Summe bei der Definition der totalen Variation in (4.9) ist eigentlich gar nicht so schlimm: Ist nämlich $f(x_{j-1}) \leq f(x_j) \leq f(x_{j+1})$, dann ist der enstprechende Teil der Summe*

$$\cdots + (f(x_j) - f(x_{j-1})) + (f(x_{j+1}) - f(x_j)) + \cdots = \cdots + f(x_{j+1}) - f(x_{j-1}) + \cdots$$

und man kann den Punkt x_j also getrost weglassen.
Mit anderen Worten: Man könnte genausogut gleich über die lokalen Extrema summieren[83]*; das erklärt wohl auch den Namen ein bißchen besser.*

Satz 4.13 *(Variationsverminderung durch Bernsteinpolynome)*
Für $f \in C(I)$ und $n \in \mathbb{N}$ gilt

$$V(B_n f) \leq V\left(f\left(\tfrac{j}{n}\right) : j = 0, \ldots, n\right) \leq V(f). \tag{4.11}$$

Bemerkung 4.14 *Es gibt noch einen anderen, geometrischen, Begriff der Variationsverminderung: Ein Operator T wird als* variationsvermindernd *bezeichnet, wenn Tf mit einer* beliebigen *Gerade höchstens so viele Schnittpunkte hat wie f selbst. Auch diese Eigenschaft besitzen die Bernsteinpolynome, aber zu deren Beweis müßte man etwas weiter ausholen, es wäre eher Stoff für eine CAGD*[84]*–Vorlesung.*

Wichtigster Bestandteil des Beweises von Satz 4.13 ist eine kleine Beobachtung über die totale Variation, die auch die Grundlage der Definition des *Stieltjes*[85]*–Integral* $\int df$ ist.

Lemma 4.15 *Für $f \in C^1(I)$ ist*

$$V(f) \leq \int_I |f'(t)|\, dt. \tag{4.12}$$

[83] Was natürlich besonders schön ist, wenn eine Funktion nur endlich viele davon haben kann, wie zum Beispiel Polynome.

[84] Computer Aided Geometric Design.

[85] Thomas Jan Stieltjes, 1856–1894, holländischer Analytiker, schaffte es, dreimal durch die Abschlußprüfung zu fallen; wegen des fehlenden Universitätsabchlusses wurde ihm später vom Ministerium (bzw. wegen eine königlichen Dekrets vom 12.3.1884) der Analysis–Lehrstuhl in Groningen nicht zuerkannt, obwohl er der Wunschkandidat der Kommission war.

Übung 4.8 Zeigen Sie: In (4.12) gilt sogar Gleichheit. ♢

Beweis: Für jede Wahl von Punkten $x_0 < \cdots < x_n \in I$ ist

$$\begin{aligned}\sum_{j=0}^{n-1} |f(x_{j+1}) - f(x_j)| &= \sum_{j=0}^{n-1} \left| \int_{x_j}^{x_{j+1}} f'(t)\ dt \right| \leq \sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} |f'(t)|\ dt = \int_{x_0}^{x_n} |f'(t)|\ dt \\ &\leq \int_I |f'(t)|\ dt,\end{aligned}$$

was sich auch auf das Supremum über alle solchen Ausdrücke, $V(f)$ überträgt. □

Beweis von Satz 4.13: Unter Verwendung von (4.2), Lemma 4.15 und Übung 4.9 erhalten wir, daß

$$\begin{aligned}V(B_n f) &\leq \int_0^1 |B_n' f(t)|\ dt = \int_0^1 \left| n \sum_{j=0}^{n-1} \Delta_{1/n} f\left(\frac{j}{n}\right)\ B_j^{n-1}(t) \right| dt \\ &\leq n \int_0^1 \sum_{j=0}^{n-1} \left| \Delta_{1/n} f\left(\frac{j}{n}\right) \right| B_j^{n-1}(t)\ dt = n \sum_{j=0}^{n-1} \left| \Delta_{1/n} f\left(\frac{j}{n}\right) \right| \underbrace{\int_0^1 B_j^{n-1}(t)\ dt}_{=\frac{1}{n}} \\ &= \sum_{j=0}^{n-1} \left| f\left(\frac{j+1}{n}\right) - f\left(\frac{j}{n}\right) \right| = V\left(f\left(\tfrac{j}{n}\right)\ :\ j = 0, \ldots, n\right) \leq V(f),\end{aligned}$$

und das war's auch schon. □

Übung 4.9 Zeigen Sie: Für $n \geq 0$ ist

$$\int_0^1 B_j^n(t)\ dt = \frac{1}{n+1}, \qquad j = 0, \ldots, n.$$

♢

Übung 4.10 Zeigen Sie: Die polynomiale Bestapproximation ist im allgemeinen nicht variationsvermindernd. ♢

4.4 Der Preis: Saturation

Wir haben gesehen, daß Bernsteinpolynome "schöne" Approximanten sind, die viele Eigenschaften der Zielfunktion auf sich übertragen. Man kann sich denken, daß es das nicht umsonst gibt – und in der Tat werden wir einen Preis bezahlen, nämlich "langsame" Konvergenz der Approximationen, also eine schlechte Approximationsordnung.

Das Phänomen mit dem wir uns hier beschäftigen wollen, die *Saturation*, zeigt sich grob gesprochen daran, daß man machen kann, was man will, besser als eine bestimmte Approximationsordnung geht's eigentlich nicht.

Definition 4.16 *Sei $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ eine stetige Funktion mit $\lim_{x\to\infty} \varphi(x) = 0$.*
Eine Folge T_n, $n \in \mathbb{N}$, von Approximationsoperatoren heißt saturiert *mit Ordnung φ, wenn*

1. *("o–Klasse", "triviale Klasse")*

$$\lim_{n\to\infty} \varphi^{-1}(n) \, \|T_n f - f\| = 0 \qquad \Longleftrightarrow \qquad T_n f = f, \quad n \in \mathbb{N}. \tag{4.13}$$

2. *("O–Klasse") für jedes $M > 0$*

$$\left\{ f \ : \ \limsup_{n\to\infty} \varphi^{-1}(n) \, \|T_n f - f\| < \infty \right\} \setminus \{ f \ : \ T_n f = f \} \neq \emptyset. \tag{4.14}$$

Mit anderen Worten: Die "optimale" Approximationsordnung $O(\varphi(n))$ wird für eine hinreichend große Klasse von Funktionen erreicht, jede "etwas bessere" Approximationsordnung $o(\varphi(n))$ hingegen nur von denjenigen Funktionen, die $T_n f = f$ erfüllen, also trivialerweise *sehr gut approximiert werden.*

Satz 4.17 *Die Bernsteinpolynome B_n sind saturiert mit Ordnung $\varphi(n) = n^{-1}$.*

Um einschätzen zu können, warum dieses Resultat so "unerfreulich" ist, muß man wissen, daß die polynomiale Bestapproximation *beliebig* schnell[86] konvergieren kann, siehe Satz 5.3
Unser erstes (und fast wichtigstes) Hilfsmittel ist die "Voronovskaja–Formel" [84], aus der wir sofort ersehen können, daß die O–Klasse wirklich groß ist, nämlich mindestens alle zweimal stetig differenzierbaren Funktionen umfasst, deren zweite Ableitung nicht identisch verschwindet[87].

Satz 4.18 *(Voronovskaja, [84])*
Für jedes $f \in C^2[0,1]$ und $x \in [0,1]$ ist

$$\lim_{n\to\infty} n \, (B_n f - f)(x) = \frac{x(1-x)}{2} \, f''(x). \tag{4.15}$$

Beweis: Wir fixieren $x \in [0,1]$ und bestimmen $f\left(\frac{j}{n}\right)$ durch eine Taylor–Entwicklung um x als

$$f\left(\frac{j}{n}\right) = f(x) + \left(\frac{j}{n} - x\right) f'(x) + \left(\frac{j}{n} - x\right)^2 \frac{f''(\xi_j)}{2}, \tag{4.16}$$

wobei $\xi_j \in \left[\frac{j}{n}, x\right]$. Multiplizieren wir nun beide Seiten von (4.16) mit[88] $B_j^n(x)$ und summieren über $j = 0, \ldots, n$, dann erhalten wir unter Verwendung einiger Identitäten aus dem Beweis von

[86] Und langsam

[87] Und das wären, nur um daran zu erinnern, gerade Π_1, die linearen Polynome. Und da $B_n p = p$ für $p \in \Pi_1$ steht zu befürchten, daß uns die nochmal begegnen werden . . ..

[88] Ja, das x hier ist dasselbe, das wir vorher festgehalten haben!

Satz 2.2, daß

$$\begin{aligned} B_n f(x) &= \sum_{j=0}^{n} f\left(\frac{j}{n}\right) B_j^n(x) = f(x) \underbrace{\sum_{j=0}^{n} B_j^n(x)}_{=1} + f'(x) \underbrace{\sum_{j=0}^{n} \left(\frac{j}{n} - x\right) B_j^n(x)}_{=x-x=0} \\ &\quad + \frac{f''(x)}{2} \underbrace{\sum_{j=0}^{n} \left(\frac{j}{n} - x\right)^2 B_j^n(x)}_{=\frac{x(1-x)}{n}} + \frac{1}{2} \underbrace{\sum_{j=0}^{n} \left(\frac{j}{n} - x\right)^2 \left(f''\left(\xi_j\right) - f''(x)\right) B_j^n(x)}_{=:R_n(x)} \\ &= f(x) + \frac{x(1-x)}{2} f''(x) + \frac{1}{2} R_n(x). \end{aligned} \tag{4.17}$$

Bleibt also die Abschätzung von R_n. Hierfür wählen wir wie im Beweis von Satz 2.2 wieder $\delta > 0$ und spalten auf, weswegen

$$\begin{aligned} |R_n(x)| &\leq \sum_{j=0}^{n} \left(\frac{j}{n} - x\right)^2 \left|f''\left(\xi_j\right) - f''(x)\right| B_j^n(x) \\ &= \sum_{\left|\frac{j}{n}-x\right| \leq \delta} \left(\frac{j}{n} - x\right)^2 \underbrace{\left|f''\left(\xi_j\right) - f''(x)\right|}_{\leq \omega(f'', \delta)} B_j^n(x) \\ &\quad \sum_{\left|\frac{j}{n}-x\right| > \delta} \left(\frac{j}{n} - x\right)^2 \underbrace{\left|f''\left(\xi_j\right) - f''(x)\right|}_{\leq 2\|f''\|} B_j^n(x) \\ &\leq \omega\left(f'', \delta\right) \underbrace{\sum_{j=0}^{n} \left(\frac{j}{n} - x\right)^2 B_j^n}_{=\frac{x(1-x)}{n}} + 2\left\|f''\right\| \frac{1}{\delta^2} \underbrace{\sum_{j=0}^{n} \left(\frac{j}{n} - x\right)^4 B_j^n}_{\leq \frac{C}{n^2}}, \end{aligned}$$

siehe Übung 4.11, also

$$n\left|R_n(x)\right| \leq \frac{\omega\left(f'', \delta\right)}{4} + \frac{C}{n\delta^2}. \tag{4.18}$$

Nun wählen wir für vorgegbenes $\varepsilon > 0$ wieder zuerst δ so klein, daß $\omega\left(f'', \delta\right) < 2\varepsilon$ und dann n so groß, $\frac{C}{n\delta^2} < \frac{\varepsilon}{2}$, so daß, nach (4.18)

$$\lim_{n\to\infty} n\left\|R_n\right\| = 0.$$

Einsetzen in (4.17) liefert somit, daß

$$\lim_{n\to\infty} \left(B_n f - f\right)(x) = \frac{x(1-x)}{2} f''(x) + \frac{1}{2} \underbrace{\lim_{n\to\infty} R_n(x)}_{=0} = \frac{x(1-x)}{2} f''(x),$$

was gerade (4.15) ist. □

Übung 4.11 Zeigen Sie: Es gibt eine Konstante C, so daß

$$\sum_{j=0}^{n} \left(\frac{j}{n} - x \right)^4 B_j^n \leq \frac{C}{n^2}.$$

Hinweis: Bestimmen Sie mit denselben Methoden wie im Beweis von Satz 2.2 das Polynom exakt. ♢

Für den Beweis des "o–Satzes" gehen wir *geometrisch* vor und verwenden eine Idee von Amel'kovič [2] bzw. von Bajšanski und Bojanić [5]; dazu benötigen wir die folgenden Beschreibung der Konvexität.

Satz 4.19 *Für $f \in C(I)$ sind äquivalent:*

1. *f ist konvex.*
2. *$B_n f$ ist konvex, $n \in \mathbb{N}$.*
3. *Für $n \in \mathbb{N}$ ist $B_n f \geq B_{n+1} f$.*
4. *Für $n \in \mathbb{N}$ ist $B_n f \geq f$.*
5. *Für $x \in [0,1]$ ist*
$$\limsup_{n \to \infty} n \left(B_n f - f \right)(x) \geq 0. \tag{4.19}$$

Korollar 4.20 *Für $f \in C(I)$ sind äquivalent*

1. *$f \in \Pi_1$.*
2. *Es ist*
$$\lim_{n \to \infty} n \, \| B_n f - f \| = 0$$
3. *Für jedes $x \in [0,1]$ ist*
$$\lim_{n \to \infty} n \left(B_n f - f \right)(x) = 0.$$

Beweis: Die Schlüsse 1) ⇒ 2) ⇒ 3) sind trivial. Ist aber die Limesbedingung in 3) erfüllt, so ist für jedes $x \in [0,1]$

$$0 = \lim_{n \to \infty} n \left(B_n f - f \right)(x) = \lim_{n \to \infty} n \underbrace{\left(f - B_n f \right)}_{= B_n(-f) - (-f)}(x),$$

also sind f und $-f$ konvex, das heißt, f ist gleichzeitig konvex *und* konkav und damit eine affine Funktion. □

Um uns den Beweis ein klein wenig einfacher machen zu können, verwenden wir eine "Formel" aus dem Reich der "angewandten" Bernsteipolynome alias *Bézier–Kurven*[89].

[89] Dem Namen Bézier sind wir ja vorher schon begegnet; Bézier–Kurven sind Bestandteil jedes Grafikprogramms, ein Hilfsmittel zur Modellierung von sogenannten "Freiformkurven".

Lemma 4.21 *Für $n \in \mathbb{N}_0$ gilt die* Graderhöhungsformel[90]

$$\sum_{j=0}^{n} f_j \, B_j^n = \sum_{j=0}^{n+1} \widehat{f}_j \, B_j^{n+1}, \qquad \widehat{f}_j = \frac{n+1-j}{n+1} \, f_j + \frac{j}{n+1} \, f_{j-1}, \quad j = 0, \ldots, n+1. \tag{4.20}$$

Beweis: Wir rechnen einfach nach, daß

$$\begin{aligned}
\sum_{j=0}^{n+1} \widehat{f}_j \, B_j^{n+1}(x) &= \sum_{j=0}^{n+1} \left(\frac{n+1-j}{n+1} \, f_j + \frac{j}{n+1} \, f_{j-1} \right) \, B_j^{n+1}(x) \\
&= \sum_{j=0}^{n} \frac{n+1-j}{n+1} \, f_j \, B_j^{n+1}(x) + \sum_{j=0}^{n} \frac{j+1}{n+1} \, f_j \, B_{j+1}^{n+1}(x) \\
&= \sum_{j=0}^{n} \frac{n+1-j}{n+1} \, f_j \, \frac{n+1}{n+1-j} (1-x) \, B_j^n(x) + \sum_{j=0}^{n} \frac{j+1}{n+1} \, f_j \, \frac{n+1}{j+1} x \, B_j^n(x) \\
&= \sum_{j=0}^{n} ((1-x) + x) \; f_j \, B_j^n(x) = \sum_{j=0}^{n} f_j \, B_j^n(x),
\end{aligned}$$

gilt[91]. □

Beweis von Satz 4.19: Die Äquivalen z von 1) und 2) haben wir ja schon (Korollar 4.10), außerdem ist 3) $\Longrightarrow$ 4) $\Longrightarrow$ 5) trivial, so daß nur noch zwei Sachen zu beweisen bleiben:

1) $\Rightarrow$ 3): Mit $\alpha_j = \frac{j}{n+1}$ ist

$$\alpha_j \, \frac{j-1}{n} + (1-\alpha_j) \; \frac{j}{n} = \frac{j}{n} - \frac{\alpha_j}{n} = \frac{(n+1)j - j}{n(n+1)} = \frac{j}{n+1},$$

weswegen mit der Graderhöhungsformel (4.20) für jedes konvexe $f \in C(I)$

$$\begin{aligned}
B_n f - B_{n+1} f &= \sum_{j=0}^{n+1} \underbrace{\left(\alpha_j \, f\left(\frac{j-1}{n} \right) + (1-\alpha_j) \; f\left(\frac{j}{n} \right) - f\left(\frac{j}{n+1} \right) \right)}_{\geq 0} B_j^{n+1} \\
&\geq 0
\end{aligned}$$

gilt.

5) $\Rightarrow$ 1): Angenommen, f wäre nicht konvex, dann gibt es Punkte $x_0 < x' < x_1$, $x' = \alpha \, x_0 + (1-\alpha) x_1$, so daß

$$f(x') > \alpha \, f(x_0) + (1-\alpha) \, f(x_1),$$

[90] Bezüglich der Monombasis läßt sich jedes Polynom vom Grad n auch als eines vom Grad $n+1$ schreiben, indem man den "Leitkoeffizienten" a_{n+1} gleich Null setzt; bei Verwendung der Basen B_j^n, $j = 0, \ldots, n$, bzw. B_j^{n+1}, $j = 0, \ldots, n+1$, ist das, wie man sieht, nicht mehr ganz so trivial.

[91] Richtig schön wird diese Formel erst nach Einführung baryzentrischer Koordinaten und in mehreren Variablen.

also

$$f(x') > \alpha f(x_0) + (1-\alpha) f(x_1) + \varepsilon\, x'(1-x')$$

für ein hinreichend kleines $\varepsilon > 0$ – schließlich ist ja $x' \in (0,1)$. Sei $\ell \in \Pi_1$ die lineare Funktion mit $\ell(x_j) = f(x_j)$, $j = 0, 1$, und

$$g(x) = \ell(x) + \varepsilon\, x(1-x) \qquad \Longrightarrow \qquad g(x') < f(x'), \quad g(x_j) > f(x_j), \quad j = 0, 1.$$

Daher gibt es ein Intervall $[a, b] \subset [x_0, x_1]$, so daß

$$(f-g)(x) > 0, \qquad x \in [a, b],$$

und es gibt eine Stelle $x^* \in [a, b]$, wo $f - g$ sein erst recht positives Maximum annimmt. Seien $a < a' < x^* < b' < b$, dann ist für $x \in [a', b']$

$$0 > (f-g)(x) - (f-g)(x^*),$$

also

$$f(x) < \widetilde{g}(x) := g(x) + (f-g)(x^*) \qquad \text{und} \qquad f(x^*) = \widetilde{g}(x^*).$$

Außerdem setzen wir $\widetilde{g}$ außerhalb von $[a', b']$ zu einer *zweimal stetig differenzierbaren* Funktion auf $[0, 1]$ fort, und zwar so, daß $\widetilde{g} \geq f$ gilt – das läßt sich ohne weiteres erreichen. Jetzt haben wir's auch schon fast geschafft: Da $\widetilde{g} \in C^2[0,1]$, können wir Satz 4.18 anwenden und erhalten, daß

$$\limsup_{n\to\infty} n\, (B_n f - f)(x^*) \leq \limsup_{n\to\infty} n \left(B_n \widetilde{g}(x^*) - \underbrace{f(x^*)}_{=\widetilde{g}(x^*)} \right) = \lim_{n\to\infty} n\, (B_n \widetilde{g} - \widetilde{g})(x^*)$$

$$= \frac{x^*(1-x^*)}{2} \widetilde{g}''(x^*) = \frac{x^*(1-x^*)}{2} \underbrace{\left(\frac{d^2}{dx^2} \varepsilon x(1-x) \right)(x^*)}_{=-2\varepsilon} = -\varepsilon\, x^*(1-x^*),$$

also

$$\limsup_{n\to\infty} (B_n f - f)(x^*) < 0,$$

im Widerspruch zu (4.19).

□

Bemerkung 4.22 *Der zweite Teil des obigen Beweises wird gerne auch als "Parabelmethode" bezeichnet.*

1. *Diese Methode geht auf H. A. Schwarz*[92]*zurück, siehe [28], der diese Idee benutzte, um zu zeigen, daß*

$$\limsup_{h\to 0^+} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \geq 0, \qquad x \in I,$$

[92] Hermann Amandus Schwarz, 1843–1921, ein Schüler von Weierstraß, Professuren in Zürich (ETH) und Göttingen, bevor er sich 1892 auf der Weierstraß–Nachfolge in Berlin "zur Ruhe setzte" (die Bemerkung stammt von Bieberbach). Arbeitete unter anderem an der Theorie der Minimalflächen und ist durch die Cauchy–Schwarz–Ungleichung "verewigt".

zur Konvexität einer stetigen[93] *Funktion f äquivalent ist.*

2. *Tatsächlich besteht die Idee darin, eine Funktion an ihrem Maximimum lokal nicht nur durch eine Gerade, also die "Tangente", sondern durch eine passend gekrümmte Parabel zu beschränken, siehe Abb. 4.2.*

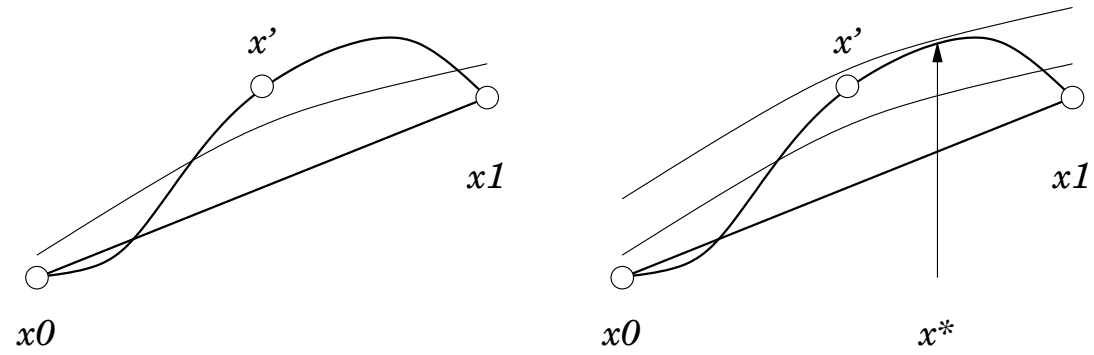


Abbildung 4.2: Die "Parabelmethode": Zuerst "quetscht" man die Parabel zwischen die Gerade und *unter* die Funktion und dann schiebt man sie an der Stelle, an der der Unterschied maximal wird, nach oben, so daß sie dort berührt und die Funktion – zumindest lokal – majorisiert.

Und in der Tat ist die Saturation, die eher gemächliche Konvergenz der Bernsteipolynome ein Preis, den man unvermeidbar für die "Shape properties" bezahlen muß. Es gilt nämlich das folgende Resultat von Berens und DeVore [6].

Satz 4.23 *Sei $T_n : C[0,1] \to \Pi_n$ eine Folge von polynomialen Operatoren*[94]*, die Positivität beliebiger Ordnung erhalten:*

$$f^{(j)} \geq 0 \qquad \Longrightarrow \qquad T_n^{(j)} f \geq 0, \quad n \in \mathbb{N}.$$

Dann gibt es ein $f \in C(I)$, so daß $B_n f \leq T_n f$, $n \in \mathbb{N}$, genauer: mit $f = (\cdot - x)^2$ ist

$$\frac{x(1-x)}{n} = B_n f(x) \leq T_n f(x)$$

mit Gleichheit dann und nur dann, wenn $T_n = B_n$.

Übrigens ist bereits *Positivität* eines linearen Operators eine ganz schön heftige Einschränkung, wie die folgenden Resultate von Korovkin zeigen – die eine ganze "Korovkin–Theorie" ausgelöst haben.

[93] Für $f \in C^2(I)$ ist das Ganze ziemlich klar, denn der (symmetrische) Differenzenquotient konvergiert gleichmäßig gegen $f''(x)$, wie man praktisch sofort aus einer Darstellung entsprechend (4.5) ersieht.

[94] Es ist schon wichtig, daß der Grad von $T_n f \leq n$ ist, denn ansonsten kann man trivialerweise "superschnelle" Approximationsoperatoren konstruieren, z.B. $T_n = B_{2^n}$ oder derartigen Blödsinn.

Satz 4.24 *Sei T_n, $n \in \mathbb{N}_0$, eine Folge* positiver *Operatoren.*

1. *Es gilt*

$$\lim_{n\to\infty} \|T_n f - f\| = 0, \quad f \in C(I) \quad \Longleftrightarrow \quad \lim_{n\to\infty} \|T_n f - f\| = 0, \quad f \in \left\{1, x, x^2\right\}.$$

2. *T_n ist saturiert mit Ordnung bestenfalls n^{-2}.*

4.5 Multivariate Bernsteinpolynome

In diesem Kapitel sehen wir uns ein paar dieser Konzepte in *mehreren* Variablen an, wobei wir nicht mehr jedes Detail beweisen wollen. Bernsteinpolynome auf Dreiecken und höherdimensionalen Simplizes wurden unabhängig voneinander von Lorentz in [45] und von Dinghas[95] [19] eingeführt.

Schlägt man sich mit *Simplizes* im $\mathbb{R}^d$ herum, ist es immer gut, die sogenannten *baryzentrischen Koordinaten* zu verwenden. Was das ist? Nun, seien $v_0, \ldots, v_d \in \mathbb{R}^d$, dann kann man (siehe Lemma 3.13) jeder Punkt $x \in \Delta := [v_j \ : \ j = 0, \ldots, d]$ als

$$x = \sum_{j=0}^{n} u_j(x)\, v_j, \qquad \sum_{j=0}^{n} u_j(x) = 1, \qquad u_j(x) \geq 0, \quad j = 0, \ldots, d, \tag{4.21}$$

darstellbar. Der Vektor $u = u(x) = (u_j(x) \ : \ j = 0, \ldots, n)$ heißt *baryzentrische Koordinaten* (siehe Abb. 4.3) von x bezüglich $v_0, \ldots, v_d$. Schreibt man (4.21) als lineares Gleichungssystem, so erhält man, daß

$$\begin{bmatrix} 1 & \ldots & 1 \\ v_{0,1} & \ldots & v_{d,1} \\ \vdots & \ddots & \vdots \\ v_{0,d} & \ldots & v_{d,d} \end{bmatrix} \begin{bmatrix} u_0(x) \\ \vdots \\ u_d(x) \end{bmatrix} = \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_d \end{bmatrix},$$

was genau dann eindeutig lösbar ist, wenn die Punkte $v_0, \ldots, v_d$ *in allgemeiner Lage* sind, das heißt, wenn die Vektoren $v_1 - v_0, \ldots, v_n - v_0$ *linear unabhängig* sind. Denn schließlich ist

$$\begin{aligned} \pm\,(d+1)!\,\mathrm{vol}_d(\Delta) &= \det \begin{bmatrix} 1 & \ldots & 1 \\ v_{0,1} & \ldots & v_{d,1} \\ \vdots & \ddots & \vdots \\ v_{0,d} & \ldots & v_{d,d} \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & \ldots & 0 \\ v_{0,1} & v_{1,1} - v_{0,1} & \ldots & v_{d,1} - v_{0,1} \\ \vdots & \vdots & \ddots & \vdots \\ v_{0,d} & v_{0,d} - v_{0,d} & \ldots & v_{d,d} - v_{0,d} \end{bmatrix} \\ &= \det \begin{bmatrix} v_{1,1} - v_{0,1} & \ldots & v_{d,1} - v_{0,1} \\ \vdots & \ddots & \vdots \\ v_{0,d} - v_{0,d} & \ldots & v_{d,d} - v_{0,d} \end{bmatrix}, \end{aligned}$$

wobei $\mathrm{vol}_d(\Delta)$ das d–dimensionale Volumen des Simplex Δ bezeichnet.

[95] Alexander Dinghas, 1908–1974, in Izmir geboren und teilweise dort, teilweise (seit 1922) in Athen aufgewachsen. Beiträge zur Funktionentheorie, insbesondere zur "Nevalinna–Theorie" oder Wachstum subharmonischer Funktionen, und zur Differentialgeometrie.

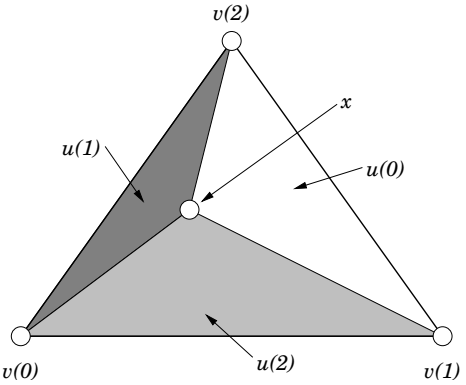


Abbildung 4.3: Geometrische Interpretation der baryzentrischen Koordinaten, dargestellt für $d = 2$: Der Punkt x unterteilt Simplex in $d + 1$ Teilsimplizes $[V \setminus \{v_j\} \cup \{x\}]$, $j = 0, \ldots, d$, und die baryzentrischen Koordinaten sind das Verhältnis des Volumens dieser Teilsimplizes zum Gesamtvolumen des "großen" Simplex. Daß sich diese Koordinaten zu 1 summieren ist dann ziemlich offensichtlich – hier ist das Ganze eben doch nur die Summe seiner Teile!

Definition 4.25 *Mit*

$$\mathbb{S}_d := \left\{ u \in \mathbb{R}^{d+1} \ : \ u_j \geq 0, \sum_{j=0}^{d} u_j = 1 \right\}$$

bezeichnen wir das Standardsimplex *in baryzentrischen Koordinaten.*

Ob wir nun Funktionen auf einem beliebigen nichtdegenerierten Simplex $\Delta \in \mathbb{R}^d$ oder gleicht auf $\mathbb{S}_d$ betrachten, das spielt nun, nach Einführung baryzentrischer Koordinaten, eben gerade keine Rolle mehr.

Definition 4.26 *1. Die* Länge *eines* Multiindex $\alpha = (\alpha_0, \ldots, \alpha_d) \in \mathbb{N}_0^{d+1}$ *ist definiert als*

$$|\alpha| = \sum_{j=0}^{d} \alpha_j.$$

2. Mit

$$\Gamma_n := \left\{ \alpha \in \mathbb{N}_0^{d+1} \ : \ |\alpha| = n \right\}, \qquad n \in \mathbb{N}_0,$$

bezeichnen wir die Gesamtheit aller (homogenen) Multiindizes der Länge n.

3. Zu $\alpha \in \Gamma_n$ ist das Bernstein–Bézier–Basispolynom $B_\alpha : \mathbb{S}_d \to \mathbb{R}$ *definiert als*

$$B_\alpha(u) = \frac{|\alpha|}{\alpha_0! \cdots \alpha_d!} u_0^{\alpha_0} \cdots u_d^{\alpha_d} =: \binom{|\alpha|}{\alpha} u^\alpha.$$

4. Zu $f \in C(\mathbb{S}_d)$ ist das n–te Bernsteinpolynom $B_n f$ definiert als

$$B_n f = \sum_{\alpha \in \Gamma_n} f\left(\frac{\alpha}{n}\right) B_\alpha.$$

Bemerkung 4.27 *Definition 4.26 ist eine echte Verallgemeinerung von $d = 1$: Im Falle des Standardsimplex $I = [0,1]$ ist $u = u(x) = (1-x, x)$ und $\alpha = (n-j, j)$.*

Betrachtet man einmal beispielsweise ein dreidimensionales Simplex, so stellt man fest, daß seine Seiten Dreiecke, also zweidimensionale Simplizes, und die Seiten dieser Seiten Streckenzüge, also eindimensionale Simplizes, sind. Außerdem sind Punkte ja auch noch nulldimensionale Simplizes. Solche Teilsimplizes können wir adressieren über $\delta \in \{0,1\}^{d+1}$, indem wir

$$\mathbb{S}_\delta := \{u \in \mathbb{S}_d \ : \ u \le \delta\} \ \simeq \ \mathbb{S}_{|\delta|-1}$$

einführen, siehe Abb 4.4. Dann ergibt sich die "Lokalitätsformel" für multivariate Bernstein-

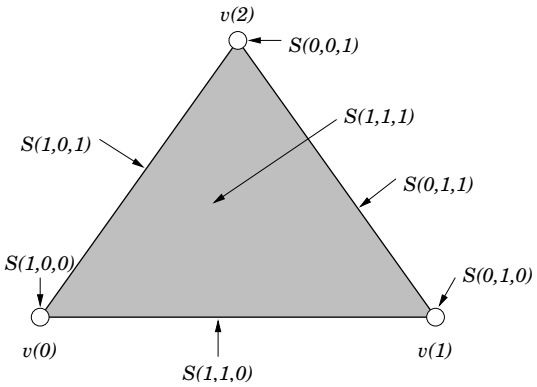


Abbildung 4.4: Ein zweidimensionales Simplex (=Dreieck) und die entsprechenden Subsimplizes $\mathbb{S}_\delta$.

polynome

$$(B_n f)|_{\mathbb{S}_\delta} = B_n\left(f|_{\mathbb{S}_\delta}\right), \qquad \delta \in \{0,1\}^{d+1}, \tag{4.22}$$

wobei das Bensteinpolynom auf der rechten Seite eines in $|\delta| - 1$ Variablen ist. Und wirklich – dieses "Innen–Rand"–Verhalten ist der Schlüssel zum Verständnis multivariater Bernsteinpolynome.

Übung 4.12 Zeigen Sie: Sind $\delta_1, \ldots, \delta_n \in \{0,1\}^{d+1}$, dann gilt

$$\delta = \bigvee_{j=1}^{n} \delta_j \qquad \Longrightarrow \qquad \mathbb{S}_\delta = \left[\mathbb{S}_{\delta_j} \ : \ j = 0, \ldots, n\right],$$

wobei $\vee$ das komponentenweise Maximum und $[\cdot]$ die konvexe Hülle bezeichnet. $\diamondsuit$

Um schon einmal einen Vorgeschmack dafür zu erhalten, wie schön man mit baryzentrischen Koordinaten rechnen kann, sehen wir uns das Gegenstück der Graderhöhungsformel an.

Proposition 4.28 *Es ist*

$$\sum_{\alpha\in\Gamma_n} f_\alpha\, B_\alpha = \sum_{\alpha\in\Gamma_{n+1}} \widehat{f}_\alpha\, B_\alpha, \qquad \Longleftrightarrow \qquad \widehat{f}_\alpha = \sum_{j=0}^{d} \frac{\alpha_j}{n+1} f_{\alpha-e_j}, \quad \alpha\in\Gamma_{n+1}. \tag{4.23}$$

Beweis: Mit

$$\begin{aligned}\sum_{\alpha\in\Gamma_{n+1}} \widehat{f}_\alpha\, B_\alpha &= \sum_{\alpha\in\Gamma_{n+1}} \sum_{j=0}^{d} \frac{\alpha_j}{n+1} f_{\alpha-e_j}\, B_\alpha = \sum_{\alpha\in\Gamma_n} f_\alpha \sum_{j=0}^{d} \frac{\alpha_j+1}{n+1}\, B_{\alpha+e_j} \\ &= \sum_{\alpha\in\Gamma_n} f_\alpha \sum_{j=0}^{d} \frac{\alpha_j+1}{n+1}\, \frac{(n+1)!}{\alpha_0!\cdots(\alpha_j+1)!\cdots\alpha_d!} u^{\alpha+e_j} = \underbrace{\left(\sum_{j=0}^{d} u_j\right)}_{=1} \sum_{\alpha\in\Gamma_n} f_\alpha\, B_\alpha\end{aligned}$$

folgt die Behauptung. □

Um unseren Beweis des "o"–Resultats auf den multivariaten Fall übertragen zu können, brauchen wir natürlich Information inwieweit Konvexität und monotone Konvergenz miteinander zu tun haben. Dazu sollten wir erst einmal klären, was Konvexität genau bedeutet und noch ein paar andere Konvexitätsbegriffe einführen.

Definition 4.29 *1. Ein Vektor $y\in\mathbb{R}^{d+1}$ heißt* Richtung *in $\mathbb{S}_d$, wenn es Punkte $u,u'\in\mathbb{S}_d$ gibt, so daß $u'=u+y$, also $y=u'-u$. Mit*

$$\mathbb{D}_d = \left\{ y\in\mathbb{R}^{d+1} \;:\; \sum_{j=0}^{d} y_j = 0 \right\}$$

bezeichnen wir die Menge aller Richtungen in $\mathbb{S}_d$.

2. Eine Funktion $f\in C(\mathbb{S}_d)$ heißt richtungskonvex *in Richtung $y\in\mathbb{D}_d$, wenn*

$$t\, f(u+y) + (1-t)\, f(u) \geq f(u+ty), \qquad t\in[0,1],$$

und konvex*, wenn sie richtungskonvex in* alle *Richtungen $y\in\mathbb{D}_d$ ist.*

3. Eine Matrix $A\in\mathbb{R}^{d+1\times d+1}$ heißt bedingt positiv definit[96]*, wenn*

$$y^T A y \geq 0, \qquad y\in\mathbb{D}_d.$$

Übung 4.13 Zeigen Sie: $f\in C(\mathbb{S}_d)$ ist genau dann konvex, wenn für alle $k\geq 0$

$$f\left(\sum_{j=0}^{k} v_j\, u^j\right) \leq \sum_{j=0}^{k} v_j\, f\left(u^j\right), \qquad v\in\mathbb{S}_k, \quad u^0,\ldots,u^k\in\mathbb{S}_d. \tag{4.24}$$

Man bezeichnet (4.24) manchmal auch als *Jensensche Ungleichung*. ♢

Und schon fällt uns ein Teilresultat von Satz 4.19 in den Schoß.

[96] Der englische "Originalbegriff" lautet *conditionally positive definite*.

Satz 4.30 *Ist $f \in C(\mathbb{S}_d)$ konvex, so ist $B_n f \geq B_{n+1} f$.*

Beweis: Für $\alpha \in \Gamma_{n+1}$ ist

$$\begin{aligned}\sum_{j=0}^{d} \frac{\alpha_j}{n+1} \frac{\alpha - e_j}{n} &= \frac{\alpha}{n(n+1)} \underbrace{\left(\sum_{j=0}^{d} \alpha_j \right)}_{=|\alpha|=n+1} - \frac{1}{n(n+1)} \underbrace{\sum_{j=0}^{d} \alpha_j \, e_j}_{=\alpha} = \alpha \underbrace{\left(\frac{1}{n} - \frac{1}{n(n+1)} \right)}_{=\frac{n}{n(n+1)}=\frac{1}{n+1}} \\ &= \frac{\alpha}{n+1}\end{aligned}$$

weswegen

$$B_n f - B_{n+1} f = \sum_{\alpha \in \Gamma_{n+1}} \underbrace{\left(\sum_{j=0}^{d} \frac{\alpha_j}{n+1} f\left(\frac{\alpha - e_j}{n} \right) - f\left(\frac{\alpha}{n+1} \right) \right)}_{\geq 0} B_\alpha \geq 0 \tag{4.25}$$

ist. □

Die Konvexität glatter Funktionen können wir nun recht einfach beschreiben.

Lemma 4.31 *Eine Funktion $f \in C^2(\mathbb{S}_d)$ ist genau dann (strikt) konvex, wenn die* baryzentrische Hesse–Matrix

$$Hf = \left[\frac{\partial^2}{\partial u_j \, \partial u_k} f \ : \ j,k = 0, \ldots, d \right]$$

(strikt) bedingt positiv definit ist.

Beweis: Für $u \in \mathbb{S}_d$ und $y \in \mathbb{D}_d$ betrachten wir die Funktion

$$F_y(t) := f(u + ty), \qquad t \in I \subset \mathbb{R},$$

die zu $C^2(I)$ gehört. Nun ist f genau dann konvex, wenn $F_y''(t) \geq 0$ für alle $y \in \mathbb{D}_d$ ist, also genau dann wenn

$$0 \leq \frac{d^2}{dt^2} f(u + ty) = y^T (Hf)\, y\, (u + ty),$$

was genau die Behauptung ist. □

Und hier haben wir den Salat: In zwei und mehr Variablen ist Konvexität eine *nichtlineare* Eigenschaft der zweiten Ableitung(en) und das muß ganz einfach für Schwierigkeiten sorgen, denn unser wesentliches Hilfsmittel waren ja Satz 4.2 und Lemma 4.4, die uns sagen, daß jede Eigenschaft, die man über Linearkombinationen von Ableitungen beschreiben kann, von Bernsteinpolynomen erhalten wird. Und tatsächlich war die folgende Beobachtung von Schmid [73], siehe [64], eine recht unerfreuliche Überraschung[97]:

[97] Zumal Popoviciu in einer Arbeit behauptet hatte, daß auch in zwei und mehr Variablen Bernsteinpolynome "offensichtlich" Konvexität erhalten würden, was wieder einmal zeigt, daß man einerseits keinem trauen soll und andererseits schon gar nicht, wenn jemand behauptet, daß es "offensichtlich" richtig wäre.

Bernsteinpolynome in zwei Variablen erhalten Konvexität nicht mehr, genauer: Das Bernsteinpolynom zu $|u_1 - u_2|$ ist nicht konvex.

Was also tun? Konvexität ist offenbar nicht die "richtige" Eigenschaft, da eine nichtlineare Eigenschaft; also nehmen wir was anderes, eine Definition, die ebenfalls auf Schmid zurückgeht.

Definition 4.32 *[73, 64, 13, 66, 67]*
Eine Funktion $f \in C(\mathbb{S}_d)$ heißt achsenkonvex[98]*, wenn sie konvex in die Richtungen $y = e_j - e_k$, $j, k = 0, \ldots, d$, ist.*

Und in der Tat sieht Achsenkonvexität eigentlich recht vielversprechend aus.

Satz 4.33 *1. Eine Funktion $f \in C^2(\mathbb{S}_d)$, ist genau dann achsenkonvex, wenn*

$$D^2_{e_j - e_k} f(u) \geq 0, \qquad u \in \mathbb{S}_d.$$

2. Eine Funktion f ist genau dann achsenkonvex, wenn $B_n f$ für alle $n \in \mathbb{N}_0$ achsenkonvex ist.

3. Ist f achsenkonvex, dann ist $B_n f \geq B_{n+1} f$.

Beweis: Für 1) brauchen wir nur zu sehen, daß

$$D_y f(u + ty) = \lim_{h \to h} \frac{f(u + ty + hy) - f(u + ty)}{h} = \frac{d}{dt} f(u + ty),$$

also ist $f \in C^2(\mathbb{S}_d)$ genau dann richtungskonvex in Richtung $y \in \mathbb{D}_d$, wenn

$$0 \leq \frac{d^2}{dt^2} f(u + ty) = D^2_y f(u + ty), \qquad u + ty \in \mathbb{S}_d,$$

und wenn wir $t = 0$ setzen, dann erhalten wir die Behauptung.
Für 2) brauchen wir wieder Ableitungen von Bernsteinpolynomen. Da, für $j = 0, \ldots, d$,

$$\frac{\partial}{\partial u_j} B_\alpha(u) = \frac{|\alpha|!}{\alpha_0! \cdots \alpha_d!} \alpha_j u^{\alpha - e_j} = |\alpha| \binom{|\alpha|}{\alpha - e_j} u^{\alpha - e_j} = |\alpha| B_{\alpha - e_j}(u), \tag{4.26}$$

ist

$$\begin{aligned} D_{e_j - e_k} B_n f &= \sum_{\alpha \in \Gamma_n} f\left(\frac{\alpha}{n}\right) |\alpha| \left(B_{\alpha - e_j} - B_{\alpha - e_k}\right) \\ &= |\alpha| \sum_{\alpha \in \Gamma_{n-1}} \left(f\left(\frac{\alpha + e_j}{n}\right) - f\left(\frac{\alpha + e_k}{n}\right)\right) B_\alpha \\ &=: |\alpha| \sum_{\alpha \in \Gamma_{n-1}} \Delta_{(e_j - e_k)/n} f\left(\frac{\alpha + e_k}{n}\right) B_\alpha, \end{aligned}$$

[98] Der Name kommt daher, daß die Kanten des Simplex auch als "Achsen" bezeichnet werden können, insbesondere, wenn man das Simplex im $\mathbb{R}^d$ so legt, daß $v_0 = 0$ und $v_j = e_j$, $j = 1, \ldots, d$, ist.

wobei

$$\Delta_y f(u) := f(u+y) - f(u), \qquad u \in \mathbb{S}_d, \quad y \in \mathbb{D}_d,$$

also

$$D^2_{e_j - e_k} = |\alpha| \left(|\alpha| - 1\right) \sum_{\alpha \in \Gamma_{n-2}} \Delta^2_{(e_j - e_k)/n} f \left(\frac{\alpha + 2e_k}{n} \right) B_\alpha, \tag{4.27}$$

was ≥ 0 ist, wenn f achsenkonvex ist, siehe Übung 4.14.
Um schließlich 3) zu beweisen, beginnen wir genau so wie beim Beweis von Satz 4.30, das heißt, wir wenden die Graderhöhungsformel auf $B_n f$ an und erhalten (4.25). Nur müssen wir jetzt etwas sorgfältiger argumentieren! Die entscheidende Beobachtung hierbei ist, daß die Simplizes

$$\Delta_\alpha = \left[\frac{\alpha - e_j}{n} \ : \ j = 0, \ldots, d \right], \qquad \alpha \in \Gamma_{n+1}$$

achsenparallel ist, das heißt, alle ihre *eindimensionalen* Kanten

$$\frac{\alpha - e_k}{n} - \frac{\alpha - e_j}{n} = \frac{e_j - e_k}{n}, \qquad j, k = 0, \ldots, d,$$

sind parallel zu den Achsen von $\mathbb{S}_d$. Und für achsenparallele Simplizes gilt eine besondere Variante der Jensen–Ungleichung (4.24), die wir gleich in Lemma 4.34 kennenlernen werden und die den Beweis komplettiert. □

Übung 4.14 Zeigen Sie: $f \in C\left(\mathbb{S}_d\right)$ ist genau dann achsenkonvex, wenn

$$\Delta^2_{h(e_j - e_k)} f\left(u\right) \geq 0, \qquad u + h(e_j - e_k) \in \mathbb{S}_d, \qquad j, k = 0, \ldots, d.$$

♢

Lemma 4.34 *[67]*
Eine Funktion $f \in C\left(\mathbb{S}_d\right)$ ist genau dann achsenkonvex wenn

$$f\left(\sum_{j=0}^{k} v_j \, u^j \right) \leq \sum_{j=0}^{k} v_j \, f\left(u^j\right), \qquad v \in \mathbb{S}_k, \tag{4.28}$$

für alle Punkte $u^0, \ldots, u^k$, die ein achsenparalleles *Simplex aufspannen.*

Bemerkung 4.35 *Gleichung (4.28) erinnert sehr stark*[99] *an die Konvexkombinationen aus Definition 3.12. In der Tat ist eine Funktion* konvex *, wenn (4.28)* für alle *Punkte gilt, die dann natürlich trivialerweise ein Simplex bilden. Achsenkonvexität liegt vor, wenn wir uns dabei auf achsenkonvexe Simplizes beschränken.*

[99] Naja, wenn man bedenkt, daß es sich ja auch hier um eine Konvexitätseigenschaft handelt, dann ist das vielleicht nicht ganz so überraschend.

Beweis: Wir beweisen (4.28) für den Fall, daß $v_j > 0$, $j = 0, \ldots, k$, der Rest folgt aus Gründen der Stetigkeit. Damit sind die Zahlen

$$\lambda_j = \frac{v_j}{1 - v_0 - \cdots - v_{j-1}}, \qquad j = 0, \ldots, k,$$

wohldefiniert, erfüllen

$$\lambda_j > 0 \qquad \text{und} \qquad 1 - \lambda_j = \frac{\overbrace{1 - v_0 - \cdots - v_j}^{= v_{j+1} + \cdots + v_k}}{1 - v_0 - \cdots - v_{j-1}} \geq 0, \qquad \text{also} \qquad 0 < \lambda_j \leq 1,$$

und wir können die Punkte

$$u^{\ell,j}, \qquad j = 0, \ldots, k, \quad \ell = j, \ldots, k,$$

mittels der Rekursion

$$u^{\ell,j+1} := \lambda_j u^{j,j} + (1 - \lambda_j)\, u^{\ell,j}, \tag{4.29}$$

initialisiert mit $u^{\ell,0} = u^\ell$, $\ell = 0, \ldots, k$, einführen. Wegen

$$u^{\ell,j+1} - u^{\ell',j+1} = \lambda_j u^{j,j} + (1 - \lambda_j)\, u^{\ell,j} - \lambda_j u^{j,j} - (1 - \lambda_j)\, u^{\ell',j} = (1 - \lambda_j) \left(u^{\ell,j} - u^{\ell',j} \right)$$

ergibt sich per Induktion, daß jede Kante $u^{\ell,j} - u^{\ell',j}$, $\ell, \ell' = j, \ldots, k$, $j = 0, \ldots, k$, parallel zu einer Achse von $\mathbb{S}_d$ ist und somit ist für $j = 0, \ldots, k-1$ und $\ell = j+1, \ldots, k$

$$f\left(u^{\ell,j+1}\right) = f\left(\lambda_j u^{j,j} + (1 - \lambda_j)\, u^{\ell,j}\right) \leq \lambda_j f\left(u^{j,j}\right) + (1 - \lambda_j)\, f\left(u^{\ell,j}\right). \tag{4.30}$$

Da außerdem für $\ell = 0, \ldots, k$

$$\lambda_\ell \prod_{j=0}^{\ell-1} (1 - \lambda_j) = \frac{v_\ell}{1 - v_0 - \cdots v_{\ell-1}} \prod_{j=0}^{\ell-1} \frac{1 - v_0 - \cdots - v_j}{1 - v_0 - \cdots v_{j-1}} = v_\ell,$$

ist

$$\begin{aligned}
u^{k,k} &= \lambda_{k-1}\, u^{k-1,k-1} + (1 - \lambda_{k-1})\, u^{k,k-1} \\
&= \lambda_{k-1} \left(\lambda_{k-2} u^{k-2,k-2} + (1 - \lambda_{k-2}) u^{k-1,k-2}\right) \\
&\quad + (1 - \lambda_{k-1}) \left(\lambda_{k-2}\, u^{k-2,k-2} + (1 - \lambda_{k-2})\, u^{k,k-2}\right) \\
&= \lambda_{k-2} u^{k-2,k-2} + (1 - \lambda_{k-1})\, \lambda_{k-2}\, u^{k-1,k-2} + (1 - \lambda_{k-1})\, (1 - \lambda_{k-2})\; u^{k,k-2} \\
&\;\;\vdots \\
&= \sum_{\ell=0}^{k} \underbrace{\lambda_\ell \prod_{j=0}^{\ell-1} (1 - \lambda_\ell)}_{= v_\ell} \underbrace{u^{\ell,0}}_{= u^\ell} = \sum_{j=0}^{k} v_j\, u^j = u.
\end{aligned}$$

Mit (4.30) ist schließlich

$$\begin{aligned} f(u) &= f\left(u^{k,k}\right) \geq \lambda_{k-1} f\left(u^{k-1,k-1}\right) + (1-\lambda_{k-1})\ f\left(u^{k,k-1}\right) \\ &\vdots \\ &\geq \sum_{\ell=0}^{k} \lambda_\ell \prod_{j=0}^{\ell-1} (1-\lambda_\ell)\ f\left(u^{\ell,0}\right) = \sum_{j=0}^{k} v_j\, f\left(u^j\right), \end{aligned}$$

also gerade (4.28).
Die Umkehrung folgt trivialerweise aus dem Fall $k = 1$. □

Leider reicht aber Achsenkonvexität trotzdem nicht; es war wieder einmal Schmid, der ein Beispiel für folgende Aussage angab:

Es gibt ein Polynom p, das nicht achsenkonvex ist, aber $B_n p \geq B_{n+1} p$ erfüllt.

Was diese Umkehrung ist, das weiß man zwar[100], man kann's aber (noch) nicht beweisen. Was noch recht einfach ist, sind die sogenannten *Umkehrsätze für Konvexität*.

Satz 4.36 *Erfüllt $f \in C(\mathbb{S}_d)$ die Bedingung $B_n f \geq f$, so nimmt f sein Maximum in einem Eckpunkt von $\mathbb{S}_d$ an.*

Solche Umkehrsätze sind motiviert durch die Tatsache, daß *konvexe* Funktionen auf $\mathbb{S}_d$ diese Eigenschaft haben[101]; ein solches Resultat wurde zuerst in [12] für den *bivariaten* Fall angegeben[102], dann von Dahmen und Micchelli in [14] mit anderen Methoden, sogenannten Operatorhalbgruppen[103], für beliebig viele Variablen behandelt; der vereinfachte Beweis, den wir hier angeben, stammt aus [65].
Beweis von Satz 4.36: Angenommen, f hätte sein Maximum an einer Stelle x^* im Inneren von $\mathbb{S}_d$. Dann ist

$$f(x^*) = f(x^*) \underbrace{\sum_{\alpha\in\Gamma_n} B_\alpha}_{=1} = \sum_{\alpha\in\Gamma_n} \underbrace{f(x^*)}_{\geq f\left(\frac{\alpha}{n}\right)} B_\alpha > B_n f(x^*),$$

da für alle α mit $\alpha_j = 0$ für ein j ja "$>$" gelten muß. Also muß f sein Maximum auf dem Rand annehmen. Mit (4.22) können wir dasselbe Argument auf den Rand anwenden und kommen so in die Ecken. □

Ganz zum Schluß noch schnell die Antwort auf die Frage: "Und was ist nun die Beschreibung der monotonen Konvergenz?".

[100] Es handelt sich um eine Form von Subharmonizität, in "lesbarster" Form sind diese Fakten wohl in [68] zu finden, auch wenn das hier wie Eigenwerbung klingen mag. Vielleicht ist's ja aber auch eine ...

[101] Wie übrigens auch subharmonische Funktionen [61] oder subharmonische Funktionen bezüglich strikt elliptischer Differentialoperatoren zweiter Ordnung, siehe [25].

[102] Im Anhang dieser Arbeit findet sich die Bemerkung, es würde bereits ein Kollege der beiden Autoren den trivariaten Fall behandeln, was technisch aber sehr viel schwieriger sei – wie sowas weitergehen kann, das läßt sich leicht an den Fingern erst einer, dann beider Hände abzählen ...

[103] Eine *Operatorhalbgruppe* ist eine Familie T_t, $t \in \mathbb{R}_+$ von Operatoren, die die Eigenschaft haben, daß sich die Halbgruppenoperation "+" in $\mathbb{R}_+$ auf die Operatoren überträgt, das heißt, daß $T_s T_t = T_{s+t}$, $s, t \in \mathbb{R}_+$ ist. Solche Operatoren besitzen eine Vielzahl von interessanten Eigenschaften, insbesondere auch im Approximationskontext, siehe [11].

1. Der *Voronovskaja–Operator* für multivariate Bernsteinpolynome hat die Form

$$\mathscr{A} = \frac{1}{2}\sum_{j=0}^{d} u_j \, D^2_{u-e_j} = \frac{1}{2}\sum_{i=0}^{m} u_i \frac{\partial^2}{\partial u_i \partial u_i} - \frac{1}{2}\sum_{i,j=0}^{m} u_i u_j \frac{\partial^2}{\partial u_i \partial u_j}$$

und erfüllt

$$\lim_{n\to\infty} n\left(B_n f - f\right) = \mathscr{A} f, \qquad f \in C^2\left(\mathbb{S}_d\right).$$

2. $\mathscr{A}$ ist ein (strikt) elliptischer[104] Differentialoperator zweiter Ordnung im Inneren von $\mathbb{S}_d$, der degeneriert, wenn man sich dem Rand von $\mathbb{S}_d$ nähert, im Inneren dieses Randes [105] aber wieder ein elliptischer Operator in entsprechend weniger Variablen ist.

3. Für jede *offene* Kugel[106] B, deren Abschluß im Inneren von $\mathbb{S}_d$ liegt, ist das *Dirichlet–Problem*

$$\mathscr{A}\phi_B(f)(u) = 0, \quad u \in B, \qquad \phi_B(f)(u) = f(u), \quad u \in \partial B,$$

eindeutig lösbar.

4. Eine Funktion $f \in C\left(\mathbb{S}_d\right)$ heißt *subharmonisch in* $\mathbb{S}_d$, wenn für jede solche Kugel B $f \leq \phi_B(f)$ auf ganz B gilt und *subharmonisch auf* $\mathbb{S}_d$, wenn sie subharmonisch in $\mathbb{S}_\delta$, $\delta \in \{0,1\}^{d+1}$ ist – das ist eine echt stärkere Bedingung.

5. Man kann nun praktisch genau wie im Beweis "5) $\Rightarrow$ 1)" von Satz 4.19 zeigen, daß

$$\limsup_{n\to\infty} n\left(B_n f - f\right)(u) \geq 0, \quad u \in \mathbb{S}_d \qquad \Longrightarrow \qquad f \text{ subharmonisch } \textit{auf } \mathbb{S}_d.$$

6. Die **Vermutung** ist nun naheliegend: Subharmonizität *auf* $\mathbb{S}_d$ ist die gesuchte Verallgemeinerung der Konvexität ...

7. Dies kann man auch tatsächlich für sehr nah verwandte Approximationsoperatoren zeigen, nämlich sogenannte *Bernstein–Durrmeyer–Operatoren*, bei denen anstelle der Punkauswertungsfunktionale Integralmittel verwendet werden:

$$f\left(\frac{\alpha}{n}\right) \to \int_{\mathbb{S}_{[\alpha]}} f(x) B_\alpha(x)\, dx,$$

natürlich mit geeigneter Normierung, so daß die konstante Funktion wieder alle Koeffizienten mit Wert 1 liefert. $\mathbb{S}_{[\alpha]}$ ist dabei die durch die von Null verschiedenen Koeffizienten von α indizierte gegebenefalls niederdimensionale Seite von $\mathbb{S}_d$.

[104] Was auch immer das ist. Nachschlagen kann man's beispielsweise in [25].
[105] Also einige $u_j = 0$, der Rest > 0.
[106] Im intuitiv baryzentrischen Sinne, was auch immer das nun schon wieder ist.

There's probably a smart way to play this, but I just can't think of it at the moment.

R. Chandler, *Trouble is my business*

5 Approximationsordnung

In diesem Kapitel beschäftigen wir uns jetzt endlich mit der *quantitativen Approximationstheorie*, genauer gesagt, mit der Frage, wie schnell die (trigonometrisch) polynomialen Bestapproximationen einer Funktion gegen diese Funktion konvergieren und was die Konvergenzraten über die Funktion aussagen, denn *daß* diese Bestapproximationen konvergieren, das wissen wir ja von den Dichtheitsaussagen, z.B. aus Satz 2.7 (Stone–Weierstraß). Die "harte" Approximationstheorie, die wir in diesem Kapitel betreiben wollen und die ohne technische Abschätzungen leider nicht auskommt, ist auch durch numerische Anwendungen motiviert, wo es immer wieder darum geht, aus der Approximierbarkeit gewisser Funktionen auf die Qualität eines Verfahrens zu schliessen. Doch zuerst einmal ein klein wenig Notation.

Definition 5.1 *Sei*[107] $I = [-1, 1]$*;*

1. *Mit*[108]
$$T_n := \operatorname{span}_{\mathbb{R}} \{1, \cos x, \sin x, \ldots, \cos nx, \sin nx\}$$
bezeichnen wir die trigonometrischen Polynome *vom Grad* $\leq n$ *und mit*
$$\Pi_n := \operatorname{span}_{\mathbb{R}} \{1, x, \ldots, x^n\}$$
die algebraischen Polynome *vom Grad* $\leq n$*.*

2. *Die* Approximationsgüte *von* T_n *bzw.* Π_n *in* $C(\mathbb{T})$ *bzw.* $C(I)$ *bezeichnen wir mit*
$$E_n^*(f) := d(f, T_n) = \inf_{p \in T_n} \|f - p\|_{\mathbb{T}}, \qquad f \in C(\mathbb{T}), \tag{5.1}$$
bzw.
$$E_n(f) := d(f, \Pi_n) = \inf_{p \in \Pi_n} \|f - p\|_I, \qquad f \in C(I). \tag{5.2}$$

[107] Schon wieder ein neues "Standardintervall"! Aber der Grund ist einfach: Um die Resultate für trigonometrische Polynome auf algebraische Polynome übertragen zu können, werden wir die Variablentransformation $x = \cos\theta$ verwenden.

[108] Es sei betont, daß es sich hierbei nicht um den *lateinischen* Buchstaben "T" (großes "t"), sondern um den *griechischen* Buchstaben "T" (großes "τ") handelt!

Bemerkung 5.2 *Die Approximationsgüten $E_n^{(*)}(f)$ sind* subadditiv *und* positiv homogen *bezüglich f: Für $f, g \in C(\mathbb{T})$ und $\lambda \in \mathbb{R}$ ist*

$$E_n^{(*)}(f+g) \le E_n^{(*)}(f) + E_n^{(*)}(g) \qquad \textit{und} \qquad E_n^{(*)}(\lambda f) = |\lambda|\, E_n^{(*)}(f). \tag{5.3}$$

Denn: Ist p Element bester Approximation von f und q Element bester Approximation von g, dann ist

$$E_n^{(*)}(f+g) \le \|f+g-(p+q)\| \le \|f-p\| + \|f-q\| = E_n^{(*)}(f) + E_n^{(*)}(g),$$

und da λp die Bestapproximation zu λf ist[109]*, ist*

$$E_n^{(*)}(\lambda f) = \|\lambda f - \lambda p\| = |\lambda|\ \|f-p\| = |\lambda|\ E_n^{(*)}(f).$$

Da $T_n \subset T_{n+1}$ bzw. $\Pi_n \subset \Pi_{n+1}$ gilt, sind $E_n^*(f)$ und $E_n(f)$ für jedes $f \in C(\mathbb{T})$ bzw. $f \in C(I)$, *monoton fallende* Nullfolgen, letzteres aufgrund der Dichtheitsaussagen. Wir werden uns in diesem Kapitel mit der Frage beschäftigen, für welche Funktionen diese Nullfolgen wie schnell gegen Null konvergieren und, umgekehrt, welche Schlüsse man aus "schneller" Konvergenz ziehen kann.

5.1 Ein Satz von Bernstein

Unsere erste Aussage wird ein Satz von Berstein aus dem Jahre 1938 sein, der uns sagt, daß beliebig gute und beliebig schlechte Approximationsordnung möglich sind.

Satz 5.3 *Zu jeder monoton fallenden Nullfolge ε_n, $n \in \mathbb{N}_0$, gibt es eine* gerade *Funktion $f \in C(\mathbb{T})$, so daß*

$$E_n^*(f) = \varepsilon_n, \qquad n \in \mathbb{N}. \tag{5.4}$$

Der Beweis beruht auf einer Folge von einfachen Beobachtungen, die wir in ein paar Lemmata zusammenstellen wollen. Aus Übung 3.10 wissen wir ja schon, daß eine gerade Funktion auch eine gerade Bestapproximation haben muß, ist also f gerade, dann ist die[110] Bestapproximation t_n^*, definiert durch $E_n^*(f) = \|f - t_n^*\|$, auch eine gerade Funktion, was heißt, daß

$$t_n^*(x) = \frac{a_{n,0}^*}{2} + \sum_{j=1}^{n} a_{n,j}^* \cos jx, \qquad x \in \mathbb{T}. \tag{5.5}$$

Lemma 5.4 *Für $f, g \in C(\mathbb{T})$ ist die Funktion*

$$\psi(\lambda) = E_n^*(f + \lambda g), \qquad \lambda \in \mathbb{R},$$

stetig in λ.

[109] Man denke nur an den Alternatensatz, Satz 3.24.

[110] Wir haben es ja mit einem Haar–Raum zu tun!

Beweis: Mit (5.6) und (5.3) erhalten wir für $\lambda, \lambda' \in \mathbb{R}$, daß

$$\begin{aligned} |\psi(\lambda) - \psi(\lambda')| &= |E_n^*(f + \lambda g) - E_n^*(f + \lambda' g)| \le E_n^*(f + \lambda g - f - \lambda' g) \\ &= |\lambda - \lambda'|\ E_n^*(g), \end{aligned}$$

woraus die Stetigkeit unmittelbar folgt. □

Übung 5.1 Zeigen Sie, daß für $f, g \in C(\mathbb{T})$ die Ungleichung

$$E_n^*(f - g) \ge |E_n^*(f) - E_n^*(g)|, \qquad n \in \mathbb{N}_0, \tag{5.6}$$

gilt.
Hinweis: Beweis von (5.3). ♢

Lemma 5.5 *Sei $f \in C(\mathbb{T})$ gerade.*

1. *Es gibt Konstanten λ_n, $n \in \mathbb{N}_0$, so daß*

$$E_n^*(f + \lambda_{n+1} \cos(n+1)(\cdot)) = E_{n+1}^*(f), \qquad n \in \mathbb{N}_0. \tag{5.7}$$

2. *Zu jedem $\varepsilon \ge E_{n+1}^*(f)$ gibt es ein $\gamma \in \mathbb{R}$, so daß*

$$E_n(f + \gamma \cos(n+1)(\cdot)) = \varepsilon. \tag{5.8}$$

Beweis: Da $E_n^*(f) = E_n^*(f + p)$ für jedes $p \in T_n$, erhalten wir mit t_{n+1}^* aus (5.5), daß

$$\begin{aligned} E_{n+1}^*(f) &= E_{n+1}^*\left(f - a_{n+1,n+1}^* \cos(n+1)(\cdot)\right) \le E_n^*\left(f - a_{n+1,n+1}^* \cos(n+1)(\cdot)\right) \\ &= E_n^*\left(f - t_{n+1}^*\right) \le \left\|f - t_{n+1}^*\right\| = E_{n+1}^*(f), \end{aligned}$$

also ist $\lambda_{n+1} := -a_{n+1,n+1}^*$ die gesuchte Konstante in 1).
Für 2) betrachten wir die Funktion

$$\psi(\lambda) = E_n^*(f + \lambda \cos(n+1)(\cdot)),$$

die, nach Teil 1), $\psi(\lambda_{n+1}) = E_{n+1}^*(f) \le \varepsilon$ erfüllt. Und da $\cos(n+1)x \in T_{n+1} \setminus T_n$, ist

$$\lim_{\lambda \to \infty} \psi(\lambda) = \infty,$$

und wegen der in Lemma 5.4 bewiesenen Stetigkeit von ψ muß auch ein γ geben, an dem ψ den Zwischenwert ε annimmt. □

Nun formulieren und beweisen wir Satz 5.3 für *endliche* Folgen $\varepsilon_0 \ge \cdots \ge \varepsilon_n$ – allerdings gleich noch mit einer "Nebenbedingung".

Proposition 5.6 *Für jede Wahl von $\varepsilon_0 \ge \cdots \ge \varepsilon_n \ge 0$ gibt es ein gerades trigonometrisches Polynom $t \in T_{n+1}$, so daß*

$$E_k^*(t) = \begin{cases} \varepsilon_k, & 0 \le k \le n, \\ 0, & k \ge n+1, \end{cases} \qquad k \in \mathbb{N}_0,$$

und $\|t\| = \varepsilon_0$.

Beweis: Wir wählen $f = 0$ in Teil 2) von Lemma 5.5, dann ist $E_{n+1}(f) = 0$ und es gibt eine Konstante γ_{n+1}, so daß

$$E_n^*\Big(\underbrace{f + \gamma_{n+1}\,\cos(n+1)(\cdot)}_{=:f_n}\Big) = \varepsilon_n.$$

Mit demselben Argument gibt es, für $k = n-1, \dots, 0$, Konstanten γ_{k+1}, so daß

$$\varepsilon_k = E_k^*\Big(\underbrace{f_{k+1} + \gamma_{k+1}\,\cos(k+1)(\cdot)}_{=:f_k}\Big)$$

Da außerdem $f_{k+1} - f_k \in T_{k+1}$ ist zudem für $\ell > k$

$$E_\ell^*\left(f_k\right) = E_\ell^*\Big(f_k + \overbrace{\sum_{j=k}^{\ell-1} \underbrace{f_{j+1} - f_j}_{\in T_{j+1}\subset T_\ell}}^{\in T_\ell}\Big) = E_\ell^*\left(f_\ell\right) = \varepsilon_\ell,$$

weswegen das so konstruierte $f_0 \in T_{n+1}$ die Forderungen

$$E_k^*\left(f_0\right) = \varepsilon_k, \qquad k = 0, \dots, n,$$

erfüllt. Schließlich setzen wir $t = f_0 - t_0^*\left(f_0\right)$, so daß

$$\|t\| = \|f_0 - t_0^*\left(f_0\right)\| = E_0^*\left(f_0\right) = \varepsilon_0$$

ist. □

Und nun haben wir eigentlich fast alle Hilfsmittel für den Beweis von Satz 5.3 beisammen – alle, bis auf eines, das wir aber in Abschnitt 5.4, genauer in Satz 5.13 beweisen werden, nämlich, daß

$$\omega\left(f, \delta\right) \le M\,\delta \sum_{j=0}^{\lfloor \delta^{-1} \rfloor} E_j^*(f), \qquad \delta > 0,\ f \in C(\mathbb{T}), \tag{5.9}$$

wobei die Konstante M von f und δ *unabhängig* ist.

Beweis von Satz 5.3: Für $n \in \mathbb{N}$ konstruieren wir nach Art von Proposition 5.6 trigonometrische Polynome $t_n \in T_{n+1}$, $n \in \mathbb{N}_0$, so daß

$$E_k^*\left(t_n\right) = \varepsilon_k, \quad k = 0, \dots, n, \qquad \text{und} \qquad E_{n+1}^*\left(t_n\right) = E_{n+2}^*\left(t_n\right) = \dots = 0.$$

Da $\|t_n\| = \varepsilon_0$ und da für jedes $\delta > 0$ und $n \in \mathbb{N}_0$

$$\omega\left(t_n, \delta\right) \le M\delta \sum_{j=0}^{\lfloor \delta^{-1} \rfloor} \underbrace{E_j^*\left(t_n\right)}_{\in\{\varepsilon_j, 0\}} \le M\,\delta \underbrace{\sum_{j=0}^{\lfloor \delta^{-1} \rfloor} \varepsilon_j}_{\to 0},$$

ist die Folge t_n gleichmäßig beschränkt und *gleichgradig stetig*[111], daher enthält sie nach dem Satz von Arzela–Ascoli[112]eine Teilfolge, die gleichmäßig gegen das gesuchte $f \in C(\mathbb{T})$ konvergiert. □

[111] Auch als *"gleichstetig"* bezeichnet.

[112] "Grundwissen Analysis"? Siehe beispielsweise [30, Satz 106.2, S. 563].

5.2 Trigonometrische Polynome I: Stetige Funktionen

Da also die Approximationsgüte $E_n^*(f)$ für $n \to \infty$ beliebig schnell und beliebig langsam gegen Null konvergieren kann, ist es jetzt an der Zeit, sich mal mit der Frage zu beschäftigen, wie man Funktionen an der Konvergenzrate von E_n^* erkennen kann und umgekehrt.

Definition 5.7 *Eine Funktion $f \in C(\mathbb{T})$ ist* lipschitzstetig[113] *von der Ordnung $0 < \alpha < 1$, wenn*

$$\sup_{x \neq x' \in \mathbb{T}} \frac{|f(x) - f(x')|}{|x - x'|^\alpha} < \infty,$$

was wir auch als $f \in C^\alpha(\mathbb{T})$ schreiben werden.

Wir können Lipschitzstetigkeit auch durch den Stetigkeitsmodul ausdrücken und erhalten dann, daß für $0 < \alpha < 1$

$$f \in C^\alpha(\mathbb{T}) \qquad \Longleftrightarrow \qquad \sup_{\delta > 0} \delta^{-\alpha}\, \omega(f, \delta) < \infty. \tag{5.10}$$

Und in der Tat ist es in erster Linie mal Lipschitzstetigkeit, die für die Konvergenzrate der Approximationsgüte verantwortlich ist.

Satz 5.8 *Für $0 < \alpha < 1$ und $f \in C(\mathbb{T})$ ist*

$$f \in C^\alpha(\mathbb{T}) \qquad \Longleftrightarrow \qquad \sup_{n \in \mathbb{N}} n^\alpha\, E_n^*(f) < \infty. \tag{5.11}$$

Wir werden den Beweis in zwei Teilen anpacken, indem wir zuerst in 5.3 die Richtung "$\Rightarrow$" zeigen, die sogenannten *Jackson–Sätze*, und uns dann in 5.4 die Richtung "$\Leftarrow$", auch als *Bernstein–Sätze* bekannt, vorknüpfen.

5.3 Trigonometrische Polynome II: Jackson–Sätze

Legen wir also los! Das erste Resultat, die Beschränkung der Approximationsgüte durch den Stetigkeitsmodul, geht auf Jackson[114] [32] zurück, siehe auch [33].

Satz 5.9 *(Jackson–Satz)*
Es gibt eine Konstante $M > 0$, so daß für alle $f \in C(\mathbb{T})$ und $n \in \mathbb{N}_0$ die Ungleichung

$$E_n^*(f) \leq M\, \omega\left(f, \frac{1}{n}\right) \tag{5.12}$$

gilt.

[113] Rudolf Lipschitz, 1832–1903, Beiträge zu Fourierreihen, algebraischer Zahlentheorie, partiellen Differentialgleichungen und Potentialtheorie.

[114] Dunham Jackson, 1888-1946, die Resultate enstanden im Rahmen seiner von Landau betreuten Dissertation. Schrieb eines der ersten Bücher über Approximationstheorie, [33].

Das entscheidende Hilfsmittel zum (überraschend einfachen) Beweis von Satz 5.9 ist wie in Abschnitt 1.2 ein *Faltungsoperator*, genauer, sogar eine Variation des *Fejér–Kerns* aus (1.13): Für $n \in \mathbb{N}$ ist der *n–te Jackson–Kern* definiert als

$$J_n(x) := \mu_n' \, F_{n-1}^2(x) = \mu_n \left(\frac{\sin \frac{n}{2} x}{\sin \frac{1}{2} x} \right)^4, \qquad \mu_n > 0, \quad x \in \mathbb{T}, \tag{5.13}$$

wobei μ_n so gewählt werden soll, daß $\int_{\mathbb{T}} J_n(t)\, dt = 1$. Ein Plot der Jackson–Kerne ist nicht besonders aufregend: Er hat eigentlich nur einen "Peek" an der Stelle $x = 0$, der "oszillierende" Teil wird praktisch unsichtbar.

Lemma 5.10 *Es gibt Konstanten $M_k > 1$, $k = 0, 1, 2$, so daß*

$$M_k^{-1} n^{-k} \le \int_0^\pi t^k \, J_n(t)\, dt \le M_k \, n^{-k}, \qquad k = 0, 1, 2. \tag{5.14}$$

Beweis: Da die Funktion $\sin x/2$ auf $[0, \pi]$ konkav ist, ist für $x \in [0, \pi]$

$$\frac{x}{\pi} = \frac{\pi - x}{\pi} \sin 0 + \frac{x}{\pi} \underbrace{\sin \frac{\pi}{2}}_{=1} \le \sin \frac{x}{2} = \int_0^{x/2} \underbrace{\cos t}_{\le 1} \, dt \le \frac{x}{2},$$

also

$$\frac{x}{\pi} \le \sin \frac{x}{2} \le \frac{x}{2}, \qquad x \in [0, \pi]. \tag{5.15}$$

Somit ist für $n \in \mathbb{N}$

$$\begin{aligned} \int_{\mathbb{T}} \left(\frac{\sin \frac{n}{2} t}{\sin \frac{1}{2} t} \right)^4 dt &= 2 \int_0^\pi \left(\frac{\sin \frac{n}{2} t}{\sin \frac{1}{2} t} \right)^4 dt \le 2 \int_0^\pi \left(\frac{\sin \frac{n}{2} t}{\frac{1}{\pi} t} \right)^4 dt = \frac{4}{n} \pi^4 \int_0^{n\pi/2} \left(\frac{\sin t}{\frac{2}{n} t} \right)^4 dt \\ &= \frac{n^3 \pi^4}{4} \int_0^{n\pi/2} \left(\frac{\sin t}{t} \right)^4 dt \le \frac{n^3 \pi^4}{4} \underbrace{\int_0^\infty \left(\frac{\sin t}{t} \right)^4 dt}_{=: M < \infty} \end{aligned}$$

und ganz analog

$$\int_{\mathbb{T}} \left(\frac{\sin \frac{n}{2} t}{\sin \frac{1}{2} t} \right)^4 dt \ge 4\, n^3 \underbrace{\int_0^{\pi/2} \left(\frac{\sin t}{t} \right)^4 dt}_{=: m > 0},$$

das heißt

$$m \, 4n^3 \le \mu_n^{-1} \le M \, \frac{n^3 \pi^4}{4}, \qquad n \in \mathbb{N}. \tag{5.16}$$

Somit ist für $k = 0, 1, 2$ und $n \in \mathbb{N}$

$$\begin{aligned} \int_{\mathbb{T}} t^k \, J_n(t)\, dt &= \mu_n \int_{\mathbb{T}} t^k \left(\frac{\sin \frac{n}{2} t}{\sin \frac{1}{2} t} \right)^4 dt \le 2\pi^4 \mu_n \int_0^\pi t^k \left(\frac{\sin \frac{n}{2} t}{t} \right)^4 dt \\ &= 2\pi^4 \mu_n \int_0^\pi \frac{\left(\sin \frac{n}{2} t \right)^4}{t^{4-k}} \le \frac{4\pi^4}{n} \, \frac{1}{4mn^3} \left(\frac{n}{2} \right)^{4-k} \int_0^{n\pi/2} \frac{\sin^4 t}{t^{4-k}} \, dt \\ &\le n^{-k} \underbrace{\frac{2^{k-4} \pi^4}{m} \int_0^\infty \frac{\sin^4 t}{t^{4-k}} \, dt}_{< \infty} \end{aligned}$$

sowie

$$\int_{\mathbb{T}} t^k \, J_n(t) \, dt \geq n^{-k} \underbrace{\frac{2^k}{M} \int_0^\infty \frac{\sin^4 t}{t^{4-k}} \, dt}_{>0},$$

woraus (5.14) unmittelbar folgt. □

Ein kleiner "Fehler" des Operators J_n ist, daß er ein trigonometrisches Polynom vom Grad $2n-2$ ist und somit auch liefert, siehe Übung 5.2, und wir hätten halt nun doch gerne ein trigonometrisches Polynom vom Grad $\leq n$. Also setzen wir $K_n = J_{\lfloor n/2 \rfloor + 1}$, $n \in \mathbb{N}$ und definieren den *Jackson–Operator* J_n als

$$J_n f := K_n * f, \qquad n \in \mathbb{N}. \tag{5.17}$$

Und obwohl die Bestapproximation *nicht* linear von f abhängt, siehe Übung 5.3 hat trotzdem dieser *lineare Faltungsoperator* bereits dieselbe Approximationsgüte.

Übung 5.2 Zeigen Sie, daß $J_n \in T_{2n-2}$ ist. ♢

Übung 5.3 Seien $t_n^*(f) \in T_n$, $n \in \mathbb{N}_0$, die trigonometrische Polynome bester Approximation zu $f \in C(\mathbb{T})$. Zeigen Sie, daß die Operatoren $T_n : f \mapsto t_n^*(f)$, $n \in \mathbb{N}_0$, *nicht* linear sind. ♢

Schließlich brauchen wir noch eine einfache Aussage über Stetigkeitsmodule.

Lemma 5.11 *Für $f \in C(\mathbb{T})$, $\delta > 0$ und $\lambda \in \mathbb{R}_+$ ist*

$$\omega(f, \lambda\, \delta) \leq (\lfloor \lambda \rfloor + 1)\ \omega(f, \delta)\,. \tag{5.18}$$

Beweis: Für $\lambda \in \mathbb{N}$ und $|h| \leq \delta$ ist

$$|\Delta_{\lambda h} f(x)| = \left| \sum_{j=0}^{\lambda-1} \Delta_h f\,(x + jh) \right| \leq \sum_{j=0}^{\lambda-1} |\Delta_h f\,(x+jh)| \leq \lambda\, \omega(f, \delta)$$

und Übergang zum Supremum liefert (5.18) für $\lambda \in \mathbb{N}$; für beliebiges $\lambda \in \mathbb{R}_+$ folgt das Ganze wegen $\lambda \leq \lfloor \lambda \rfloor + 1$ und der Monotonie des Steigkeitsmoduls. □

Beweis von Satz 5.9: Für $x \in \mathbb{T}$ und $n \in \mathbb{N}$ ist, wegen der Normierung von K_n und da K_n gerade ist,

$$\begin{aligned}
f(x) - J_n f(x) &= f(x) \underbrace{(K_n * 1)}_{=1} - (K_n * f)\,(x) = \int_{\mathbb{T}} (f(x) - f(x-t))\ K_n(t)\, dt \\
&= \int_{-\pi}^{0} (f(x) - f(x-t))\, K_n(t)\, dt + \int_0^\pi (f(x) - f(x-t))\, K_n(t)\, dt \\
&= \int_0^\pi (f(x) - f(x+t)) \underbrace{K_n(-t)}_{=K_n(t)}\, dt + \int_0^\pi (f(x) - f(x-t))\, K_n(t)\, dt \\
&= -\int_0^\pi (f(x+t) - 2f(x) + f(x-t))\, K_n(t)\, dt \\
&= -\int_0^\pi (\Delta_t f(x) + \Delta_{-t} f(x))\, K_n(t)\, dt.
\end{aligned}$$

Unter Verwendung von Lemma 5.11 und Lemma 5.10 ergibt sich dann, daß für $x \in \mathbb{T}$

$$\begin{aligned}
|f(x) - J_n f(x)| &\le \int_0^\pi |\Delta_t f(x) + \Delta_{-t} f(x)| \underbrace{K_n(t)}_{\ge 0}\, dt \le 2 \int_0^\pi \underbrace{\omega\,(f,t)}_{=\omega\left(f,nt\frac{1}{n}\right)}\, K_n(t)\, dt \\
&\le 2 \int_0^\pi \underbrace{(nt+1)}_{\ge \lfloor nt \rfloor + 1} \omega\left(f, \frac{1}{n}\right) K_n(t)\, dt = 2\,\omega\left(f, \frac{1}{n}\right) \left(n \int_0^\pi t\, K_n(t)\, dt + \int_0^\pi K_n(t)\, dt \right) \\
&\le 2\,\omega\left(f, \frac{1}{n}\right) \left(\frac{n}{\lfloor n/2 \rfloor + 1} M_1 + 1 \right) \le \underbrace{2\,(2M_1+1)}_{=:M}\, \omega\left(f, \frac{1}{n}\right),
\end{aligned}$$

womit (5.8) bewiesen ist. □

Übung 5.4 Bestimmen Sie eine *explizite* obere Schranke für die Konstante M aus (5.8). ♢

Soweit also zum Jackson–Satz! Wir könnten natürlich auch einen expliziten Wert für die Konstante M angeben, der zwar asymptotisch ziemlich irrelevant wäre, aber doch irgendwie von Interesse ist. Und in der Tat war die Suche nach möglichst guten, wenn nicht sogar *besten* Konstanten in (5.8) durchaus von Interesse. Eine "scharfe" Version des Jackson–Satzes wurde von Korneĭčuk um 1962 angegeben, wahrscheinlich in [40][115], siehe auch [42]. Was man sieht, ist daß $\frac{1}{n}$ eigentlich die "falsche" Schrittweite für den Stetigkeitsmodul war: Um zu exakten Konstanten zu kommen, mußte die Schrittweite in den dem Torus $\mathbb{T}$ "angemesseneren" Wert $\frac{\pi}{n}$ abgeändert werden.

Satz 5.12 *Für $f \in C(\mathbb{T})$ und $n \in \mathbb{N}$ ist*

$$E_n^*(f) \le \omega\left(f, \frac{\pi}{n+1}\right) \qquad \textit{und} \qquad E_n^*(f) \le \frac{1}{2}\,\overline{\omega}\left(f, \frac{\pi}{n+1}\right)$$

wobei $\overline{\omega}\,(f,\cdot)$ die konkave Majorante[116] *von $\omega(f,\cdot)$ ist. Die Konstanten 1 bzw $\frac{1}{2}$ können nicht verbessert werden.*

Bleibt noch unser "Fernziel", nämlich Satz 5.8; eine Hälfte davon können wir nun auch tatsächlich schon beweisen.

Beweis von Satz 5.8, $\Rightarrow$: Sei $f \in C^\alpha(\mathbb{T})$, d.h., es ist $n^\alpha \omega\,(f, n^{-1})$ unabhängig von n beschränkt. Somit ist nach Satz 5.9

$$\sup_{n\in\mathbb{N}} n^\alpha E_n^*(f) \le \sup_{n\in\mathbb{N}} n^\alpha M\, \omega\left(f, \frac{1}{n}\right) = M \sup_{n\in\mathbb{N}} n^\alpha\, \omega\left(f, \frac{1}{n}\right) < \infty.$$

□

[115] Zur Beachtung: Wie [39, 41] ist diese Arbeit in den *"Doklady"* erschienen, wo meist nur die Resultate *ohne* Beweis angekündigt werden, was gut für Prioritäten, aber schlecht für's Verständnis ist. Außerdem ist so mancher Beweis eines "Doklady–Resultats" nie wirklich erschienen, was auch immer das bedeuten mag ...

[116] Das ist die kleinste *konkave* Funktion (in δ), die $\ge \omega(f,\cdot)$ ist.

5.4 Trigonometrische Polynome III: Bernstein–Sätze

Jetzt aber endlich zur Umkehrung. Das Gegenstück zu dem Jackson–Satz, der sogenannte *Bernstein–Satz* schätzt jetzt den Stetigkeitsmodul über die Approximationsgüte ab, allerdings erhalten wir keine Umkehrungen der Form $\omega\left(f, \frac{1}{n}\right) \leq M E_n^*(f)$, die man auch gerne als *"Umkehrsätze vom starken Typ"* bezeichnet, sondern nur "schwache" Umkehrungen – die aber für unsere Zwecke immer noch stark genug sein werden.

Satz 5.13 *(Bernstein–Satz)*
Es gibt eine Konstante $M > 0$, so daß für $f \in C(\mathbb{T})$ und $\delta > 0$

$$\omega(f, \delta) \leq M\,\delta \sum_{n=0}^{\lfloor \delta^{-1} \rfloor} E_n^*(f) \tag{5.19}$$

ist.

Auch für Satz 5.13 brauchen wir Hilfsmittel. Wir beginnen mit einem Resultat, das die Norm der Ableitung eines trigonometrischen Polynoms mit der Norm des Polynoms verknüpft.

Proposition 5.14 *Für $p \in T_n$ gilt die* Bernsteinsche Ungleichung

$$\|p'\| \leq n \|p\|. \tag{5.20}$$

Beweis: Wäre (5.20) falsch, dann gäbe es ein $p \in T_n$ und $x^* \in \mathbb{T}$ mit

$$|p'(x^*)| = \|p'\| = nM > n\|p\|, \qquad M > \|p\|,$$

und wir können annehmen, daß $p'(x^*) = nM$ ist. Weil p' an x^* ein *Maximum* hat, muß $p''(x^*) = 0$ sein. Die Funktion

$$q(x) = M \sin n(x - x^*) - p(x), \qquad x \in \mathbb{T},$$

hat nun die Eigenschaft, daß für $k = 0, \ldots, 2n-1$

$$q(x_k) := q\left(x^* + \frac{2k+1}{2n}\pi\right) = M \underbrace{\sin\left(n\,\frac{2k+1}{2n}\pi\right)}_{=\sin\left(k+\frac{1}{2}\right)\pi = (-1)^k} - p(x_k) = (-1)^k \mu_k,$$

wobei $\mu_k > 0$ ist, also hat q mindestens $2n$ Nullstellen[117]. Nach dem Satz von Rolle hat auch q' mindestens $2n$ Nullstellen und da

$$q'(x^*) = nM \cos n(x^* - x^*) - p'(x^*) = nM - nM = 0$$

[117] Rund um den Einheitskreis!

ist, ist x^* eine davon. Nochmalige Anwendung des Satzes von Rolle ergibt dann, daß q'' ebenfalls mindestens $2n$ Nullstellen *zwischen* den Nullstellen von q' hat, also mindestens $2n$ Nullstellen *verschieden* von x^* und da, nach obiger Überlegung

$$q''(x^*) = -n^2 M \underbrace{\sin n(x^* - x^*)}_{=0} - \underbrace{p''(x^*)}_{=0} = 0$$

ist, hat q'' mindestens $2n+1$ Nullstellen, weswegen q konstant sein müßte – aber eine konstante Funktion mit so vielen echten Vorzeichenwechseln, die muß erst noch erfunden werden. Also haben wir den gewünschten Widerspruch. □

Übung 5.5 Beweisen Sie, daß die Ungleichung (5.20) *scharf* ist, das heißt, es gibt (mindestens) ein p, so daß Gleichheit angenommen wird. ♢

Übung 5.6 Zeigen Sie: Ist p ein trigonometrischen Polynom und ist $p^{(k)} = 0$ für ein $k \in \mathbb{N}$, dann ist p konstant. ♢

Um Satz 5.13 in allgemeinerer Form beweisen zu können, brauchen wir noch etwas mehr Terminologie, nämlich Stetigkeitsmodule *höherer Ordnung*.

Definition 5.15 *Zu $f \in C(\mathbb{T})$ und $r \in \mathbb{N}$ ist der r–te* Glättemodul $\omega_r(f, \delta)$ *definiert als*

$$\omega_r(f, \delta) := \sup_{0<h\leq\delta} \|\Delta_h^r f\|, \tag{5.21}$$

wobei natürlich $\omega = \omega_1$ gilt.

Lemma 5.16 *(Einfache Eigenschaften des Glättemoduls)*

1. *Für $f \in C(\mathbb{T})$ und $r \in \mathbb{N}$ ist*

$$\omega_r(f, \delta) \leq 2^{r-1} \omega(f, \delta). \tag{5.22}$$

2. *Für $f \in C^r(\mathbb{T})$ ist*

$$\omega_r(f, \delta) \leq \delta^r \left\| f^{(r)} \right\|. \tag{5.23}$$

Beweis: Da

$$\|\Delta_h^r f\| = \left\|\Delta_h^{r-1} f(\cdot + h) - \Delta_h^{r-1} f\right\| \leq \underbrace{\left\|\Delta_h^{r-1} f(\cdot + h)\right\|}_{=\left\|\Delta_h^{r-1} f\right\|} + \left\|\Delta_h^{r-1} f\right\| = 2\left\|\Delta_h^{r-1} f\right\|,$$

folgt (5.22) ganz einfach induktiv und (5.23) ist eine unmittelbare Konsequenz von (4.4). □

Nun aber zu unserem zentralen Resultat dieses Abschnitts, aus dem Satz 5.13 als Spezialfall $r = 1$ unmittelbar folgt.

Satz 5.17 *(Bernstein–Satz für Glättemodule) Es gibt Konstanten M_r, $r \in \mathbb{N}$, so daß für jedes $f \in C(\mathbb{T})$ und $\delta > 0$*

$$\omega_r(f,\delta) \le M_r\,\delta^r \sum_{n=0}^{\lfloor \delta^{-1} \rfloor} (n+1)^{r-1}\,E_n^*(f). \tag{5.24}$$

Beweis: Für $n \in \mathbb{N}$ sei $t_n \in T_n$ das jeweilige Polynom bester Approximation an f. Für $\delta \ge 1$ wird (5.24) zu

$$\omega_r(f,\delta) \le M_r\,\delta^r\,E_0^*(f),$$

was wegen

$$\omega_r(f,\delta) = \omega_r(f-t_0,\delta) \le 2^r\,\|f-t_0\| = 2^r\,E_0^*(f)$$

mit $M_r \ge 2^r$ immer erfüllt werden kann. Interessant wird's also für $0 < \delta \le 1$. Wegen der Linearität der Differenzen erhalten wir, daß für $x \in \mathbb{T}$, $0 < h \le \delta$ und $k \in \mathbb{N}_0$

$$\begin{aligned} |\Delta_h^r f(x)| &= |\Delta_h^r\,(f-t_{2^k})\,(x) + \Delta_h^r t_{2^k}(x)| \le |\Delta_h^r\,(f-t_{2^k})\,(x)| + |\Delta_h^r t_{2^k}(x)| \\ &\le h^r \left\|t_{2^k}^{(r)}\right\| + 2^r \underbrace{\|f-t_{2^k}\|}_{=E_{2^k}^*(f)} = h^r \left\|t_{2^k}^{(r)}\right\| + 2^r\,E_{2^k}^*(f). \end{aligned} \tag{5.25}$$

Wir müssen jetzt den Ausdruck $\left\|t_{2^k}^{(r)}\right\|$ abschätzen. Dazu verwenden wir die Abkürzung $E(n) := E_n^*(f)$. Unter Verwendung der Bernstein–Ungleichung (5.20) erhalten wir jetzt für $x \in \mathbb{T}$, daß

$$\begin{aligned} \left\|t_{2^k}^{(r)}\right\| &= \left\|t_{2^k}^{(r)} - \underbrace{t_0^{(r)}}_{=0}\right\| = \left\|t_1^{(r)} - t_0^{(r)} + \sum_{j=1}^{k}\left(t_{2^j}^{(r)} - t_{2^{j-1}}^{(r)}\right)\right\| \\ &\le \left\|t_1^{(r)} - t_0^{(r)}\right\| + \sum_{j=1}^{k}\left\|t_{2^j}^{(r)} - t_{2^{j-1}}^{(r)}\right\| \le 1^r\,\|t_1 - t_0\| + \sum_{j=1}^{k} 2^{jr}\,\|t_{2^j} - t_{2^{j-1}}\| \\ &= \|t_1 - f + f - t_0\| + \sum_{j=1}^{k} 2^{jr}\,\|t_{2^j} - f + f - t_{2^{j-1}}\| \\ &\le \underbrace{E(0)+E(1)}_{\le 2E(0)} + \sum_{j=1}^{k} 2^{jr} \underbrace{\left(E\left(2^j\right) + E\left(2^{j-1}\right)\right)}_{\le 2E(2^{j-1})} \le 2\left(E(0) + 2^r \sum_{j=0}^{k-1} 2^{jr}\,E\left(2^j\right)\right). \end{aligned}$$

Nun ist aber

$$\begin{aligned} \sum_{j=0}^{k-1} 2^{rj}\,E\left(2^j\right) &= 2^r \sum_{j=0}^{k-1} 2^{(r-1)(j-1)} E\left(2^j\right) \underbrace{\left(2^j - 2^{j-1}\right)}_{=2^{j-1}} \le 2^r \sum_{j=0}^{k-1} \int_{2^{j-1}}^{2^j} \underbrace{t^{r-1}}_{\ge 2^{(r-1)(j-1)}} \underbrace{E(t)}_{\ge E(2^j)}\,dt \\ &= 2^r \int_1^{2^{k-1}} t^{r-1}\,E(t)\,dt = 2^r \sum_{j=0}^{2^{k-1}-1} \int_j^{j+1} \underbrace{t^{r-1}}_{\le (j+1)^{r-1}} \underbrace{E(t)}_{\le E(j)}\,dt \\ &\le 2^r \sum_{j=1}^{2^k} (j+1)^{r-1}\,E(j). \end{aligned}$$

Der langen Rechnung kurzer Sinn: Es ist

$$\left\| t_{2^k}^{(r)} \right\| \leq 2^{2r+1} \sum_{j=0}^{2^k} (j+1)^{r-1} \; E(j), \qquad k \in \mathbb{N}_0. \tag{5.26}$$

Außerdem ist aber auch

$$\sum_{j=0}^{2^k} (j+1)^{r-1} \; E(j) \geq E\left(2^k\right) \sum_{j=0}^{2^k} \int_j^{j+1} t^{r-1}\, dt = E\left(2^k\right) \int_0^{2^k} t^{r-1}\, dt = E\left(2^k\right) \frac{2^{rk}}{r},$$

also

$$2^r E\left(2^k\right) \leq r 2^r \, 2^{-rk} \sum_{j=0}^{2^k} (j+1)^{r-1} \; E(j). \tag{5.27}$$

Setzen wir nun (5.26) und (5.27) in (5.25) ein, dann erhalten wir, daß

$$|\Delta_h^r f(x)| \leq \underbrace{2^r \, (2+r)}_{=:M_r/2} \left(h^r + 2^{-rk}\right) \sum_{j=0}^{2^k} (j+1)^{r-1} \; E(j),$$

und wenn wir k so wählen[118], daß $2^k \leq h^{-1} < 2^{k+1}$, dann erhalten wir tatsächlich (5.24). □

Beweis von Satz 5.8, "⇐": Sei also nun $C > 0$ eine Konstante, so daß $n^\alpha E_n^*(f) < C$. Dann ist, nach Satz 5.13

$$\begin{aligned}
\omega\left(f, \delta\right) &\leq M\, \delta \sum_{n=0}^{\lfloor \delta^{-1} \rfloor} E_n^*(f) = M \left(\delta\, E_0^*(f) + \delta \sum_{n=1}^{\lfloor \delta^{-1} \rfloor} E_n^*(f) \right) \\
&\leq M \left(\delta\, E_0^*(f) + C \sum_{n=1}^{\lfloor \delta^{-1} \rfloor} n^{-\alpha} \right) \leq M\delta\, E_0^*(f) + M\delta\, C \sum_{n=1}^{\lfloor \delta^{-1} \rfloor} \int_{n-1}^{n} t^{-\alpha}\, dt \\
&\leq M\delta\, E_0^*(f) + MC\delta\, \frac{\left(\delta^{-1}+1\right)^{1-\alpha} - 1}{1-\alpha} = M\delta\, E_0^*(f) + MC\, \delta^\alpha \frac{(1+\delta)^{1-\alpha} - \delta^{1-\alpha}}{1-\alpha} \\
&= \delta^\alpha \underbrace{\left(M\delta^{1-\alpha} E_0^*(f) + MC\, \frac{(1+\delta)^{1-\alpha} - \delta^{1-\alpha}}{1-\alpha} \right)}_{=:M' < \infty} = M'\, \delta^\alpha,
\end{aligned}$$

also ist $f \in C^\alpha(\mathbb{T})$. □

[118] Hier brauchen wir $\delta \leq 1$!

5.5 Trigonometrische Polynome IV: Differenzierbare Funktionen

Bisher haben wir die Approximationsgüte stetiger Funktionen auf $\mathbb{T}$ durch trigonometrische Polynome über eine *Glattheitseigenschaft* der Funktionen beschrieben – allerdings nur für Exponenten *echt* zwischen 0 und 1. In diesem Abschnitt übertragen wir dies nun auf *alle* nichtganzzahligen rellen Exponenten und erhalten so die folgende Aussage.

Satz 5.18 *Für $k \in \mathbb{N}_0$, $0 < \alpha < 1$ und $f \in C(\mathbb{T})$ ist*

$$f \in C^{k+\alpha}(\mathbb{T}) \qquad \Longleftrightarrow \qquad \sup_{n\in\mathbb{N}} n^{k+\alpha}\, E_n^*(f) < \infty. \tag{5.28}$$

Eine Richtung von Satz 5.18 folgt wieder recht unmittelbar aus der folgenden Verallgemeinerung des Jackson–Satzes, Satz 5.9.

Proposition 5.19 *(Jackson–Satz für differenzierbare Funktionen)*
Für jedes $k \in \mathbb{N}$ gibt es eine Konstante M_k, so daß für alle $f \in C^k(\mathbb{T})$ und $n \in \mathbb{N}_0$ die Ungleichung

$$E_n^*(f) \le M_k\, n^{-k}\, \omega\left(f^{(k)}, \frac{1}{n}\right) \tag{5.29}$$

gilt.

Beweis: Wir werden zeigen, daß es eine Konstante M gibt, mit der die Ungleichung

$$E_n^*(f) \le M\, n^{-1}\, E_n^*\left(f'\right), \qquad f \in C^1(\mathbb{T}), \tag{5.30}$$

gilt, woraus durch Iteration und Satz 5.9 für $n \ge k$

$$\begin{aligned} E_n^*(f) &\le M\, n^{-1} E_n^*\left(f'\right) \le M^2\, n^{-2} E_n^*\left(f''\right) \le \cdots \le M^k\, n^{-k} E_n^*\left(f^{(k)}\right) \\ &\le \underbrace{M^k\, \widetilde{M}}_{=:M_k}\, n^{-k}\, \omega\left(f^{(k)}, \frac{1}{n}\right) \end{aligned}$$

folgt. Und auch (5.30) ist eine eher einfache[119] Angelegenkeit! Dazu sei $p \in T_n$ definiert als

$$f' * K_n(x) =: p(x) =: \frac{p_0}{2} + \sum_{k=1}^{n} p_k\, \cos kx + p'_k\, \sin kx, \qquad x \in \mathbb{T},$$

wobei K_n wieder den *Jackson–Kern* aus (5.13) bezeichnet. Wegen der Normierung $1 * K_n = 1$, $n \in \mathbb{N}$, ist[120]

$$\begin{aligned} \pi\, p_0 &= \int_{\mathbb{T}} p(t)\, dt = \int_{\mathbb{T}} f' * K_n(t)\; dt = \int_{\mathbb{T}} \int_{\mathbb{T}} f'(s)\, K_n\left(t-s\right)\; ds\, dt \\ &= \int_{\mathbb{T}} \int_{\mathbb{T}} f'(s)\, K_n(t)\; dt\, ds = \underbrace{\int_{\mathbb{T}} f'(s)\; ds}_{=f(\pi)-f(-\pi)=0}\; \underbrace{\int_{\mathbb{T}} K_n(t)\; dt}_{=1*K_n=1} = 0. \end{aligned}$$

[119] Aber bekanntlich ist ja alles relativ ...

[120] Wenn man genau hinschaut, dann verbirgt sich in dieser Rechnung das "Prinzip" daß die Fouriertransformierte der Faltung zweier Funktionen das Produkt der individuellen Fouriertransformationen ist (siehe Satz 6.13) – das wiederum ist von kaum zu überschätzender Bedeutung in der Signalverarbeitung.

Außerdem gibt es nach Satz 5.9[121] und Satz 5.13[122] Konstanten M und M', so daß

$$\|f' - p\| = \|f' - J_n f'\| \leq M\omega\left(f', \frac{1}{n}\right) \leq M\, M' E_n^*(f')\,. \tag{5.31}$$

Nun definieren wir $q \in T_n$ durch

$$q(x) := \sum_{k=1}^{n} \frac{p_k}{k}\, \sin kx - \frac{p_k'}{k}\, \cos kx, \qquad x \in \mathbb{T},$$

also $p = q'$. Dann ist, mit Satz 5.9 und (5.23) aus Lemma 5.16, sowie unter Verwendung von (5.31)

$$\begin{aligned} E_n(f) &= E_n(f-q) \leq M\,\omega\left(f-q, \frac{1}{n}\right) \leq M\, n^{-1}\, \|f' - q'\| = M\, n^{-1}\, \|f' - p\| \\ &= M'\, n^{-1}\, E_n(f')\,, \end{aligned}$$

wobei M' eine geeignete Konstante ist, und das ist gerade (5.30), womit der Beweis komplett ist. □

Die Umkehrung basiert auf dem folgenden Resultat.

Proposition 5.20 *Ist für ein $k \in \mathbb{N}$*

$$\sum_{n=0}^{\infty} n^{k-1}\, E_n^*(f) < \infty, \tag{5.32}$$

dann ist $f \in C^k(\mathbb{T})$ und es gibt eine Konstante $M > 0$, so daß

$$E_n^*\left(f^{(k)}\right) \leq M \sum_{\ell=n}^{\infty} \ell^{k-1}\, E_\ell^*(f), \qquad n \in \mathbb{N}. \tag{5.33}$$

Beweis: Sei $t_n \in T_n$ die Bestapproximierende zu f, dann ist, da $t_n \to f$ für $n \to \infty$,

$$f(x) - t_n(x) = \sum_{j=0}^{\infty} \left(t_{2^{j+1}n} - t_{2^j n}\right)(x), \qquad x \in \mathbb{T},$$

also

$$f = t_n + \sum_{j=0}^{\infty} \left(t_{2^{j+1}n} - t_{2^j n}\right), \tag{5.34}$$

wobei die Reihe auf der rechten Seite gleichmäßig konvergiert:

$$\left\| f - \left(t_n + \sum_{j=0}^{m} \left(t_{2^{j+1}n} - t_{2^j n}\right)\right)\right\| = \|f - t_{2^{m+1}n}\| \to 0, \qquad \text{für } m \to \infty.$$

[121] Genauer: dessen Beweis!

[122] Beziehungsweise dessen Anwendung zum Beweis von Satz 5.8, "⇐".

Nun differenzieren wir die Reihe auf der rechten Seite von (5.34) gliedweise und erhalten mit Hilfe der Bernstein–Ungleichung (5.20), und wie im Beweis von Satz 5.17 daß

$$\begin{aligned}
\sum_{j=0}^{\infty} \left\| t^{(k)}_{2^{j+1}n} - t^{(k)}_{2^j n} \right\| &\le \sum_{j=0}^{\infty} \left(2^{j+1}\, n\right)^k \left\| t_{2^{j+1}n} - t_{2^j n} \right\| \\
&\le \sum_{j=0}^{\infty} \left(2^{j+1}\, n\right)^k \left(\left\| t_{2^{j+1}n} - f \right\| + \left\| f - t_{2^j n} \right\| \right) \le 2 \sum_{j=0}^{\infty} \left(2^{j+1}\, n\right)^k \underbrace{E^*_{2^j n}(f)}_{=:E(2^j n)} \\
&= 2 \sum_{j=1}^{\infty} 2^{jk}\, n^k\, E\left(2^{j-1} n\right) = 2^k \sum_{j=1}^{\infty} 2^{(j-1)(k-1)}\, n^k\, E\left(2^{j-1} n\right) \left(2^j - 2^{j-1}\right) \\
&\le 2^k\, n^k \int_1^{\infty} t^{k-1} E\left(nt\right)\, dt = 2^k\, n^k\, \frac{1}{n} \int_n^{\infty} \left(\frac{t}{n}\right)^{k-1} E\left(t\right)\, dt = 2^k \int_n^{\infty} t^{k-1} E\left(t\right)\, dt \\
&\le 2^k \sum_{j=n}^{\infty} (j+1)^{k-1}\, E(j) = 2^k \sum_{j=n}^{\infty} (j+1)^{k-1}\, E_j(f),
\end{aligned}$$

was für $n \to \infty$ gegen 0 konvergiert; die abgeleitete Reihe auf der rechten Seite von (5.34) konvergiert also auch gleichmäßig. Jetzt verwenden wir wieder mal ein bißchen Analysis[123] und erkennen so, daß $f^{(k)} \in C(\mathbb{T})$ existiert und daß die Ableitungen dagegen konvergieren. Mit $M = 4^k$ folgt dann auch (5.33), da $(j+1)/j = 1 + \frac{1}{j} \le 2$ für $j \in \mathbb{N}$ ist. □

Korollar 5.21 *Für $f \in C(\mathbb{T})$ und $0 < \alpha < 1$ gilt:*

$$\sup_{n\in\mathbb{N}} n^{k+\alpha} E^*_n(f) < \infty \qquad \Longrightarrow \qquad f \in C^k(\mathbb{T}) \quad \textit{und} \quad \sup_{n\in\mathbb{N}} n^{\alpha} E^*_n\left(f^{(k)}\right) < \infty. \tag{5.35}$$

Beweis: Sei $M_f := \sup_n n^{k+\alpha} E^*_n(f)$, dann ist

$$\sum_{n=1}^{\infty} n^{k-1}\, E^*_n(f) \le M_f \sum_{n=1}^{\infty} n^{k-1} n^{-k-\alpha} = M_f \underbrace{\sum_{n=1}^{\infty} \frac{1}{n^{1+\alpha}}}_{<\infty} < \infty,$$

weswegen $f \in C^k(\mathbb{T})$ ist. Und nach (5.33) ist außerdem

$$E^*_n\left(f^{(k)}\right) \le \sum_{\ell=n}^{\infty} \ell^{k-1}\, E^*_\ell(f) \le M_f \sum_{\ell=n}^{\infty} \frac{1}{\ell^{1+\alpha}} \le M_f \underbrace{\int_n^{\infty} t^{-1-\alpha}\, dt}_{=\alpha^{-1} n^{-\alpha}} = \frac{M_f}{\alpha} n^{-\alpha},$$

was den Beweis komplettiert. □

Der kapitelabschließende[124] Beweis ist nunmehr fast eine reine Formsache.

[123] Beispielsweise, aus einem Standardlehrbuch für Analysis, [30, Satz 104.6, S. 554].

[124] Dieses Wort ist zu schön, um dem Zerstückelungs Wahn der "Recht"schreibreform anheimzufallen.

Beweis von Satz 5.18: Ist $f \in C^{k+\alpha}(\mathbb{T})$, dann ist nach Satz 5.19 für $n \in \mathbb{N}$

$$n^{k+\alpha}\, E_n^*(f) \le M_k \underbrace{\; n^\alpha\, \omega\left(f^{(k)}, \frac{1}{n}\right)}_{<\infty} < \infty,$$

was "$\Rightarrow$" beweist; für "$\Leftarrow$" müssen wir nur Korollar 5.21 mit Satz 5.8 kombinieren. □

5.6 Trigonometrische Polynome V: Die Zygmund–Klasse

Bleibt also eigentlich nur eine Frage, nämlich: Was ist mit den ganzzahligen Approximationsordnungen, also mit $\sup n^k E_n^*(f) < \infty$? Daß wir nicht so ganz einfach $\alpha \to 0$ oder $\alpha \to 1$ gehen lassen können, das zeigen uns schon die Beweise von Satz 5.18 und Satz 5.8, "$\Leftarrow$", wo durch α beziehungsweise $1-\alpha$ dividiert wurde. Außerdem machen ja auch die Teilklassen von $C(\mathbb{T})$, die durch

$$\lim_{\delta\to 0} \delta^\epsilon\, \omega(f,\delta) = 0, \qquad \epsilon \in \{0,1\}$$

definiert sind, nur wenig Sinn: für $\epsilon = 0$ erhält man ganz $C(\mathbb{T})$, für $\epsilon = 1$ gerade die konstanten Funktionen, siehe Übung 4.3. Die Antwort hierauf wurde von Zygmund[125] [87] gegeben, braucht aber noch ein klein wenig Notation.

Definition 5.22 *Wir definieren die* verallgemeinerten Lipschitz–Klassen

$$C^{\alpha,r}(\mathbb{T}) := \left\{ f \in C(\mathbb{T}) \;:\; \sup_{\delta>0} \delta^{-\alpha}\, \omega_r(f,\delta) < \infty \right\}, \tag{5.36}$$

so daß $C^\alpha(\mathbb{T}) = C^{\alpha,1}(\mathbb{T})$, $\alpha \in (0,1)$.
Die verallgemeinerte Lipschitz–Klasse $C^{1,2}(\mathbb{T})$ *heißt* Zygmund–Klasse *und mit*

$$C_0^{1,2}(\mathbb{T}) := \left\{ f \in C(\mathbb{T}) \;:\; \lim_{\delta\to 0} \delta^{-1}\, \omega_2(f,\delta) = 0 \right\}$$

bezeichnen wir die "glatten" *Funktionen*[126]

Bemerkung 5.23 *Hier ein paar Informationen über die verallgemeinerten Lipschitz–Klassen.*

1. *So verallgemeinert sind die Lipschitz–Klassen eigentlich gar nicht: es gilt nämlich*

$$C^{\alpha,1}(\mathbb{T}) = C^{\alpha,2}(\mathbb{T}), \qquad 0 < \alpha < 1,$$

aber, und darauf kommt es an,

$$C^{1,1}(\mathbb{T}) \subset C^{1,2}(\mathbb{T}), \tag{5.37}$$

siehe Übung 5.7

[125] Antoni Zygmund, 1900–1992, (harmonischer) Analytiker, baute in Chicago eine der stärksten Analysis–Schulen auf. Außerdem, wen wundert's, Beiträge zur Fourier–Analysis und deren Anwendung auf partielle Differentialgleichungen.

[126] Englisch: *"smooth"*, siehe auch den Titel von [87].

2. *Auch der Name "glatt" für die "o"–Zygmund–Klasse $C_0^{1,2}(\mathbb{T})$ ist durchaus nicht unberechtigt, denn diese Funktionen sind an einer dichten Menge von Punkten differenzierbar, siehe Übung 5.8. Allerdings ist das nicht allzu viel, denn es gibt Funktionen in $C_0^{1,2}(\mathbb{T})$, die nur an einer Menge vom Lebesgueschen Maß 0 differenzierbar sind.*

Übung 5.7 Gegeben seien die Funktionen

$$f(x) = \sum_{k=1}^{\infty} \frac{\cos kx}{k^2} \qquad \text{und} \qquad g(x) = \sum_{k=1}^{\infty} \frac{\sin kx}{k^2}.$$

Zeigen Sie:

1. $f, g \in C^{1,2}(\mathbb{T})$.
2. $g \notin C^{1,1}(\mathbb{T})$.

♢

Übung 5.8 Zeigen Sie: Ist $f \in C_0^{1,2}(\mathbb{T})$, dann gibt es im Inneren jedes nichttrivialen Intervalls $I \subset \mathbb{T}$ einen Punkt x, an dem f differenzierbar ist.
Hinweis: Ist $[a, b]$ ein nichttriviales kompaktes Intervall und $\ell \in \Pi_1$ diejenige affine Funktion, die $\ell(a) = f(a)$ und $\ell(b) = f(b)$ erfüllt, dann besitzt $f - \ell$ an einem Punkt $x^* \in (a, b)$ ein Extremum. Dort ist f differenzierbar. ♢

Und tatsächlich beantwortet die Zygmund–Klasse die Frage nach den Konvergenzordnungen mit ganzzahligen Exponenten.

Satz 5.24 *Für $f \in C(\mathbb{T})$ und $k \in \mathbb{N}$ gilt*

$$\sup_{n\in\mathbb{N}} n^k \, E_n^*(f) < \infty \qquad \Longleftrightarrow \qquad f \in C^{k-1}(\mathbb{T}) \quad \textit{und} \quad f^{(k-1)} \in C^{1,2}(\mathbb{T}). \tag{5.38}$$

5.7 Algebraische Polynome

Jetzt aber zu den *algebraischen* Polynomen und deren Approximationsgüte auf dem Intervall $I = [-1, 1]$. Die grundlegende Idee ist die Transformation $x = \cos\xi$, die $\mathbb{T}$ auf $[-1, 1]$ abbildet[127], aber die Supremumsnorm nicht verändert: Für $f \in C(I)$ und $p \in \Pi_n$ ist

$$\|f - p\|_I = \max_{x\in[-1,1]} |f(x) - p(x)| = \max_{\xi\in\mathbb{T}} |f(\cos\xi) - p(\cos\xi)| = \left\|\widetilde{f} - \widetilde{p}\right\|_{\mathbb{T}}$$

wobei $\widetilde{f} \in C(\mathbb{T})$ und $\widetilde{p} \in T_n$ *gerade* Funktionen sind.

Übung 5.9 Zeigen Sie:

[127] Zwar nicht eineindeutig, aber das stört nicht wirklich

1. Ist $p \in \Pi_n$, dann ist

$$\widetilde{p}(\xi) = p(\cos\xi) = \frac{p_0}{2} + \sum_{k=1}^{n} p_k \cos k\xi, \qquad \xi \in \mathbb{T}.$$

2. Ist $\widetilde{p} \in T_n$ ein *gerades* trigonometrisches Polynom, dann gibt es ein Polynom $p \in \Pi_n$, so daß $\widetilde{p}(\xi) = p(\cos\xi)$.

♢

Proposition 5.25 *(Jackson–Satz für algebraische Polynome)*
Es gibt eine Konstante $M > 0$, so daß[128]

$$E_n(f) := \inf_{p\in\Pi_n} \|f - p\| \le M\,\omega\left(f, \frac{1}{n}\right), \qquad f \in C(I). \tag{5.39}$$

Beweis: Sei $\widetilde{p} \in T_n$ die Bestapproximierende zu $\widetilde{f} = f(\cos\cdot)$, dann ist $\widetilde{p}$ eine gerade Funktion und es gibt, siehe Übung 5.9, ein $p \in \Pi_n$, so daß $\widetilde{p} = p(\cos\cdot)$. Damit ist

$$E_n(f) \le \|f - p\|_I = \left\|\widetilde{f} - \widetilde{p}\right\|_{\mathbb{T}} \le M\,\omega\left(\widetilde{f}, \frac{1}{n}\right) \le M\,\omega\left(f, \frac{1}{n}\right),$$

da für $h > 0$ die Ungleichung $\cos(\xi + h) - \cos\xi \le h$ gilt und somit für $\xi \in \mathbb{T}$

$$\begin{aligned}\sup_{0<h\le\delta}\left|f(\cos(\xi+h)) - f(\underbrace{\cos\xi}_{=:x})\right| &\le \sup_{0<h\le\delta} |f(x + (\cos(\xi+h) - \cos\xi)) - f(x)| \\ &\le \sup_{0<h\le\delta} |f(x+h) - f(x)| \le \omega(f,\delta)\end{aligned}$$

ist. □

Was aber ist mit der Umkehrung, mit den Bernsteinsätzen. Nun, die können leider nicht mehr funktionieren, wie das folgende Beispiel zeigt.

Beispiel 5.26 *Wir betrachten $f(x) = \sqrt{1-x^2}$, $x \in I$. Mit der Substitution $x = \cos\xi$ ist dann*

$$\widetilde{f}(\xi) = \sqrt{1 - \cos^2\xi} = \sqrt{\sin^2\xi} = |\sin\xi|, \qquad \xi \in \mathbb{T}.$$

Nun ist

$$E_n(f) = E_n^*\left(\widetilde{f}\right) \sim \omega\left(\widetilde{f}, \frac{1}{n}\right), \qquad n \in \mathbb{N}, \tag{5.40}$$

[128] Achtung: Die Definition des Stetigkeitsmoduls für Funktionen in $C(I)$ unterscheidet sich in einem kleinen aber feinen Detail von der für Funktionen in $C(\mathbb{T})$, da man jetzt darauf achten muß, daß man *im Intervall* bleibt. Nachdem aber aus dem Kontext klar sein wird, mit welchem Typ von Funktionen man es zu tun hat, werden wir in beiden Fällen das Symbol "ω" verwenden.

wobei zwei Folgen a_n, b_n, $n \in \mathbb{N}$, die asymptotische *Relation $a_n \sim b_n$ erfüllen, wenn es Konstanten $m, M > 0$ gibt, so daß*

$$m\,|a_n| \le |b_n| \le M\,|a_n|\,, \qquad n \in \mathbb{N},$$

gilt. Gemäß Übung 5.10 und Lemma 5.16 ist

$$\omega\left(\widetilde{f}, \delta\right) = \omega\left(|\sin\cdot|\,, \delta\right) \le \omega\left(\sin\cdot, \delta\right) \le \delta\,\|\cos\cdot\| = \delta$$

sowie, wegen der Konkavität des Cosinus,

$$\left|\Delta_h \widetilde{f}(0)\right| = \int_0^h \cos t\,dt \ge h\cos h \ge 2\,\frac{h(1-h)}{\pi},$$

also ist mit (5.40)

$$E_n(f) \sim \omega\left(\widetilde{f}, \frac{1}{n}\right) \sim \frac{1}{n}. \tag{5.41}$$

Andererseits ist aber, für $\delta < 1$

$$\begin{aligned}\omega\left(f,\delta\right) &\ge |\Delta_\delta f(-1)| = \left|\sqrt{1-(-1+\delta)^2} - \sqrt{1-(-1)^2}\right| = \left|\sqrt{2\delta-\delta^2}\right| = \sqrt{\delta}\,\sqrt{2-\delta} \\ &\ge \sqrt{\delta}.\end{aligned}$$

Somit ist also $\omega\left(f, \frac{1}{n}\right) \sim \sqrt{\frac{1}{n}}$, weswegen $E_n(f) \sim \omega\left(f, \frac{1}{n}\right)$ wie im Fall der trigonometrischen Polynome offensichtlich unmöglich ist.

Übung 5.10 Zeigen Sie: Für $f \in C(\mathbb{T})$ und $\delta > 0$ ist $\omega\left(|f|\,, \delta\right) \le \omega(f, \delta)$. ♢

Irgendwas funktioniert also nicht so ganz mit unserer schönen Substitution und es muß wohl irgendwas mit dem Rand zu tun haben. Einen ersten Hinweis, woran es liegen könnte, bekommen wir, wenn wir einmal annehmen, daß $f \in C^1(I)$ ist, denn dann ist ja, unter Verwendung von[129] (5.30)

$$\begin{aligned}E_n(f) &= E_n^*\left(\widetilde{f}\right) \le M\,\frac{1}{n}\,E_n^*\left(\widetilde{f}'\right) \le \frac{M}{n}\,\left\|\widetilde{f}'\right\|_{\mathbb{T}} = \frac{M}{n}\,\left\|\frac{d}{d\xi} f\left(\cos\xi\right)\right\|_{\mathbb{T}} \\ &= \frac{M}{n}\,\left\||\sin\cdot|\;\, f'\left(\cos\cdot\right)\right\|_{\mathbb{T}} = \frac{M}{n}\,\left\|\sqrt{1-\cos^2\cdot}\;\, f'\left(\cos\cdot\right)\right\|_{\mathbb{T}} \\ &= \frac{M}{n}\,\left\|\sqrt{(1+\cdot)(1-\cdot)}\;\, f'\right\|_I.\end{aligned}$$

Mit anderen Worten:

> *Es muß also etwas mit dem "Abstand" $\sqrt{1-x^2} = \sqrt{(1-x)(1+x)}$ des Punktes x von den Randpunkten ± 1 von I zu tun haben, die Approximationsgüte wird also am Rand* besser.

[129] Manchmal kann man einen Beweis ja irgendwann nochmals recyclen.

Um nun Proposition 5.25 so verschärfen zu können, daß wir auch die Chance einer Umkehrung besitzen, brauchen wir noch ein klein wenig Notation.

Definition 5.27 *Für $n \in \mathbb{N}$ und $x \in I$ setzen wir*

$$\delta_n(x) = \max\left\{\sqrt{1-x^2}, \frac{1}{n}\right\}.$$

Nun können wir das folgende Resultat angeben, das auf Timan [82] zurückgeht.

Satz 5.28 *Es gibt eine Konstante $M > 0$, so daß für alle Funktionen $f \in C(I)$ und $n \in \mathbb{N}$ ein Polynom $p_n \in \Pi_n$ existiert, das die Ungleichung*

$$|f(x) - p_n(x)| \le M\,\omega\left(f, \delta_n(x)\right), \qquad x \in I, \tag{5.42}$$

erfüllt.

Und Satz 5.28 hat nun wieder eine Umkehrung, die es ermöglicht, aus der Approximationsordnung auf die Glattheit der Funktion zu schließen. Doch dafür brauchen wir noch den Begriff des Stetigkeitsmoduls "per se".

Definition 5.29 *Eine Abbildung $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ heißt* Stetigkeitsmodul*, wenn*

1. *ω monoton steigend ist,*
2. *$\omega(t) \to 0$ für $t \to 0$ gilt,*
3. *ω* subadditiv *ist, das heißt*

$$\omega\left(t + t'\right) \le \omega(t) + \omega(t'), \qquad t, t' \in \mathbb{R}_+.$$

Übung 5.11 Zeigen Sie: Ist ω ein Stetigkeitsmodul, so existiert eine Funktion f, so daß $\omega(\delta) = \omega(f, \delta)$. ♢

Die Variante des Bernstein–Satzes für algebraische Polynome auf I kann man dann wie folgt angeben.

Satz 5.30 *Seien $f \in C(I)$ und ω ein Stetigkeitsmodul. Gibt es Polynome $p_n \in \Pi_n$, so daß*

$$|f(x) - p_n(x)| \le \omega\left(\delta_n(x)\right), \qquad x \in I, \tag{5.43}$$

dann ist für $\delta > 0$

$$\omega\left(f, \delta\right) \le M\,\delta \sum_{n=1}^{\lfloor \delta^{-1} \rfloor} \omega\left(\frac{1}{n}\right), \tag{5.44}$$

wobei die Konstante M von f und δ unabhängig ist.

Bemerkung 5.31 *1. Wir werden nur Satz 5.28 beweisen. Der Beweis von Satz 5.30 ist von der Grundidee her so ähnlich wie der des Bernstein–Satzes 5.17, nur eben in Details wegen des Auftretens von $\delta_n(x)$ deutlich komplizierter. Der Beweis kann beispielsweise in [47, Theorem 4, S. 73–75] nachgeschlagen werden.*

2. *Satz 5.28 hat natürlich auch wieder ein Gegenstück mit höheren Approximationsordnungen, bei der, wie in Proposition 5.19, dem Jackson–Satz für differenzierbare Funktionen, der Stetigkeitsmodul dann auf entsprechende* Ableitungen *von f angewandt wird, siehe [47, Theorem 2, S. 66].*

3. *Es ist nicht so ganz unmittelbar offensichtlich, daß Satz 5.30 wirklich ein Gegenstück zum Bernstein–Satz 5.13 ist, denn es taucht ja nirgendwo $E_n(f)$ auf. Aber natürlich sollte man die p_n in (5.43) schon als Bestapproximationen sehen – ansonsten könnte man ja mit einem "kleineren" Stetigkeitsmodul ein "besseres" Ergebnis erzielen.*

4. *Man kann es auch so sehen: $E_n(f)$ als* globale[130] *Größe verträgt sich halt nicht mit dem* lokalen *Charakter der Polynomapproximation.*

Bringen wir's also hinter uns ...

Beweis von Satz 5.28: Wir setzen wieder $\widetilde{f} := f(\cos\cdot)$, dann $t_n := J_n\widetilde{f} = \widetilde{f} * K_n$ und wählen schließlich p_n so, daß $t_n = p_n(\cos\cdot)$. Also ist, für $x = \cos\xi$

$$\begin{aligned}
|f(x) - p_n(x)| &= |f(\cos\xi) - (K_n * f(\cos\cdot))(\xi)| \\
&\le \int_{\mathbb{T}} |f(\cos\xi) - f(\cos(\xi+\vartheta))|\ K_n(\vartheta)\ d\vartheta \le \int_{\mathbb{T}} \omega(f, |\cos(\xi+\vartheta) - \cos\xi|)\ K_n(\vartheta)\ d\vartheta \\
&\le \int_{\mathbb{T}} \left(\frac{|\cos(\xi+\vartheta) - \cos\xi|}{\delta_n(x)} + 1 \right) \omega(f, \delta_n(x))\ K_n(\vartheta)\ d\vartheta \\
&\le \omega(f, \delta_n(x)) \underbrace{\int_{\mathbb{T}} \left(\frac{|\cos(\xi+\vartheta) - \cos\xi|}{\delta_n(x)} + 1 \right) K_n(\vartheta)\ d\vartheta}_{=:J}.
\end{aligned}$$

Wir müssen also noch zeigen, daß J unabhängig von n beschränkt ist. Nach den Additionstheoremen und mit der inzwischen wohlvertrauten Abschätzung $|\sin t| \le |t|$ ist zuerst einmal

$$\begin{aligned}
|\cos(\xi+\vartheta) - \cos\xi| &= 2\left|\sin\frac{\vartheta}{2}\ \sin\left(\xi + \frac{\vartheta}{2}\right)\right| = 2\left|\sin\frac{\vartheta}{2}\left(\sin\xi\ \cos\frac{\vartheta}{2} + \cos\xi\ \sin\frac{\vartheta}{2}\right)\right| \\
&\le 2\underbrace{\left|\cos\frac{\vartheta}{2}\right|}_{\le 1}\left|\sin\frac{\vartheta}{2}\ \sin\xi\right| + 2\underbrace{|\cos\xi|}_{\le 1}\ \sin^2\frac{\vartheta}{2} \le \frac{\vartheta^2}{2} + |\vartheta|\ |\sin\xi| \le \frac{\vartheta^2}{2} + |\vartheta|\ \delta_n(x),
\end{aligned}$$

da ja

$$\delta_n(x) = \max\left\{\sqrt{1-x^2}, \frac{1}{n}\right\} = \max\left\{\sqrt{1-\cos^2\xi}, \frac{1}{n}\right\} = \max\left\{|\sin\xi|, \frac{1}{n}\right\} \ge |\sin\xi|.$$

[130] Schließlich ist das ja die *Norm* der besten Abweichung!

Also ist, unter Verwendung von Lemma 5.10[131]

$$\begin{aligned} J &= \int_{\mathbb{T}} \left(\frac{\vartheta^2}{\delta_n(x)} + |\vartheta| + 1 \right) K_n(\vartheta)\, d\vartheta \\ &= \frac{1}{\delta_n(x)} \underbrace{\int_{\mathbb{T}} \vartheta^2 K_n(\vartheta)\, d\vartheta}_{\leq M_2 n^{-2}} + \underbrace{\int_{\mathbb{T}} |\vartheta| K_n(\vartheta)\, d\vartheta}_{\leq M_1 n^{-1}} + \underbrace{\int_{\mathbb{T}} K_n(\vartheta)\, d\vartheta}_{=1} \\ &= \frac{M_2}{n^2} \underbrace{\frac{1}{\delta_n(x)}}_{\leq 1/n} + \frac{M_1}{n} + 1 \leq 1 + \frac{M_1 + M_2}{n} \leq 1 + M_1 + M_2 =: M, \end{aligned}$$

was unseren Beweis auch schon komplettiert. □

[131] Hurrah, wir recyclen schon wieder! Und zwar schlampig, denn eigentlich ist ja $K_n = J_{\lfloor n/2 \rfloor + 1}$, aber diesen "Fehler" stecken wir in die Konstanten M_1 und M_2.

I disapprove of certainties [. . .] They limit one's range of vision. Doubt is one aspect of width.

S. Rushdie, *Grimus*

6 Approximation mit translationsinvarianten Räumen

In diesem Kapitel behandeln wir die Approximation von Funktionen auf $\mathbb{R}$, also wieder einem Bereich *ohne* Rand. Eine bzw. **die** Gruppenoperation auf $\mathbb{R}$ sind Addition und Subtraktion und diese Invarianz sollte sich dann auch auf "vernünftige" Maße auf $\mathbb{R}$ übertragen, zumindest, wenn man $\mathbb{R}$ nach wie vor als additive[132] Gruppe betrachten möchte[133]. Insbesondere sind die Räume $C(\mathbb{R})$ oder

$$L_p(\mathbb{R}) := \left\{ f : \mathbb{R} \to \mathbb{R} \ : \ \|f\|_p := \left(\int_{\mathbb{R}} |f(t)|^p \, dt \right)^{1/p} < \infty \right\}, \qquad 1 \le p < \infty,$$

den wir hier vor allem betrachten wollen[134], ebenso *translationsinvariant*

$$f \in L_p(\mathbb{R}) \qquad \Longleftrightarrow \qquad \tau_y f := f(\cdot + y) \in L_p(\mathbb{R}), \quad y \in \mathbb{R}, \tag{6.1}$$

wie die Norm selbst: $\|f\|_p = \|\tau_y f\|_p$, $y \in \mathbb{R}$. Für "praktische" Zwecke sind solche Räume, die invariant unter *beliebigen* Verschiebungen sind, aber zu groß[135], weswegen man sich gerne auf Räume beschränkt, die nur unter einer "diskreten" Menge von Translationen invariant sind – hier bieten sich natürlich die *ganzzahligen* Translationen $f(\cdot + k)$, $k \in \mathbb{Z}$, an.

Bemerkung 6.1 *Wegen des unendlichen Trägers des Lebesgue*[136]*–Maßes auf* $\mathbb{R}$ *verhalten sich die Räume* $L_p(\mathbb{R})$ *schon ein wenig anders als die Räume* $L_p(I)$ *zu einem kompakten Intervall*

[132] Als *multiplikative* Gruppe nimmt man besser $\mathbb{R}_+$ – warum eigentlich nicht $\mathbb{R}_-$?

[133] Maße, die invariant unter einer Gruppenoperation des Integrationsbereichs sind, werden als *Haar-Maße* bezeichnet und stellen die Grundlage der abstrakten harmonischen Analysis dar, siehe [35].

[134] Wir erweitern so ganz nebenbei einmal unseren Horizont und untersuchen Approximationsprozesse in einer p–Norm.

[135] Wenn man vielleicht einmal von $C(\mathbb{R})$ und $L_p(\mathbb{R})$ absieht, aber beispielsweise die Approximation von Elementen von $L_p(\mathbb{R})$ durch $L_p(\mathbb{R})$ ist nicht sonderlich interessant.

[136] Henri Léon Lebesgue, 1875–1941, (Mit-)Begründer der Maßtheorie, Erfinder des Lebesgue–Integrals (was nun nicht so sehr überrascht), wichtige Beiträge zur Fourier–Analysis. Von ihm stammt das bemerkenswerte antibourbakistische Zitat:

> *Reduced to general theories, mathematics would be a beautiful form without content. It would quickly die.*

(gefunden auf der "History of Mathematics"–Website)

I. Im letzteren Fall gilt beispielsweise immer, daß $L_p(I) \subset L_1(I)$, $1 < p \leq \infty$, ist[137]*, auf $\mathbb{R}$ gehört aber zur Integrierbarkeit immer noch ein gewisses Abklingen von $|f(x)|^p$ für $x \to \pm\infty$, was mit fallendem Exponenten p nicht mehr der Fall sein muß.*

6.1 Translationsinvariante Räume

Wir nennen einen Teilraum $V \subset L_2(\mathbb{R})$ *translationsinvariant*[138], englisch *"shift invariant"*, wenn

$$f \in V \qquad \Longleftrightarrow \qquad f(\cdot \pm k) \in V, \quad k \in \mathbb{Z}.$$

Demnach sind die einfachsten translationsinvarianten Räume also Räume der Form[139]

$$V = \text{span } \{\varphi(\cdot - k) \ : \ k \in \mathbb{Z}\} = \left\{ \sum_{k \in \mathbb{Z}} c_k \, \varphi(\cdot - k) \ : \ c_k \in \mathbb{R}, \, k \in \mathbb{Z} \right\}, \tag{6.2}$$

die von den Translaten *einer* Funktion φ erzeugt werden. Solche Räume werden gern als *principal*[140] *shift invariant spaces* ("PSI") bezeichnet. Die nächste Stufe wären dann "FSI"–Räume[141], die von den (ganzzahligen) Translaten von *endlich vielen* Funktionen aufgespannt werden. Wir bleiben aber einfach und werden uns daher nur mit Räumen der Form 6.2 befassen.

Definition 6.2 *Mit $\ell(\mathbb{Z}) = \{c : \mathbb{Z} \to \mathbb{R}\}$ bezeichnen wir den Vektorraum aller doppeltunendlichen Folgen, die wir bequemerweise als* diskrete *Funktionen $c : \mathbb{Z} \to \mathbb{R}$ schreiben werden und mit $\ell_p(\mathbb{Z})$, $1 \leq p < \infty$, den Vektorraum derjenigen Folgen $c \in \ell(\mathbb{Z})$, die*

$$\|c\|_p := \left(\sum_{j \in \mathbb{Z}} |c(j)|^p \right)^{1/p} < \infty$$

erfüllen. Entsprechend ist

$$\ell_\infty(\mathbb{Z}) = \left\{ c \in \ell(\mathbb{Z}) \ : \ \|c\|_\infty := \sup_{j \in \mathbb{Z}} |c(j)| < \infty \right\}.$$

Zwischen den Folgenräumen $\ell(\mathbb{Z})$, bzw. $\ell_p(\mathbb{Z})$, $1 \leq p \leq \infty$ und der Funktion φ kann man nun *Faltung* definieren, indem man das Integral im kontinuierlichen Gegenstück

$$f * g := \int_{\mathbb{R}} f(\cdot - t) \, g(t) \, dt$$

[137] Das einzige was passieren kann ist, daß eine Funktion irgendwo "gegen Unendlich geht" und das ist mit Exponent 1 immer langsamer als mit Exponent > 1

[138] Nach dem, was wir gerade gesagt haben, wäre "ganzzahlig translationsinvariant" wohl der korrektere Begriff, aber man muß ja nicht unnötig mit Worten um sich werfen.

[139] Ob wir hier nun "$+k$" oder "$-k$" schreiben, das ist offensichtlich irrelevant. Warum wir's nun gerade so und nicht anders machen, wird aber hoffentlich bald klar werden.

[140] Das Wort "principal" wurde wohl im Anklang an *"principal ideals"*, auf Deutsch *"Hauptideale"* gewählt, das sind (polynomiale) Ideale, die von einem einzigen Polynom erzeugt werden. Und nun übersetze man das mal ins Deutsche ...

[141] **F**initely generated **S**hift **I**nvariant spaces

von (1.8) durch eine Summe ersetzt und so

$$\varphi * c := \sum_{j\in\mathbb{Z}} \varphi\left(\cdot - j\right)\ c(j) \tag{6.3}$$

erhält, vorausgesetzt, die Reihe auf der rechten Seite konvergiert. Mit dieser Notation, die auch manchmal als *semidiskrete Faltung* bezeichnet wird, siehe z.B. [34], können wir dann unsere translationsinvarianten Räume besonders einfach schreiben.

Definition 6.3 *Für eine Funktion $\varphi : \mathbb{R} \to \mathbb{R}$ bezeichnen wir mit*

$$\mathbb{S}\left(\varphi\right) := \left\{\varphi * c \ :\ c \in \ell(\mathbb{Z})\right\}$$

den algebraischen Span *von φ und definieren außerdem*

$$\mathbb{S}_p\left(\varphi\right) := \left\{\varphi * c \ :\ c \in \ell_p(\mathbb{Z})\right\}, \qquad 1 \le p \le \infty.$$

Übung 6.1 Zeigen Sie:

1. Die Faltung $\varphi * c$ ist linear in φ und c.
2. Die Räume $\mathbb{S}(\varphi)$ und $\mathbb{S}_p(\varphi)$, $1 \le p \le \infty$, sind (ganzzahlig) translationsinvariant.

♢

Zum "Aufwärmen" und vor allem um ein Gefühl für die verwendeten Begriffe zu bekommen, jetzt erst einmal eine einfache Beobachtung.

Lemma 6.4 *Ist $\varphi \in L_1(\mathbb{R})$, dann ist $\mathbb{S}_1(\varphi) \subset L_1(\mathbb{R})$.*

Beweis: Es ist

$$\begin{aligned}
\left\|\varphi * c\right\|_1 &= \int_{\mathbb{R}} \left|\sum_{j\in\mathbb{Z}} \varphi\left(t-j\right)\ c(j)\right|\ dt \le \int_{\mathbb{R}} \sum_{j\in\mathbb{Z}} \left|\varphi\left(t-j\right)\ c(j)\right|\ dt \\
&= \underbrace{\sum_{j\in\mathbb{Z}} |c(j)|}_{=\|c\|_1} \underbrace{\int_{\mathbb{R}} \left|\varphi\left(t-j\right)\right|\ dt}_{=\|\varphi\|_1} = \|\varphi\|_1\ \|c\|_1 .
\end{aligned}$$

□

Für $p > 1$ ist die Sache leider nicht mehr so einfach! Am leichtesten sieht man das im Fall $p = \infty$ mit $\varphi \equiv 1$ und $c \equiv 1$, denn dann divergiert $\varphi * c$ überall und die Norm ist in keinster Weise beschränkt. Die einfachste Zusatzbedingung an φ besteht darin, zu fordern, daß φ *kompakten Träger* hat.

Proposition 6.5 *Hat $\varphi \in L_p(\mathbb{R})$ kompakten Träger, dann ist $\mathbb{S}_p(\varphi) \subset L_p(\mathbb{R})$.*

Beweis: Der etwas sorgfältigere Beweis basiert auf der Hölder–Ungleichung

$$\|c \cdot d\|_1 \leq \|c\|_p \, \|d\|_q \,, \qquad c \in \ell_p(\mathbb{Z}),\, d \in \ell_q(\mathbb{Z}), \quad \frac{1}{p} + \frac{1}{q} = 1, \tag{6.4}$$

für Folgen. Da wir bei der Bildung von $\mathbb{S}_p(\varphi)$ außerdem immer φ durch $\varphi\,(\cdot - j)$ für beliebiges $j \in \mathbb{Z}$ ersetzen können, können wir annehmen, daß $\varphi(x) = 0$, $x \notin [0, N]$, für ein $N > 0$. Nun ist

$$\begin{aligned}
\|\varphi * c\|_p^p &= \int_{\mathbb{R}} \left| \sum_{j\in\mathbb{Z}} \varphi\,(t-j)\;c(j) \right|^p \, dt \leq \sum_{k\in\mathbb{Z}} \int_k^{k+1} \left(\sum_{j\in\mathbb{Z}} |\varphi\,(t-j)\,c(j)| \right)^p \, dt \\
&= \sum_{k\in\mathbb{Z}} \int_0^1 \left(\sum_{j\in\mathbb{Z}} |\varphi\,(t+k-j)\,c(j)| \right)^p \, dt = \sum_{k\in\mathbb{Z}} \int_0^1 \left(\sum_{j\in\mathbb{Z}} |\varphi\,(t-j)\,c(j+k)| \right)^p \, dt \\
&= \sum_{k\in\mathbb{Z}} \int_0^1 \left(\sum_{j=0}^{N-1} |\varphi\,(t-j)\,c(j+k)| \right)^p \, dt = \sum_{k\in\mathbb{Z}} \int_0^1 \left(\left\| \chi_{[0,N-1]} \cdot \varphi(t-\cdot) c(\cdot+k) \right\|_1 \right)^p \, dt,
\end{aligned}$$

wobei $\chi_{[0,N-1]}$ die charakteristische Funktion des Intervalls $[0, N-1]$ ist[142]. Nun liefert (6.4), daß

$$\begin{aligned}
\|\varphi * c\|_p^p &\leq \sum_{k\in\mathbb{Z}} \int_0^1 \underbrace{\left\| \chi_{[0,N-1]} \right\|_q^p}_{=N^{p/q}=N^{p-1}} \, \|\varphi(t+\cdot) c(k-\cdot)\|_p^p \, dt \\
&= N^{p-1} \sum_{k\in\mathbb{Z}} \int_0^1 \sum_{j\in\mathbb{Z}} |\varphi(t-j)|^p \; |c(j+k)|^p \, dt = N^{p-1} \underbrace{\sum_{j\in\mathbb{Z}} |\varphi(t-j)|^p}_{=\|\varphi\|_p^p} \underbrace{\sum_{k\in\mathbb{Z}} |c(k)|^p}_{=\|c\|_p},
\end{aligned}$$

also

$$\|\varphi * c\|_p \leq N^{1/q} \, \|\varphi\|_p \, \|c\|_p \,, \tag{6.5}$$

was unsere Behauptung beweist. □

Übung 6.2 Beweisen Sie Proposition 6.5 für $p = \infty$. ♢

Bemerkung 6.6 *1. Wenigstens einmal sollte man so einen Beweis gesehen haben, denn die Abschätzungen sind durchaus typisch für die Arbeit mit translationsinvarianten Räumen und Wavelets in $L_p(\mathbb{R})$. Im Falle $p = 2$ wird of unmittelbar über die Fouriertransformierte argumentiert, dazu gleich mehr.*

2. Gleichung (6.5) sagt uns auch, warum der Fall $p = 1$ so einfach war: Ist $p = 1$, dann ist $1/q = 0$ und somit spielt die Größe des Trägers in diesem Fall und nur in diesem Fall schlichtweg keine Rolle. Und dann können wir auch getrost darauf verzichten.

[142] Ein schöner Nebeneffekt unserer Notation ist, daß wir auch Funktionen jederzeit als Folgen auffassen können – die Umkehrung geht natürlich aus guten Gründen nicht.

Sieht man sich den Beweis von Proposition 6.5 etwas genauer an, so stellt man fest, daß der kompakte Träger von φ bereits die ganze Geschichte nach L_1 verlagert.

Korollar 6.7 *Hat $\varphi \in L_p(\mathbb{R})$ kompakten Träger, dann ist $\varphi \in L_1(\mathbb{R})$.*

Beweis: Sei wieder $[0, N]$ der Träger von φ, dann lifert die Höldersche Ungleichung $\|fg\|_1 \leq \|f\|_p \|g\|_g$ für Funktionen, daß

$$\|\varphi\|_1 = \left\|\chi_{[0,N]} \cdot \varphi\right\|_1 \leq \left\|\chi_{[0,N]}\right\|_q \|\varphi\|_p = N^{1/q} \|\varphi\|_p .$$

□

Beispiel 6.8 *Uns fehlt noch ein griffiges Beispiel für eine "gute" Funktion φ, die $\mathbb{S}(\varphi)$ erzeugt. Dazu bezeichnen wir mit $\chi = \chi_{[0,1]}$ die charakteristische Funktion von $[0, 1]$ und definieren als*

$$N_0 = \chi, \qquad N_{j+1} = \chi * N_j = \int_{\mathbb{R}} \chi(t)\, N_j\,(\cdot - t)\; dt = \int_0^1 N_j\,(\cdot - t)\; dt, \qquad j \in \mathbb{N}_0,$$

den kardinalen B–Spline N_j *der Ordnung $j \in \mathbb{N}_0$. Diese Funktionen haben Träger $[0, j]$, sind $j - 1$–mal stetig differenzierbar*[143] *und stückweise Polynome vom Grad j auf jedem Intervall der Form $(k, k + 1)$.*

Übung 6.3 Beweisen Sie die in Beispiel 6.8 aufgeführten Eigenschaften der kardinalen B–Splines. ♢

Das Approximationsproblem mit dem wir uns hier beschäftigen wollen, soll aber nicht nur *einen* Raum $\mathbb{S}(\varphi)$ oder $\mathbb{S}_p(\varphi)$ verwenden, denn das wäre einfach zu wenig Flexibilität, wir wollen vielmehr auch noch die Skalierung σ_h, definiert durch $\sigma_h f = f\,(h\cdot)$, $f : \mathbb{R} \to \mathbb{R}$, $h \in \mathbb{R}_+$, der Funktion φ zulassen. Mit anderen Worten, für $h > 0$ sollen die Räume

$$V_h = \sigma_h\, \mathbb{S}\,(\varphi) = \left\{ (\varphi * c)\,(h\cdot) = \sum_{j\in\mathbb{Z}} \varphi(h \cdot - j)\, c(j) \;:\; c \in \ell(\mathbb{Z}) \right\}$$

zur Approximation zugelassen werden. Auch zur Konstruktion eines Elements dieser Räume genügt es immer noch, einzig die Funktion φ zu kennen.

Bemerkung 6.9 *Hat φ den kompakten Träger $[0, N]$, dann ist genau dann $\varphi(hx - j) \neq 0$, wenn $hx \in j + [0, N] = [j, j + N]$, also genau dann, wenn $x \in h^{-1}j + [0, h^{-1}N]$. Mit anderen Worten: Die Funktion φ wird um den Faktor h gestaucht und um j/h verschoben.*

Beispiel 6.10 *Ist $\varphi = N_0$, dann ist $V_1 = \mathbb{S}(\varphi)$ gerade die Menge aller Funktionen, die auf den ganzzahligen Intervallen $[j, j + 1]$, $j \in \mathbb{Z}$, konstant sind, wohingegen V_h gerade aus denjenigen Funktionen besteht, die auf $[j/h, (j + 1)/h]$ konstant sind.*

[143] Eine -1–mal stetig differenzierbare Funktion darf sogar unstetig sein, wohingegen 0–malige stetige Differenzierbarkeit einfach Stetigkeit ist.

6.2 Ein bißchen Fourieranalysis

Ein wichtiges Hilfsmittel bei der Betrachtung von Wavelets, aber auch in der Signalverarbeitung generell ist, vor allem in L_2 die *Fouriertransformierte* einer Funktion. Wir werden hier im wesentlichen den *Kalkül* bereitstellen und uns weniger um die theoretischen Konzepte der Fourieranalysis kümmern; für diese sei auf [35] verwiesen. Bei der Definition werden wir $f \in L_1(\mathbb{R})$ voraussetzen, was in unserem Kontext von Funktionen mit *kompaktem* Träger[144] völlig ausreichend ist, denn dann ist, siehe Corollar 6.7, f auch schon eine L_1–Funktion.

Definition 6.11 *Für $f \in L_1(\mathbb{R})$ definieren wir die* Fouriertransformierte $\widehat{f} : \mathbb{R} \to \mathbb{C}$ *als*

$$\widehat{f}(\xi) := f^\wedge(\xi) := \int_{\mathbb{R}} f(t)\, e^{-i\xi t}\, dt, \qquad \xi \in \mathbb{R}, \tag{6.6}$$

und die Fouriertransformierte einer Folge $c \in \ell_1(\mathbb{Z})$ als diskretes Gegenstück, die trigonometrische Reihe

$$\widehat{c}(\xi) := c^\wedge(\xi) := \sum_{k\in\mathbb{Z}} c(k)\, e^{-ik\xi}, \qquad \xi \in \mathbb{R}. \tag{6.7}$$

Bemerkung 6.12

1. *In ihrer physikalischen oder technischen Interpretation liefert die Fouriertransformierte eines "Signals" (das man als Aplitudenfunktion der Zeit ansieht), den Anteil der entsprechenden Frequenz an diesem Signal.*

2. *Die Bedingung $f \in L_1(\mathbb{R})$ garantiert, daß die $\widehat{f}(\xi)$ für alle $\xi \in \mathbb{R}$ existiert:*

$$\left|\widehat{f}(\xi)\right| \le \int_{\mathbb{R}} |f(t)| \underbrace{\left|e^{-i\xi t}\right|}_{=1} dt = \|f\|_1 . \tag{6.8}$$

   Allerdings ist das "nur" hinreichend, aber eben nicht notwendig für die Existenz der Fouriertransformierten.

3. *Manchmal wird die Fouriertransformierte auch noch mit dem Vorfaktor $(2\pi)^{1/2}$ versehen, wir werden bald sehen, warum. Man sollte also bei der Verwendung von Literatur immer gut aufpassen, welche Normierung dort gewählt ist, sonst kann so ein konstanter Faktor für üble Fehler sorgen.*

4. *Man kann die Fouriertransformierte auch für allgemeinere "Funktionen"klassen als L_1 definieren, beispielsweise für* temperierte Distributionen*, siehe z.B. [35, 86].*

5. *Außerdem gibt es die Fouriertransformation nicht nur auf $\mathbb{R}$ oder $\mathbb{R}^n$ sondern auf lokal kompakten abelschen Gruppen unter Verwendung des Haar–Maßes; dann sieht man, daß die Fouriertransformierte auf der* dualen Gruppe *definiert ist. Das soll uns aber hier nicht stören, in unserem einfachen aber bedeutenden Spezialfall spielt $\mathbb{R}$ beide Rollen.*

[144] Die einen "schönen" translationsinvarianten Raum erzeugen.

6. *Die zwei bedeutendsten Gruppen- bzw Halbgruppenoperationen auf* $\mathbb{R}$ *sind die* Translation *und die* Skalierung *die durch die beiden Operatoren* τ_y *und* σ_h*, definiert als*

$$\tau_y f = f(\cdot + y) \qquad \textit{und} \qquad \sigma_h f = f(h\cdot)$$

realisiert werden sollen.

Als nächstes stellen wir ein paar einfache Eigenschaften der Fouriertransformierten zusammen – daß die Fouriertransformierte linear in f bzw. c ist, das braucht ja wohl nicht mehr besonders betont werden.

Satz 6.13 *(Eigenschaften der Fouriertransformierten)*

1. *Für* $f \in L_1(\mathbb{R})$ *und* $y \in \mathbb{R}$ *ist*

$$(\tau_y f)^\wedge(\xi) = e^{iy\xi}\,\widehat{f}(\xi), \qquad \xi \in \mathbb{R}. \tag{6.9}$$

2. *Für* $f \in L_1(\mathbb{R})$ *und* $h \in \mathbb{R}$ *ist*

$$(\sigma_h f)^\wedge(\xi) = \frac{\widehat{f}\,(h^{-1}\xi)}{h}, \qquad \xi \in \mathbb{R}. \tag{6.10}$$

3. *Für* $f, g \in L_1(\mathbb{R})$ *bzw.* $c, d \in \ell_1(\mathbb{Z})$ *ist* $f * g \in L_1(\mathbb{R})$ *bzw.* $c * d \in \ell_1(\mathbb{Z})$ *und es gilt für* $\xi \in \mathbb{R}$

$$(f * g)^\wedge(\xi) = \widehat{f}(\xi)\,\widehat{g}(\xi), \qquad \textit{bzw.} \qquad (c * d)^\wedge(\xi) = \widehat{c}(\xi)\,\widehat{d}(\xi). \tag{6.11}$$

4. *Für* $f \in L_1(\mathbb{R})$ *und* $c \in \ell_1(\mathbb{Z})$ *ist* $f * c \in L_1(\mathbb{R})$ *und*

$$(f * c)^\wedge(\xi) = \widehat{f}(\xi)\,\widehat{c}(\xi), \xi \in \mathbb{R}. \tag{6.12}$$

5. *Sind* $f, f' \in L_1(\mathbb{R})$*, dann gilt*

$$\left(\frac{d}{dx} f\right)^\wedge(\xi) = i\xi\,\widehat{f}(\xi), \qquad \xi \in \mathbb{R}. \tag{6.13}$$

6. *Sind* $f, xf \in L_1(\mathbb{R})$*, dann ist* $\widehat{f}$ *differenzierbar und es gilt*

$$\frac{d}{d\xi}\widehat{f}(\xi) = (-ix\,f)^\wedge(\xi), \qquad \xi \in \mathbb{R}. \tag{6.14}$$

7. *Sind* $f, \widehat{f} \in L_1(\mathbb{R})$*, dann ist*

$$f(x) = \left(\widehat{f}\right)^\vee(x) := \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{f}(\vartheta)\,e^{ix\vartheta}\,d\vartheta \tag{6.15}$$

Die Operation $f \mapsto f^\vee$ *bezeichnet man als* inverse Fouriertransformation[145].

[145] Die Gründe sind ja wohl offensichtlich.

Beweis: Für 1) berechnen wir

$$\begin{aligned}(\tau_y f)^\wedge(\xi) &= \int_{\mathbb{R}} f(t+y)\, e^{-i\xi t}\, dt = \int_{\mathbb{R}} f(t)\, e^{-i\xi(t-y)}\, dt = e^{iy\xi} \int_{\mathbb{R}} f(t)\, e^{-i\xi t}\, dt \\ &= e^{iy\xi}\, \widehat{f}(\xi),\end{aligned}$$

während 2) ganz ähnlich mit

$$(\sigma_h f)^\wedge(\xi) = \int_{\mathbb{R}} f(ht)\, e^{-i\xi t}\, dt = \frac{1}{h} \int_{\mathbb{R}} f(t)\, e^{-i(\xi/h)t}\, dt = \frac{\widehat{f}\left(\frac{\xi}{h}\right)}{h}$$

bewiesen wird. Die erste Aussage von 3) folgt aus

$$\|f * g\|_1 = \int_{\mathbb{R}} \left| \int_{\mathbb{R}} f(t) g(s-t)\, dt \right| ds \le \int_{\mathbb{R}} \int_{\mathbb{R}} |f(t)g(s)|\, dt\, ds = \|f\|_1\, \|g\|_1$$

bzw.

$$\|c * d\|_1 = \sum_{j\in\mathbb{Z}} \left| \sum_{k\in\mathbb{Z}} c(k)\, d(j-k) \right| \le \sum_{j,k\in\mathbb{Z}} |c(k)\, d(j)| = \|c\|_1\, \|c\|_1,$$

der zweite, etwas interessantere Teil hingegen aus

$$\begin{aligned}(f * g)^\wedge(\xi) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(s) g(t-s)\, ds \right) e^{-i\xi t} dt = \int_{\mathbb{R}} \int_{\mathbb{R}} f(s)\, e^{i\xi s}\, g(t-s)\, e^{i\xi(t-s)}\, ds dt \\ &= \widehat{f}(\xi)\, \widehat{g}(\xi),\end{aligned}$$

bzw.

$$\begin{aligned}(c * d)^\wedge(\xi) &= \sum_{j\in\mathbb{Z}} \left(\sum_{k\in\mathbb{Z}} c(k)\, d(j-k) \right) e^{-ij\xi} = \sum_{j,k\in\mathbb{Z}} c(k)\, e^{-ik\xi}\, d(j-k)\, e^{-i(j-k)\xi} \\ &= \widehat{f}(\xi)\, \widehat{g}(\xi).\end{aligned}$$

4) folgt aus Lemma 6.4 und

$$\begin{aligned}(f * c)^\wedge(\xi) &= \int_{\mathbb{R}} \sum_{k\in\mathbb{Z}} f(t-k)\, c(k) e^{-i\xi t}\, dt = \int_{\mathbb{R}} \sum_{k\in\mathbb{Z}} f(t-k)\, e^{-i\xi(t-k)}\, c(k)\, e^{-ik\xi} \\ &= \widehat{f}(\xi)\, \widehat{c}(\xi).\end{aligned}$$

oder auch direkt unter Verwendung von (6.9). Für 5 verwenden wir partielle Integration[146], um

$$(f')^\wedge(\xi) = \int_{\mathbb{R}} \frac{df}{dt}(t) e^{-i\xi t}\, dt = -\int_{\mathbb{R}} f(t) \frac{d}{dt} e^{-i\xi t}\, dt = i\xi \int_{\mathbb{R}} f(t) e^{-i\xi t}\, dt = i\xi\, \widehat{f}(\xi).$$

[146] Daß dies gerechtfertigt ist, liegt an der Tatsache, daß für $f \in L_1(\mathbb{R})$ immer $\lim_{x\to\pm\infty} |f(x)| = 0$ sein muß und daß die stetigen Funktionen mit kompaktem Träger bezüglich der Norm $\|\cdot\|_1$ *dicht* in $L_1(\mathbb{R})$ sind. Deswegen muß man sich um "Randwerte" hier nicht kümmern.

6) erhalten wir, indem wir für $h > 0$ den Differenzenquotient

$$\frac{\widehat{f}\,(\xi+h) - \widehat{f}(\xi)}{h} = \int_{\mathbb{R}} f(t)\,\frac{e^{-i(\xi+h)t} - e^{-i\xi t}}{h}\,dt = \int_{\mathbb{R}} f(t)\,e^{-i\xi t}\,\frac{e^{-iht}-1}{h}\,dt$$

betrachten; das Integral existiert, weil $xf \in L_1(\mathbb{R})$ und da

$$\lim_{h\to 0}\frac{e^{-iht}-1}{h} = \lim_{h\to 0}(-it)\,e^{-iht} = -it$$

ist, folgt (6.14). Der Beweise von 7) ist ein klein wenig aufwendiger und verwendet die *Fejér–Kerne*

$$F_\lambda := \lambda F\,(\lambda\cdot)\,, \quad \lambda > 0, \qquad F(x) := \frac{1}{2\pi}\int_{-1}^{1}(1-|t|)\;e^{ixt}\,dt, \quad x \in \mathbb{R},$$

auf $\mathbb{R}$, die die Eigenschaft haben, daß für jedes $f \in L_1(\mathbb{R})$

$$\lim_{\lambda\to\infty}\|f - f * F_\lambda\| = 0, \tag{6.16}$$

siehe [35, S. 124–126], also auch $f * F_\lambda \to f$ punktweise fast überall[147]. Dann ist aber für $x \in \mathbb{R}$

$$\begin{aligned}
f * F_\lambda(x) &= \frac{1}{2\pi}\int_{\mathbb{R}} f\,(t)\,\left(\lambda\int_{-1}^{1}(1-|\vartheta|)\,e^{i(x-t)\lambda\vartheta}\,d\vartheta\right)\,dt\\
&= \frac{1}{2\pi}\int_{\mathbb{R}} f\,(t)\int_{-\lambda}^{\lambda}\left(1-\frac{|\vartheta|}{\lambda}\right)e^{i(x-t)\vartheta}\,d\vartheta\,dt\\
&= \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\vartheta|}{\lambda}\right)\underbrace{\int_{\mathbb{R}} f\,(t)\;e^{-it\vartheta}\,dt}_{=\widehat{f}(\vartheta)}\;e^{ix\vartheta}\,d\vartheta\\
&= \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\vartheta|}{\lambda}\right)\,\widehat{f}(\vartheta)\,e^{ix\vartheta}\,d\vartheta\\
&= \underbrace{\frac{1}{2\pi}\int_{|\vartheta|\le\sqrt{\lambda}}\overbrace{\left(1-\frac{|\vartheta|}{\lambda}\right)}^{1-1/\sqrt{\lambda}\le\,\cdot\,\le 1}\widehat{f}(\vartheta)\,e^{ix\vartheta}\,d\vartheta}_{\to\frac{1}{2\pi}\int_{\mathbb{R}}\widehat{f}(\vartheta)\,e^{ix\vartheta}\,d\vartheta} + \underbrace{\frac{1}{2\pi}\int_{|\vartheta|\ge\sqrt{\lambda}}\left(1-\frac{|\vartheta|}{\lambda}\right)\,\widehat{f}(\vartheta)\,e^{ix\vartheta}\,d\vartheta}_{\to 0},
\end{aligned}$$

weil $\widehat{f} \in L_1(\mathbb{R})$. □

Übung 6.4 Beweisen Sie *ohne* Verwendung von (6.13) die folgende Aussage: Sind $f, f' \in L_1(\mathbb{R})$, dann ist $(f')^\wedge(0) = 0$.
Hinweis: Partielle Integration. ♢

[147] Zumindest für eine Teilfolge, siehe [23, S. 96].

Beispiel 6.14 *Berechnen wir doch mal zu Übungszwecken so eine Fouriertransformierte, und zwar die der kardinalen B–Splines $N_j = \chi * \cdots * \chi$. Insbesondere ist also*

$$\widehat{N_0}(\xi) = \widehat{\chi}(\xi) = \int_{\mathbb{R}} \chi(t)\, e^{-i\xi t}\, dt = \int_0^1 e^{-i\xi t}\, dt = \left.\frac{e^{-i\xi t}}{-i\xi}\right|_{t=0}^1 = \frac{1-e^{-i\xi}}{i\xi}$$

und somit, nach (6.11),

$$\widehat{N_j}(\xi) = (\widehat{\chi}(\xi))^{j+1} = \left(\frac{1-e^{-i\xi}}{i\xi}\right)^{j+1}.$$

Übung 6.5 Die *zentrierten B–Splines* M_j, $j \in \mathbb{N}_0$, sind definiert als

$$M_j = \underbrace{\chi_{[-1/2,1/2]} * \cdots * \chi_{[-1/2,1/2]}}_{j+1}.$$

Zeigen Sie:

1. Diese Funktionen sind gerade: $M_j(-x) = M_j(x)$, $x \in \mathbb{R}$.
2. Für $j \in \mathbb{N}_0$ ist
$$\widehat{M_j}(\xi) = \left(\frac{\sin \xi/2}{\xi/2}\right)^{j+1}, \qquad \xi \in \mathbb{R}.$$

♢

Ist $f \in L_1(\mathbb{R})$, dann ist für $\xi, \eta \in \mathbb{R}$

$$\left|\widehat{f}\,(\xi+\eta) - \widehat{f}(\xi)\right| \le \int_{\mathbb{R}} |f(t)|\, \underbrace{\left|e^{-i\xi t}\right|}_{=1} \left|e^{-i\eta t} - 1\right|\, dt,$$

was auf der rechten Seite unabhängig von ξ ist und mit $\eta \to 0$ gegen Null konvergiert, denn für jedes $\varepsilon > 0$ gibt es ein $N > 0$, so daß

$$\int_{|t|>N} |f(t)|\, dt < \varepsilon$$

ist, während wir, durch Wahl eines hinreichend kleinen Wertes von η, die Funktion $|e^{-i\eta t} - 1|$ auf $[-N, N]$ so klein machen können, wie wir wollen. Der langen Rede kurzer Sinn:

> *Ist $f \in L_1(\mathbb{R})$, so ist $\widehat{f} \in C_u(\mathbb{R})$, dem Vektorraum der gleichmäßig stetigen und gleichmäßig beschränkten*[148] *Funktionen auf $\mathbb{R}$.*

Außerdem kann man sogar sagen, wie sich die Fouriertransformierte für $|\xi| \to \infty$ benimmt.

[148] Siehe (6.8).

Proposition 6.15 *(Riemann–Lebesgue–Lemma)*
Ist $f \in L_1(\mathbb{R})$, so ist

$$\lim_{\xi \to \pm\infty} \widehat{f}(\xi) = 0. \tag{6.17}$$

Beweis: Ist auch $f' \in L_1(\mathbb{R})$ so folgt (6.17) sofort mittels (6.13) und (6.8):

$$\|f'\|_1 \geq \left|(f')^\wedge(\xi)\right| = |\xi| \left|\widehat{f}(\xi)\right|, \qquad \xi \in \mathbb{R},$$

also $\left|\widehat{f}(\xi)\right| \leq \|f'\|_1 / |\xi| \to 0$ für $|\xi| \to \infty$. Für beliebiges $f \in L_1(\mathbb{R})$ und differenzierbares $g \in L_1(\mathbb{R})$ mit[149] mit $\|f - g\|_1 \leq \varepsilon$ ist

$$\|f - g\|_1 \geq \left|\widehat{f}(\xi) - \widehat{g}(\xi)\right| \geq \left|\widehat{f}(\xi)\right| - |\widehat{g}(\xi)|,$$

also

$$\lim_{|\xi| \to \infty} \left|\widehat{f}(\xi)\right| \leq \lim_{|\xi| \to \infty} |\widehat{g}(\xi)| + \|f - g\|_1 \leq \varepsilon$$

und da man ε beliebig klein wählen kann, folgt die Behauptung. □

Wie sieht es nun auf anderen L_p–Räumen, $p \neq 1$, insbesondere mit $L_2(\mathbb{R})$ aus[150]? Hier nutzt man aus, daß $L_1(\mathbb{R}) \cap L_p(\mathbb{R})$ eine *dichte Teilmenge* von $L_p(\mathbb{R})$ ist. Für L_2 gibt es nun noch eine besonders schöne Eigenschaft.

Satz 6.16 *(Parseval[151]/Plancherel)*
Für $f, g \in L_1(\mathbb{R}) \cap L_2(\mathbb{R})$ ist

$$\int_{\mathbb{R}} f(t)\, g(t)\, dt = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(\vartheta)\, \overline{\widehat{g}(\vartheta)}\, d\vartheta, \tag{6.18}$$

also insbesondere, mit $f = g$,

$$\|f\|_2 = \frac{1}{\sqrt{2\pi}} \left\|\widehat{f}\right\|_2. \tag{6.19}$$

Diese Aussage hilft uns nun, die Definition der Fouriertransformierten auf $L_2(\mathbb{R})$ zu übertragen: Zu $f \in L_2(\mathbb{R})$ betrachtet man eine Folge

$$f_n := \chi_{[-n,n]} \cdot f \quad \in L_1(\mathbb{R}) \cap L_2(\mathbb{R}), \qquad n \in \mathbb{N},$$

[149] Man beachte, daß sogar die *unendlich oft stetig differenzierbaren Funktionen mit kompaktem Träger* eine dichte Teilmenge von $L_1(\mathbb{R})$ bilden.

[150] $L_2(\mathbb{R})$ spielt in der Signalverarbeitung schon deswegen so eine wesentliche Rolle, weil das gerade die Signale (und die sind normalerweise nicht unbedingt stetig) mit *endlicher Energie* sind – eine ziemlich realistische Annahmen, oder nicht?

[151] Marc–Antoine Parseval des Chênes, 1755–1836, Zeitgenosse von Fourier, der ziemlich heftig in die Wirren der franzßischen Revolution verwickelt wurde, publizierte überhaupt nur 5 (in Worten: "fünf") Arbeiten, die er aber allesamt der *Académie des Sciences* vorlegte.

die für $n \to \infty$ in der Norm $\|\cdot\|_2$ gegen f konvergiert. Da

$$\left\| \widehat{f}_{n+k} - \widehat{f}_n \right\|_2 = \left\| (f_{n+k} - f)^\wedge \right\|_2 = \|f_{n+k} - f_n\|_2 , \qquad k, n \in \mathbb{N},$$

ist $\widehat{f}_n$ eine Cauchyfolge und konvergiert gegen eine Funktion in $L_2(\mathbb{R})$, die wir $\widehat{f}$ nennen wollen.
Beweis von Satz 6.16: Wir definieren

$$h(x) = \int_{\mathbb{R}} f(t)\, g\,(t-x)\; dt = (f * g(-\cdot))\,(x), \qquad x \in \mathbb{R},$$

und erhalten, daß $h(0) = \int fg$. Außerdem ist

$$\widehat{h}(\xi) = \widehat{f}(\xi)\; \underbrace{(g(-\cdot))^\wedge\,(\xi)}_{=\overline{\widehat{g}(\xi)}} = \widehat{f}(\xi)\,\overline{\widehat{g}(\xi)}, \qquad \xi \in \mathbb{R}.$$

Sind nun f und g so "brav", daß $\widehat{f}, \widehat{g} \in L_2(\mathbb{R})$ ist[152], dann ist mit (6.15)

$$\frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(\vartheta)\,\overline{\widehat{g}(\vartheta)}\, d\vartheta = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{h}(\vartheta)\, e^{i0\theta}\, d\vartheta = h(0) = \int_{\mathbb{R}} f(t)\, g(t)\, dt,$$

was (6.18) liefert. Und die *Plancherel–Identität* (6.19) ist dann eine unmittelbare Konsequenz aus der *Parseval–Formel* (6.18). □

Eine wichtige Identität haben wir aber noch – und auf ihr werden die wesentlichen Resultate der folgenden Abschnitte beruhen; außerdem werden wir sehen, daß sie auf sehr interessante Art die Fouriertransformierte und die Fourierreihe miteinander verbindet.

Satz 6.17 *(Poisson[153]–Summenformel)*
Für $f \in L_1(\mathbb{R})$ ist

$$\sum_{k\in\mathbb{Z}} f\,(2k\pi) = \frac{1}{2\pi} \sum_{k\in\mathbb{Z}} \widehat{f}\,(k) \qquad \textit{und} \qquad \sum_{k\in\mathbb{Z}} f(k) = \sum_{k\in\mathbb{Z}} \widehat{f}\,(2k\pi)\,. \tag{6.20}$$

Beweis: Wir setzen

$$g(x) = \sum_{k\in\mathbb{Z}} f\,(x + 2k\pi)\,, \qquad x \in \mathbb{R}, \tag{6.21}$$

und bemerken, daß $g\,(x + 2\pi) = g(x)$, also g eine 2π–periodische Funktion ist, und wegen

$$\begin{aligned} \|g\|_{\mathbb{T},1} &= \int_{\mathbb{T}} |g(t)|\; dt = \int_{\mathbb{T}} \left| \sum_{k\in\mathbb{Z}} f\,(t + 2k\pi) \right| dt \le \sum_{k\in\mathbb{Z}} \int_0^{2\pi} |f\,(t + 2k\pi)|\; dt \\ &= \int_{\mathbb{R}} |f\,(t)|\; dt = \|f\|_{\mathbb{R},1} \end{aligned}$$

[152] Das ist beispielsweise der Fall, wenn f und g differenzierbar sind; dies folgt aus (6.13) und dem Riemann–Lebesgue–Lemma, Proposition 6.15.

[153] Siméon Denis Poisson, 1781–1840, studierte bei Laplace und Legendre, Beiträge zur Fourier–Analysis und Wahrscheinlichkeitstheorie ("Poisson–Verteilung"), schrieb zwischen 300 und 400 Arbeiten, auch über Elektrizität, Magnetismus und Astronomie.

ist $g \in L_1(\mathbb{T})$ und insbesondere wohldefiniert – die Summe in (6.21) divergiert nicht allzu unmotiviert. Die Fourierkoeffizienten g_k von g haben nun nach (1.3) die Form

$$g_k = \frac{1}{2\pi}\int_{\mathbb{T}} g(t)\, e^{-ikt}\, dt = \frac{1}{2\pi}\int_{\mathbb{T}} \sum_{\ell\in\mathbb{Z}} f\left(t+2\ell\pi\right)\, e^{-ikt}\, dt = \frac{1}{2\pi}\int_{\mathbb{R}} f(t)\, e^{-ikt}\, dt = \frac{1}{2\pi}\,\widehat{f}(k)$$

und, angenommen die Partialsummen der Fourierreihe von g würden konvergieren[154], erhalten so, daß

$$\frac{1}{2\pi}\sum_{k\in\mathbb{Z}} \widehat{f}(k) = \sum_{k\in\mathbb{Z}} g_k \underbrace{e^{ik0}}_{=1} = g(0) = \sum_{k\in\mathbb{Z}} f\left(0+2k\pi\right) = \sum_{k\in\mathbb{Z}} f\left(2k\pi\right),$$

was die erste Identität liefert. Mit deren Hilfe und (6.10) ergibt sich dann, daß

$$\sum_{k\in\mathbb{Z}} f(k) = \sum_{k\in\mathbb{Z}} \left(\sigma_{(2\pi)^{-1}} f\right)\left(2k\pi\right) = \frac{1}{2\pi}\,\sum_{k\in\mathbb{Z}} \left(\sigma_{(2\pi)^{-1}} f\right)^{\wedge}(k) = \sum_{k\in\mathbb{Z}} \widehat{f}\left(2k\pi\right).$$

□

Schließlich werden wir die Fouriertransformation noch nutzen, um einen etwas anderen Begriff der Differenzierbarkeit einzuführen, der nicht auf Eigenschaften der Funktion, sondern auf Eigenschaften der Fouriertransformierten beruht. Um diese Idee zu motivieren, betrachten wir einmal eine Funktion $f \in L_1(\mathbb{R})$, deren Ableitung ebenfalls zu $L_1(\mathbb{R})$ gehört. Da, nach (6.13), $\left(f'\right)^{\wedge}(\xi) = i\xi\widehat{f}(\xi)$ und nach dem Riemann–Lebesgue–Lemma

$$0 = \lim_{\xi\to\pm\infty} \left|\left(f'\right)^{\wedge}(\xi)\right| = \lim_{\xi\to\pm\infty} \left|i\xi\widehat{f}(\xi)\right| = \lim_{\xi\to\pm\infty} |\xi|\, \left|\widehat{f}(\xi)\right|$$

ist, muß die (gleichmäßig stetige) Funktion $\widehat{f}$ also die Eigenschaft

$$\left|\widehat{f}(\xi)\right| \le C_f\left(1+|\xi|\right)^{-1}, \qquad \xi\in\mathbb{R},$$

für eine Konstante $C_f > 0$ haben[155] – auf kompakten Intervallen ist ja $\left|\widehat{f}(\cdot)\right|$ beschränkt und außerhalb davon schlägt die asymptotische Aussage des Riemann–Lebesgue–Lemma durch. Mit anderen Worten:

Ist $f \in L_1(\mathbb{R})$ differenzierbar, so ist

$$\left\|\left(1+|\cdot|\right)\,\widehat{f}(\cdot)\right\|_{\infty} < \infty.$$

Wie aber sieht's mit der Umkehrung aus? Verlangen wir ein bißchen "mehr", nämlich, daß

$$\left\|\left(1+|\cdot|\right)^{\alpha}\,\widehat{f}(\cdot)\right\|_{1} < \infty, \qquad \alpha > 0,$$

[154] Ansonsten müssten wir ein Summationsverfahren, beispielsweise den Féjer–Kern verwenden.

[155] Man beachte: Hier hängt die Konstante sehr wohl von f ab, nur eben nicht von ξ!

dann ist, unter Verwendung der *inversen Fouriertransformation* (6.15), nun plötzlich $f \in C_u(\mathbb{R})$, ja sogar $f \in C_u^\alpha(\mathbb{R})$. Trotzdem passen die beiden Enden wieder einmal nicht zusammen, es sei denn, wir verwenden wieder die Plancherel–Identität (6.18) aus Satz 6.16 und die Tatsache, daß die Fouriertransformierte auf $L_2(\mathbb{R})$ eine *Isomorphie*[156] ist, um Differenzierbarkeit für L_2–Funktionen über die Fouriertransformierte zu definieren.

Definition 6.18 *Für $\alpha > 0$ definieren die* Sobolev[157]–Klassen

$$W_2^\alpha(\mathbb{R}) = \left\{ f \in L_2(\mathbb{R}) \ : \ (1 + |\cdot|)^\alpha \left| \widehat{f}(\cdot) \right| \in L_2(\mathbb{R}) \right\}. \tag{6.22}$$

Das minimale $\alpha > 0$ für das (6.22) für ein gegebenes $f \in L_2(\mathbb{R})$ noch erfüllt ist, bezeichnet man als kritischen Index *von f.*

Einen großen Unterschied gibt es aber zwischen der klassischen Differenzierbarkeit und der Sobolev–Differenzierbarkeit: Erstere ist eine *lokale* Bedingung an die Funktion, während letztere, wegen des Übergangs zur Fouriertransformierten eine *globale* Eigenschaft der Funktion f ist.

6.3 Polynomreproduktion und die Strang–Fix–Bedingungen

In diesem Abschnitt befassen wir uns mit der Frage, wie gut man Funktionen mit den Räumen V_h approximieren kann, und zwar in Abhängigkeit vom "Diskretisierungsparamter" $h > 0$. Es ist ja noch ziemlich naheliegend, daß für $h \to \infty$, also immer feinere Abtastung, die Ergebnisse wohl immer besser werden und daß der Approximationsfehler sogar gegen Null gehen wird, aber wir werden auch feststellen, daß dies umso schneller geht, je glatter die zu approximierende Funktion ist – varausgesetzt, unser Raum $\mathbb{S}(\varphi)$ ist imstande, Polynome eines gewissen Grades zu (re-)produzieren. Daß man dafür wirklich $\mathbb{S}(\varphi)$ braucht und nicht mit dem Teilraum $\mathbb{S}_p(\varphi)$ auskommt, zeigt Übung 6.6.

Übung 6.6 Zeigen Sie: Hat $\varphi \in L_p(\mathbb{R})$ kompakten Träger, dann ist $\mathbb{S}_p(\varphi) \cap \Pi = \{0\}$. ♢

Wir werden die Approximationsordnung in zwei Schritten angehen: Zuerst werden wir untersuchen, wann die Translate unserer Funktio φ imstande sind, Polynome von einem bestimmten Grad zu (re-)produzieren und dann werden wir uns ansehen, welche Folgen das für die Approximationseigenschaften der Räume $\sigma_{1/h}\mathbb{S}(\varphi)$ für $h \to 0$ hat. Das mit dem "(re-)produzieren" hat übrigens einen einfachen Grund, mit dem wir auch beginnen wollen.

Proposition 6.19 *Ist $\varphi \in C_{00}(\mathbb{R})$ eine stetige Funktion mit kompaktem Träger*[158] *Ist $\Pi_n \subset$*

[156] Zumindest bis auf die Konstante $\sqrt{2\pi}^{-1}$.

[157] Sergei Lvovich Sobolev, 1908–1989, eine der ganz herausragenden Gestalten in der Theorie der partiellen Differentialgleichungen.

[158] Auch wenn wir in der Norm von L_2 approximieren, werden die Funktionen, mit denen wir das tun wollen, doch zumeist zu $C_{00}(\mathbb{R})$ gehören; wenn diese Funktionen z.B. inverse Fouriertransformationen sein wollen, dann bleibt ihnen ja sowieso nichts anderes übrig, als gleichmäßig stetig zu sein und wenn man "richtig" mit ihnen rechnen will, dann muß man entweder kompakten Träger oder ziemlich flottes Abklingen fordern.

$\mathbb{S}(\varphi)$, *dann ist für jedes* $p \in \Pi_n$ *auch*[159]

$$\varphi * p = \sum_{j\in\mathbb{Z}} \varphi(\cdot - j)\ p(j) \quad \in \Pi_n \qquad \textit{und} \qquad \deg \varphi * p \le \deg p. \tag{6.23}$$

Mit anderen Worten: "Produktion" und "Reproduktion" von Polynomen sind beinahe äquivalent.

Beweis: Sei $\Pi_n \ni p = \varphi * c$ für ein $c \in \ell(\mathbb{Z})$, das ja existieren muß, weil $\Pi_n \subset \mathbb{S}(\varphi)$. Dann ist

$$\begin{aligned}
\varphi * p &= \sum_{j\in\mathbb{Z}} \varphi(\cdot - j)\ p(j) = \sum_{j\in\mathbb{Z}} \varphi(\cdot - j) \sum_{k\in\mathbb{Z}} \varphi(j-k)\, c(k) \\
&= \sum_{j,k\in\mathbb{Z}} \varphi(\cdot - j - k)\ \varphi(j)\, c(k) = \sum_{j\in\mathbb{Z}} \varphi(j) \sum_{k\in\mathbb{Z}} \underbrace{\varphi(\cdot - j - k)\ c(k)}_{\varphi * c(\cdot - j) = p(\cdot - j)} \\
&= \sum_{j\in\mathbb{Z}} \varphi(j)\, p(\cdot - j).
\end{aligned} \tag{6.24}$$

Ist nun $p \in \Pi_n$, so ist auch $p(\cdot - j)$ ein Polynom[160] vom selben Grad wie p in Π_n und da φ kompakten Träger hat, ist die Summe auf der rechten Seite eine *endliche* Linearkombination von Polynomen vom Grad $\le \deg p$, also wieder in ein Polynom vom Grad $\le \deg p$. □

Was natürlich schön wäre, das wäre $\deg \varphi * p = \deg p$ in (6.23), nur leider können wir das nicht erwarten, wie das folgende Beispiel zeigt.

Beispiel 6.20 *Sei*

$$\varphi := \begin{cases} -1, & x \in [-1, 0), \\ 1, & x \in [0, 1), \\ 0, & sonst, \end{cases}$$

dann ist

$$2 = \sum_{j\in\mathbb{Z}} (-1)^j\, \varphi(\cdot - j),$$

also $\Pi_0 \subset \mathbb{S}(\varphi)$*, aber*

$$\varphi * 1 = \sum_{j\in\mathbb{Z}} \varphi(\cdot - j) = 0,$$

der Grad wird also echt "kleiner"! Nun gut, dieses Funktion φ *ist ja auch nicht stetig, aber erstens wurden im obigen Beweis ja eigentlich nur kompakter Träger und* gleichmäßige Beschränktheit *von* φ *verwendet und zweitens kann man natürlich auch Beispiele höherer Ordnung angeben. Die sind halt dann bloß nicht mehr so einfach.*

[159] Die Faltung in (6.23) ist zuerst einmal mehrdeutig! Da es aber nur genau ein Polynom gibt (welches?!), das zu $L_p(\mathbb{R})$ für irgendein $1 \le p < \infty$ gehört, wollen wir Polynome immer nur im "diskreten Sinne" mit Funktionen falten, also Polynome p hier als die "Vektoren" oder Folgen $(p(j)\ :\ j \in \mathbb{Z}) \in \ell(\mathbb{Z})$ auffassen.

[160] Ja, die Polynome von einem bestimmten Höchstgrad bilden einen *endlichdimensionalen* translationsinvarianten Raum – wer häte gedacht, daß es so was gibt?

Übung 6.7 Zu der Funktion φ aus Beispiel 6.20 betrachten wir $\psi = \chi * \varphi$. Zeigen Sie:

1. $\Pi_1 \subset \mathbb{S}(\psi)$.
2. Es gibt eine Funktion $p \in \Pi_1$ mit $\deg p = 1$ und $\deg \psi * p < 1$.

♢

Und trotzdem brauchen wir gar nicht so viel mehr von φ zu fordern, nämlich nur, daß das, was in Beispiel 6.20 gerade schiefgegangen ist, eben *nicht* passiert.

Korollar 6.21 *Erfüllt φ neben den Voraussetzungen von Proposition 6.19 auch noch*

$$0 \neq \varphi * 1(0) = \sum_{j\in\mathbb{Z}} \varphi(j), \tag{6.25}$$

so ist $\deg \varphi * p = \deg p$ *für alle* $p \in \Pi_n$.

Beweis: Nehmen wir der Einfachheit halber an, daß $p(x) = x^k + \cdots$, $k \leq n$, was man durch geeignete Normierung ja immer erreichen kann. Daß $\deg \varphi * p \leq \deg p = k$ ist, das wissen wir ja schon aus Proposition 6.19. Würde aber die strikte Ungleichung "$<$" gelten, so ergibt (6.24) daß

$$0 = (\varphi * p)^{(k)} = \sum_{j\in\mathbb{Z}} \varphi(j)\ \underbrace{p^{(k)}(\cdot - j)}_{=k!} = k! \sum_{j\in\mathbb{Z}} \varphi(j) \neq 0,$$

was natürlich einen Widerspruch darstellt. □

Definition 6.22 *Eine Funktion $f \in L_1(\mathbb{R})$ erfüllt die* Strang–Fix–Bedingungen *der Ordnung $r \geq 0$ wenn*

1. $\widehat{f}(0) \neq 0$.
2. *für $j = 0, \ldots, r$ ist*
$$\widehat{f}^{(j)}(2k\pi) = 0, \qquad k \in \mathbb{Z} \setminus \{0\}. \tag{6.26}$$

Bemerkung 6.23 *Diese Bedingungen an die Fouriertransforierte einer Funktion wurden erstmals von Schoenberg*[161] *in [74] aufgestellt und untersucht, ihren Namen haben sie jedoch von der Arbeit [80]*[162]

Satz 6.24 *Erfüllt die Funktion $\varphi \in C_{00}(\mathbb{R})$ die Strang–Fix–Bedingungen der Ordnung n, dann ist $\Pi_n \subset \mathbb{S}(\varphi)$.*

[161] Isaac J. Schoenberg, 1903–1990, Studium der Mathematik in Berlin und Göttingen, befasste sich unter anderem mit analytischer Zahlentheorie, totaler Positivität, isometrischen Einbettungen metrischer Räume in Hilberträume. In der Numerik als "Vater der Splines" am bekanntesten. Schoenberg war mit Landaus Tochter Charlotte verheiratet, seine Schwester mit Hans Rademacher.

[162] Es ist bemerkenswert, daß dies eine der wenigen mathematischen Arbeiten ist, deren Autoren *nicht* in alphabetischer Reihenfolge aufgeführt werden.

Wir werden Satz 6.24 sogar in einer etwas allgemeineren Form beweisen, indem wir Erhaltung polynomialer Räume für solche translationsivarianten Räume beweisen, die (6.25) erfüllen.

Satz 6.25 *Sei $\varphi \in C_{00}(\mathbb{R})$ und $\varphi * 1(0) \neq 0$. Dann ist*

$$\Pi_n \subseteq \mathbb{S}(\varphi) \qquad \Longleftrightarrow \qquad \widehat{\varphi}^{(j)}\,(2k\pi) = 0, \quad j = 0, \ldots, n,\ k \in \mathbb{Z} \setminus \{0\}. \tag{6.27}$$

Satz 6.24 folgt nun aus Satz 6.25 als unmittelbare Anwendung der Poissonschen Summenformel (6.20), denn erfüllt φ auch nur die Strang–Fix–Bedingungen der Ordnung 0, so ist

$$\varphi * 1 = \sum_{k\in\mathbb{Z}} \varphi(k) = \sum_{k\in\mathbb{Z}} \widehat{\varphi}(2k\pi) = \underbrace{\widehat{\varphi}(0)}_{\neq 0} + \sum_{k\in\mathbb{Z}\setminus\{0\}} \underbrace{\widehat{\varphi}(2k\pi)}_{=0} \neq 0 \tag{6.28}$$

und Satz 6.24 ist gerade die Richtung "$\Longleftarrow$" von Satz 6.25.
Beweis von Satz 6.25: Beginnen wir mit der Richtung "$\Longrightarrow$", für die wir die Voraussetzung $\varphi * 1 \neq 0$ noch nicht einmal brauchen werden. Dabei betrachten wir für $j = 0, \ldots, n$ und festes $x \in \mathbb{R}$ die Funktion $\psi(t) = (-t)^j\,\varphi(x+t)$ und erhalten, da $\psi \in C_{00}(\mathbb{R})$, über die Poissonsche Summenformel (6.20), (6.9) und (6.14), daß

$$\begin{aligned} p(x) \;&:=\; \sum_{k\in\mathbb{Z}} k^j \varphi(x-k) = \sum_{k\in\mathbb{Z}} \psi(k) = \sum_{k\in\mathbb{Z}} \widehat{\psi}(2k\pi) = \sum_{k\in\mathbb{Z}} \left((-\cdot)^j\,\varphi(\cdot + x)\right)^{\wedge}(2k\pi) \\ &=\; \sum_{k\in\mathbb{Z}} i^j\,\frac{d^j}{d\xi^j}\left(\widehat{\varphi}(\xi)e^{ix\xi}\right)(2k\pi) = \sum_{k\in\mathbb{Z}} i^j \sum_{\ell=0}^{j} \binom{j}{\ell} \widehat{\varphi}^{(\ell)}(2k\pi)\;\; (ix)^{j-\ell}\; e^{2i\pi kx}, \end{aligned}$$

also

$$\sum_{k\in\mathbb{Z}} k^j \varphi(x-k) = \sum_{k\in\mathbb{Z}} i^j \sum_{\ell=0}^{j} \binom{j}{\ell} \widehat{\varphi}^{(\ell)}(2k\pi)\;\; (ix)^{j-\ell}\; e^{2i\pi kx}, \qquad x \in \mathbb{R}. \tag{6.29}$$

Per Induktion über j können wir annehmen, daß alle Terme mit $\ell < j$ in der Summe auf der rechten Seite verschwinden[163], und so erhalten wir, daß für jedes $x \in \mathbb{R}$

$$p(x) = \sum_{k\in\mathbb{Z}} i^j\,\widehat{\varphi}^{(j)}(2k\pi)\,e^{2i\pi kx}$$

ist. Die Funktion auf der linken Seite ist ein Polynom[164] in x nach Proposition 6.19, was auf der rechten Seite steht hingegen periodisch mit Periode 1, da $e^{2ik\pi} = 1$ für alle $k \in \mathbb{Z}$ und deswegen bleibt p nichts anderes übrig, als konstant zu sein, was aber gerade $\widehat{\varphi}^{(j)}(2k\pi) = 0$ für $k \neq 0$ impliziert.

[163] Der Induktionsanfang, $j = 0$, ist trivialerweise erfüllt, denn dann haben wir einfach keine Bedingung.
[164] Vom Grad $\leq n$, aber das ist eher sekundär.

Für die andere Richtung, "$\Longleftarrow$", verwenden wir nochmals (6.29) und erhalten, nach Einsetzen der Strang–Fix–Bedingungen $\widehat{\varphi}^{(j)}(2k\pi) = 0$, $k \in \mathbb{Z} \setminus \{0\}$, daß

$$\begin{aligned}
&\sum_{k\in\mathbb{Z}} k^j \varphi(x-k) \\
&= \; i^j \sum_{\ell=0}^{j} \binom{j}{\ell} \widehat{\varphi}^{(\ell)}(0) \; (ix)^{j-\ell} + \sum_{k\in\mathbb{Z}\setminus\{0\}} i^j \sum_{\ell=0}^{j} \binom{j}{\ell} \underbrace{\widehat{\varphi}^{(\ell)}(2k\pi)}_{=0} \; (ix)^{j-\ell} \; e^{2i\pi kx} \\
&= \; \sum_{\ell=0}^{j} \underbrace{i^{j+\ell} \binom{j}{\ell} \widehat{\varphi}^{(j-\ell)}(0)}_{=:c_\ell} \; x^\ell = \sum_{\ell=0}^{j} c_\ell \, x^\ell \quad \in \Pi_j,
\end{aligned}$$

und da $c_j = (-1)^j \widehat{\varphi}(0) \neq 0$ ist[165], ist $\deg p = j$. Damit sind aber die Polynome

$$p_j := \sum_{k\in\mathbb{Z}} k^j \varphi(\cdot - k), \qquad j = 0, \ldots, n,$$

linear unabhängig und bilden demzufolge eine Basis von Π_n. Insbesondere gilt dann aber auch $\Pi_n \subset \mathbb{S}(\varphi)$. □

Beispiel 6.26 *Hier ein paar Beispiele für Funktionen, die die Strang–Fix–Bedingungen erfüllen (oder auch nicht):*

1. *Der (kardinale) B–Spline N_j erfüllt eine Strang–Fix–Bedingung der Ordnung j. Da*

$$\widehat{N_j}(\xi) = \left(\frac{1 - e^{-i\xi}}{i\xi} \right)^{j+1} =: \psi^{j+1}(\xi), \qquad \textit{also} \qquad \widehat{N_j}(0) = 1$$

mit $\psi(2k\pi) = 0$, hat $\widehat{N_j}$ an $2k\pi$, $k \in \mathbb{Z} \setminus \{0\}$ eine Nullstelle der Ordnung j, also

$$\widehat{N_j}^{(\ell)}(2k\pi) = 0, \qquad \ell = 0, \ldots, j, \quad k \in \mathbb{Z} \setminus \{0\}.$$

Damit erfüllen die B–Splines die Strang–Fix–Bedingungen und besitzen die Fähigkeit zur Polynomreproduktion.

2. *Wie sieht es nun mit der Funktion φ aus Beispiel 6.20 aus? Da wir in L_1 auch $\varphi = \chi - \chi(\cdot + 1)$ schreiben können[166], erhalten wir, daß*

$$\widehat{\varphi}(\xi) = \widehat{\chi}(\xi) - e^{i\xi}\widehat{\chi}(\xi) = \left(1 - e^{i\xi}\right) \left(\frac{1 - e^{-i\xi}}{i\xi} \right).$$

An den Stellen $2k\pi$ verschwinden also tatsächlich $\widehat{\varphi}$, was aber für die Probleme sorgt, ist die Tatsache, daß hier auch $\widehat{\varphi}(0) = 0$ ist.

[165] Hier gehen wieder die Annahme $\varphi * 1 \neq 0$ und die Poissonsche Summationsformel ein, genauer, Gleichung (6.28).

[166] Funktionen in L_p–Räumen sind nur bis auf Menge vom Maß Null eindeutig bestimmt!

3. *Das ändert sich auch nicht groß, wenn wir in Verallgemeinerung von Übung 6.7 die immer glatteren Funktionen*

$$\psi_j := \underbrace{\chi * \cdots * \chi}_{j} * \varphi = N_{j-1} * \varphi, \qquad j \in \mathbb{N}, \qquad \psi_0 = \varphi, \tag{6.30}$$

einführen, also

$$\begin{aligned} \widehat{\psi_j}(\xi) &= \widehat{\varphi}(\xi)\, \widehat{N_{j-1}}(\xi) = \left(1 - e^{i\xi}\right) \left(\frac{1 - e^{-i\xi}}{i\xi}\right) \left(\frac{1 - e^{-i\xi}}{i\xi}\right)^j \\ &= \left(1 - e^{i\xi}\right) \left(\frac{1 - e^{-i\xi}}{i\xi}\right)^{j+1} = \left(1 - e^{i\xi}\right)\, \widehat{N_j}(\xi), \end{aligned}$$

die sogar

$$\widehat{\psi_j}^{(\ell)}(2k\pi) = 0, \qquad \ell = 0, \ldots, j+1, \quad k \in \mathbb{Z} \setminus \{0\}$$

erfüllen, aber wegen $\widehat{\psi_j}(0) = 0$ *knapp an den Strang–Fix–Bedingungen scheitern.*

Bemerkung 6.27 *Die Funktionen* ψ_j, $j \in \mathbb{N}_0$, *aus (6.30), die, wie man sich leicht überzeugt, stückweise Polynome vom Grad* $\le j$ *mit globaler Differenzierbarkeitsordnung* $j-1$ *sind*[167]*, sind auch Beispiele, daß die "Komponente"* $\widehat{\varphi}(0) \neq 0$ *der Strang–Fix–Bedingungen auch* notwendig *ist: Die anderen Bedingungen an den Stellen* $2k\pi$, $k \in \mathbb{Z} \setminus \{0\}$ *erfüllt nämlich* ψ_j *sogar von der Ordnung* $j+1$ *und müßte ja dann sogar Polynome der Ordnung* $j+1$ *erzeugen können – etwas abwegig, wie man sich leicht klarmachen kann, wenn man beispielsweise versucht, die Funktion* x *durch stückweise konstante Funktionen zu kombinieren. Wer's lieber bewiesen hat: Da* ψ_j *stückweise zu* Π_j *gehört, ist für ein* x *im Inneren dieser "Stücke"*[168] $\psi_j^{(j+1)}(x-k) = 0$ *für alle* $k \in \mathbb{Z}$*, also auch*

$$(\psi_j * c)^{(j+1)}(x) = \sum_{k \in \mathbb{Z}} \underbrace{\psi_j^{(j+1)}(x-k)}_{=0}\, c_k = 0, \qquad c \in \ell(\mathbb{Z}),$$

weswegen es kein $c \in \ell(\mathbb{Z})$ *geben kann, so daß* $x^{j+1} = \psi * c(x)$, $x \in \mathbb{R}$.

Damit haben wir also das Problem, ob der Raum $\mathbb{S}(\varphi)$ Polynome einer bestimmten Ordnung enthält oder nicht, ziemlich erschöpfend abgehandelt und über die Fouriertransformierte von φ beschrieben. Es gibt natürlich auch detailliertere Beschreibungen, die mit schwächeren Voraussetzungen als $\varphi \in C_{00}(\mathbb{R})$ auskommen, siehe z.B. [34], aber im Prinzip lebt alles davon, daß man die Poissonsche Summenformel auf geeignete Art und Weise anwenden kann.

[167] Es gilt also $\psi_j \in C_{00}^{j-1}(\mathbb{R})$!

[168] Die ganzzahlige Intervalle sind!

6.4 Approximationsordnung

Der Nutzen der Polynomerhaltung bei der Bestimmung der Approximationsgüte wird nun darin liegen, daß wir die Tatsache ausnutzen, daß für eine hinreichend oft differenzierbare Funktion f das *Taylorpolynom*

$$T_n f = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!} \, (\cdot - x)^k$$

der Ordnung $n \in \mathbb{N}_0$ eine gute *lokale* Approximation der Funktion darstellt. Kann nun das Taylorpolynom durch $\mathbb{S}(\varphi)$ *exakt* dargestellt werden, so muß nur noch dieser lokal recht kleine Fehler approximiert werden und wenn man das mit der Lokalität von φ kombiniert (das heißt, mit dem kompakten Träger von φ), dann kommen wir zu guten Abschätzungen für den Fehler bei der Approximation aus $\mathbb{S}(\varphi)$. Nun, das ist die hehre Idee, bei deren Realisierung wir uns natürlich schon noch mit den Details herumschlagen müssen.

Bemerkung 6.28 *Mit*

$$L_p^n(\mathbb{R}) = \left\{ f \in L_p(\mathbb{R}) \ : \ f^{(n)} \in L_p(\mathbb{R}) \right\}$$

bezeichnen wir die Funktionen, die eine Ableitung in L_p besitzen. Bezüglich der Norm

$$\|f\|_{p,n} = \sum_{j=0}^{n} \left\| f^{(j)} \right\|_p$$

ist $L_p^n(\mathbb{R})$ dann ein Banachraum, in dem $C_{00}^n(\mathbb{R})$ dicht[169] *liegt.*

Nun können wir unser Approximationsresultat auch "schon" formulieren.

Satz 6.29 *Die Funktion $\varphi \in C_{00}(\mathbb{R})$ erfülle die Strang–Fix–Bedingungen der Ordnung n. Dann gibt es zu jedem $f \in L_p^{n+1}(\mathbb{R})$ und $h > 0$ ein $c_h \in \ell(\mathbb{Z})$, so daß*

$$\left\| f - \sigma_{1/h} \left(\varphi * c_h \right) \right\|_p \le C \, h^{n+1} \left\| f^{(n+1)} \right\|_p . \tag{6.31}$$

Wir können das auch noch anders sagen: Erfüllt φ die Strang–Fix–Bedingungen, dann erlaubt $\mathbb{S}(\varphi)$ sehr gute Approximation an zwar nicht alle Funktionen[170], aber doch an differenzierbare Funktionen. Und solche Sätze habe wir schon kennengelernt: Satz 6.29 ist nämlich nichts anderes als das Gegenstück zu den *Jackson–Sätzen* für differenzierbare Funktionen, siehe Proposition 5.19.

Wir werden einen *linearen* Operator, der (6.31) erfüllt sogar *explizit* konstruieren, und zwar einen sogenannten *Quasiinterpolanten*[171] der Form

$$Q_h f := \sum_{k \in \mathbb{Z}} \lambda_k \left(\sigma_h f \right) \, \varphi \left(h^{-1} \cdot - k \right) =: \left(\varphi * \lambda \right) \left(h^{-1} \cdot \right) \left(\sigma_h \, f \right), \qquad h > 0,$$

[169] Natürlich wieder bezüglich der Norm $\| \cdot \|_{p,k}$.

[170] Wer würde das auch erwarten?

[171] Oder, wie G. G. Lorentz mal gesagt haben soll (sinngemäß): "Entweder der Operator interpoliert, dann ist er ein Interpolant, oder er tut's nicht, dann ist er kein Interpolant. Was also ist ein Quasiinterpolant?"

wobei die λ_k, $k \in \mathbb{Z}$, noch zu bestimmende lineare Funktionale sind. Unser erstes Ziel besteht darin, diejenigen Funktionale zu bestimmen, die dafür sorgen, daß Q_h, eingeschränkt auf Π_n die Identität ist, das heißt, daß $Q_h p = p$, $p \in \Pi_n$, natürlich immer vorausgesetzt, daß φ die Strang–Fix–Bedingungen der Ordnung n erfüllt. Unter dieser Voraussetzung garantieren Satz 6.24 und Proposition 6.19 die Existenz von Polynomen $p_j \in \Pi_j$, $j = 0, \ldots, n$, so daß

$$\varphi * p_j(x) = x^j, \qquad x \in \mathbb{R}. \tag{6.32}$$

Und mit Hilfe dieser Polynome setzen wir unseren Quasiinterpolanten zusammen.

Lemma 6.30 *Erfüllt φ die Strang–Fix–Bedingungen der Ordnung n, dann gilt für die Funktionale*

$$\lambda_k(f) := \sum_{j=0}^{n} \frac{f^{(j)}(k)}{j!}\, p_j(0), \qquad f \in C^n(\mathbb{R}), \quad k \in \mathbb{Z}, \tag{6.33}$$

daß

$$Q_h p := \varphi * \lambda(p) = p, \qquad p \in \Pi_n. \tag{6.34}$$

Beweis: Es genügt, denn Fall $h = 1$ zu betrachten: Ist nämlich $Q_1 p = \varphi * \lambda(p) = p$, so ist natürlich auch

$$Q_h\, \sigma_{h^{-1}} p = (\varphi * \lambda)\left(h^{-1}\cdot\right)(\sigma_h \sigma_{h^{-1}} p) = (Q_1 p)\left(h^{-1}\cdot\right) = \sigma_{h^{-1}} p,$$

und ersetzt man x durch hx in dieser Gleichung, so folgt $Q_h p = p$.
Für $\ell \in \mathbb{Z}$ und $j = 0, \ldots, n$ ist

$$(x-\ell)^j = (\varphi * p_j)(\cdot - \ell) = \sum_{k\in\mathbb{Z}} \varphi(\cdot - \ell - k)\; p_j(k) = \sum_{k\in\mathbb{Z}} \varphi(\cdot - k)\; p_j(k-\ell) \tag{6.35}$$

Nun hat aber jedes $p \in \Pi_n$ an der Stelle ℓ die (endliche) Taylorentwicklung

$$p(x) = \sum_{j=0}^{n} \frac{p^{(j)}(\ell)}{j!}\; (x-\ell)^j \tag{6.36}$$

Setzen wir nun (6.35) in (6.36) ein, dann erhalten wir, daß

$$\begin{aligned} p(x) &= \sum_{j=0}^{n} \frac{p^{(j)}(\ell)}{j!} \sum_{k\in\mathbb{Z}} \varphi(\cdot - k)\; p_j(k-\ell) = \sum_{k\in\mathbb{Z}} \varphi(\cdot - k) \underbrace{\sum_{j=0}^{n} \frac{p^{(j)}(\ell)}{j!}\, p_j(k-\ell)}_{=:q_\ell(k)} \\ &= \varphi * q_\ell(x), \end{aligned}$$

wobei $q_\ell \in \Pi_n$. Nach Übung 6.8 gilt dann für beliebige $\ell, \ell' \in \mathbb{Z}$, daß $q_\ell = q_{\ell'}$, also insbesondere

$$q_0(k) = q_k(k) = \sum_{j=0}^{n} \frac{p^{(j)}(k)}{j!}\, p_j(0), \qquad k \in \mathbb{Z},$$

woraus unmittelbar $p = \varphi * q_0 = Q_1 p$, also (6.34) folgt. □

Übung 6.8 Zeigen Sie: Erfüllt φ die Strang–Fix–Bedingungen der Ordnung n, dann gilt für $p \in \Pi_n$

$$\varphi * p \equiv 0 \qquad \Longleftrightarrow \qquad p = 0.$$

♢

Beweis von Satz 6.29: Es genügt, zu fordern, daß $f \in C_{00}^{n+1}(\mathbb{R})$ ist; das folgt aus Bemerkung 6.28 und der Stetigkeit der Normen auf beiden Seiten von (6.31), die ja beide durch $\|\cdot\|_{p,n+1}$ majorisiert werden. Außerdem nehmen wir wieder an, daß $\operatorname{supp}\varphi \subseteq [0, N]$. Für $x \in \mathbb{R}$ und $h > 0$ sei

$$T_n f := \sum_{j=0}^{n} \frac{f^{(j)}(x)}{j!} \, (\cdot - x)^j$$

das Taylor–Polynom der Ordnung n an f bezüglich der Stelle x. Dann ist, da $f(x) = T_n f(x) = Q_h \left(T_n f\right)(x)$

$$\begin{aligned} |f(x) - Q_h f(x)| &= \left| Q_h (\underbrace{f - T_n f}_{=:g})(x) \right| = |Q_h g(x)| = \left| \sum_{k \in \mathbb{Z}} \lambda_k \left(g(h\cdot)\right) \, \varphi\left(h^{-1}x \cdot - k\right) \right| \\ &= \left| \sum_{k \in x/h + (-N,0)} \lambda_k \left(g(h\cdot)\right) \, \varphi\left(h^{-1}x \cdot - k\right) \right| \end{aligned}$$

Für beliebiges $1 \le p < \infty$ und $q = (p-1)/p$, also $1/p + 1/q = 1$, ist dann, mit unserem "Trick" aus dem Beweis von Proposition 6.5

$$|f(x) - Q_h f(x)|^p \le N^{p-1} \sum_{k \in x/h + (-N,0)} \left| \lambda_k \left(g(h\cdot)\right) \, \varphi\left(h^{-1}x \cdot - k\right) \right|^p \tag{6.37}$$

Als nächstes schauen wir uns mal den Ausdruck $\lambda_k \left(g(h\cdot)\right)$, $k \in \mathbb{Z}$, an. Dazu bemerken wir zuerst einmal, daß für $y \in \mathbb{R}$ und $\ell = 0, \ldots, n$ die Gleichung[172]

$$\left(T_n f\right)^{(\ell)}(y) = \sum_{j=0}^{n} \frac{f^{(j)}(x)}{j!} \underbrace{\frac{d^\ell}{dy^\ell} (y-x)^j}_{j!/(j-\ell)!(y-x)^{j-\ell}} = \sum_{j=0}^{n} \frac{f^{(j)}(x)}{(j-\ell)!} (y-x)^{j-\ell} = \sum_{j=0}^{n-\ell} \frac{f^{(j+\ell)}}{j!} (y-x)^j,$$

also

$$\left(T_n f\right)^{(\ell)} = T_{n-\ell} f^{(\ell)}, \qquad \ell = 0, \ldots, n, \tag{6.38}$$

gilt. Da $g = f - T_n f$ ist, ergibt sich also nach (6.33)

$$\begin{aligned} \lambda_k \left(g(h\cdot)\right) &= \sum_{j=0}^{n} \frac{p_j(0)}{j!} \, h^j g^{(j)}(hk) = \sum_{j=0}^{n} \frac{h^j \, p_j(0)}{j!} \left(f - T_n f\right)^{(j)}(hk) \\ &= \sum_{j=0}^{n} \frac{h^j \, p_j(0)}{j!} \left(f^{(j)} - T_{n-j} f^{(j)}\right)(hk). \end{aligned}$$

[172] Unter Verwendung der Konvention, daß $j! = \infty$ für $j < 0$, also insbesondere $1/j! = 0$.

Nach der *Taylor–Formel mit Integralrestglied*,

$$(f - T_n f)(y) = \frac{1}{n!} \int_x^y f^{(n+1)}(t)\,(y-t)^n\,dt, \qquad f \in C^{n+1}(\mathbb{R}), \quad y \in \mathbb{R}, \tag{6.39}$$

siehe z.B. [29, S. 285][173] oder Übung 6.9, ist somit

$$\lambda_k\left(g(h\cdot)\right) = \sum_{j=0}^{n} \frac{h^j\,p_j(0)}{j!(n-j)!} \int_x^{hk} f^{(n+1)}(t)\,(hk-t)^{n-j}\,dt. \tag{6.40}$$

Jetzt können wir unsere Bausteinchen zusammensetzen! Da die Summe nur über solche Kombinationen x, k läuft, für die $x/h - k \in [0, N]$, also $x - hk \in h[0, N]$, also $|x - hk| \le Nh$ gilt, können wir nun (6.37) nach x integrieren, (6.40) einsetzen, um so

$$\begin{aligned}
\|f - Q_h f\|_p^p &= \int_{\mathbb{R}} |f(x) - Q_h f(x)|^p\,dx \\
&\le N^{p-1} \int_{\mathbb{R}} \sum_{k \in x/h + (-N,0)} \left|\lambda_k\left(g(h\cdot)\right)\,\varphi\left(h^{-1}x - k\right)\right|^p\,dx \\
&\le ((n+1)N)^{p-1} \int_{\mathbb{R}} \sum_{k \in x/h+(-N,0)} \sum_{j=0}^{n} \underbrace{\left|\frac{p_j(0)}{j!(n-j)!}\right|^p}_{\le C^p} \times \\
&\qquad \times \left| h^j \int_0^{x-hk} f^{(n+1)}(x+t)\,(hk-x-t)^{n-j}\,dt\,\varphi\left(h^{-1}x-k\right)\right|^p dx \\
&\le ((n+1)N)^{p-1} C^p \int_{\mathbb{R}} \sum_{k \in x/h+[-N,0]} \sum_{j=0}^{n} h^{jp} \underbrace{|x-hk|^{p-1}}_{\le (Nh)^{p-1}} \int_0^{x-kh} \underbrace{\left|(hk-x-t)^{n-j}\right|^p}_{\le (2Nh)^{(n-j)p} \le (2N)^{np} h^{(n-j)p}} \times \\
&\qquad \times \left|f^{(n+1)}(x+t)\right|^p \underbrace{\left|\varphi\left(h^{-1}x-k\right)\right|^p}_{\le \|\varphi\|_\infty^p}\,dt\,dx \\
&\le \underbrace{(n+1)^p N^p C^p N^{p-1} N^{np}\,2^{np}\,\|\varphi\|_\infty^p}_{=:C_1^p}\,h^{np}\,h^{p-1} \int_{\mathbb{R}} \int_0^{hN} \left|f^{(n+1)}(x+t)\right|^p\,dt\,dx \\
&= C_1^p h^{np}\,h^{p-1} \int_0^{hN} \int_{\mathbb{R}} \left|f^{(n+1)}(x)\right|^p dx\,dt = N C_1^p\,h^{(n+1)p} \left\|f^{(n+1)}\right\|_p^p
\end{aligned}$$

zu erhalten, was (6.31) mit der Konstanten

$$C := 2^n N^{n+2}(n+1)\,\|\varphi\|_\infty \tag{6.41}$$

liefert. □

Übung 6.9 Beweisen Sie die Taylor–Formel (6.39).
Hinweis: Partielle Integration. ♢

[173] Wer hofft, dort die Lösung von Übung 6.9 zu finden, muß leider enttäuscht werden, denn es ist dort auch nur als Übungsaufgabe aufgelistet.

Bemerkung 6.31 *Was hat das nun für einen Sinn, solche Konstanten wie in (6.41) so explizit anzugeben? Nun, man sieht ihnen an, wie klein man h wählen muß, um überhaut mal langsam sowas wie ein vernüfntiges Resultat zu bekommen; eine Minimalforderung ist demnach auf jeden Fall $h < (2N)^{-1}$, was ja auch geometrisch sinnvoll ist: Die Funktion sollte zumindest so gestaucht werden, daß man verschiedene φs vernünftig miteinander vergleichen kann.*
Im Beweis ging auch ein, daß $\varphi \in C_{00}(\mathbb{R})$, denn ansonsten könnte ja die Norm $\|\varphi\|_\infty$ unendlich werden und die Aussage von Satz 6.29 bedeutungslos machen durch $C = \infty$. Es geht auch mit schwächeren Bedingungen, aber dann wird der Beweis auch technisch aufwendiger. Für Details siehe [34] oder [18, Kap. ?].

Weia!
Waga!
Woge, du Welle,
walle zur Wiege!
Wagalaweia!
Wallala weiala weia!

R. Wagner, *Das Rheingold*

7 Wavelets

Jetzt also zur "modernen" Approximationstheorie, nämlich zu Wavelets und deren Approximationsfähigkeit. Wir wir gleich sehen werden, sind Wavelets ein Spezialfall von translationsinvarianten Räumen. Auch wenn es nicht unbedingt notwendig ist, werden wir Wavelets hier "nur" in $L_2(\mathbb{R})$ betrachten; dies beruht vor allem auf der Tatsache, daß wir es hier mit einem *Hilbertraum* zu tun haben, dessen Norm auf dem inneren Produkt

$$\langle f, g\rangle := \int_{\mathbb{R}} f(t)\,\overline{g(t)}\,dt, \qquad f, g \in L_2(\mathbb{R}),$$

beruht; die komplexe Konjugation bringen wie mal vorsichtshalber ins Spiel, um auch Fouriertransformierte gegeneinander integrieren zu können.

7.1 Multiresolution Analysis

Die *diskreten Wavelets*[174] führt man am besten über den Begriff der *Multiresolution Analysis* ein, die auf Mallat zurückgeht, siehe z.B. [54].

Definition 7.1 *Eine Folge $V_j \subset L_2(\mathbb{R})$, $j \in \mathbb{Z}$, von linearen Räumen heißt* Multiresolution Analysis[175]*, abgekürzt* MRA*, wenn*

1. *die Räume V_j* verschachtelt[176] *sind, das heißt, $\cdots V_{-1} \subset V_0 \subset V_1 \subset \cdots \subset L_2(\mathbb{R})$ und wenn*

$$\lim_{j\to-\infty} V_j = \{0\} \qquad \textit{sowie} \qquad \overline{\lim_{j\to\infty} V_j} = L_2(\mathbb{R}). \tag{7.1}$$

gilt.

[174] Im Gegensatz zur *Wavelettransformation*, die eher in die harmonische Analysis (Verallgemeinerungen der Fourieranalysis) gehört.

[175] Man hat auch, in "deutschen" Kreisen, schon das grausige Unwort "Multiresolutions–Analyse" gehört, aber da "Mehrauflösungsanalyse" auch nicht gut klingt, bleiben wir doch lieber beim englischen Terminus Technicus.

[176] Englisch *"nested"*.

2. *die Räume* translationsinvariant *sind, das heißt, wenn für* $j \in \mathbb{Z}$

$$f \in V_j \qquad \Longleftrightarrow \qquad f(\cdot + k) \in V_j, \quad k \in \mathbb{Z}. \tag{7.2}$$

3. *die Räume* Skalenräume *sind, das heißt, wenn für* $j \in \mathbb{Z}$

$$f \in V_j \qquad \Longleftrightarrow \qquad f(2\,\cdot) \in V_{j+1} \tag{7.3}$$

4. V_0 *von einer* Skalierungsfunktion φ *erzeugt wird, also*

$$V_0 = \mathbb{S}_2(\varphi) = \{\varphi * c \,:\, c \in \ell_2(\mathbb{Z})\}, \tag{7.4}$$

wobei die Translate von φ *sogar eine* Riesz[177]–Basis *von* V_0 *bilden, das heißt, es gibt Konstanten* $A, B > 0$*, so daß*

$$A\,\|c\|_{\ell_2(\mathbb{Z})} \le \|\varphi * c\|_{L_2(\mathbb{R})} \le B\,\|c\|_{\ell_2(\mathbb{Z})}, \qquad c \in \ell_2(\mathbb{Z}). \tag{7.5}$$

Bemerkung 7.2 *1. Man kann eine MRA auch nur für* V_j*,* $j \in \mathbb{N}_0$*, definieren; einen wirklichen Unterschied machen die immer niedriger auflösenden Räume eigentlich nicht, viel "wichtiger" sind die* V_j *mit* $j \ge 0$*.*

2. *Die Bedingung (7.5) bezeichnet man auch als die* Stabilität *von* φ*, genauer der Translate von* φ*. Besonders einfache stabile Funktionen sind die, die* orthonormale *Translate haben, denn dann ist*

$$\begin{aligned}\|\varphi * c\|_2^2 &= \langle \varphi * c, \varphi * c\rangle = \int_{\mathbb{R}} \left(\sum_{j\in\mathbb{Z}} \varphi(t-j)\,c_j\right)\left(\sum_{k\in\mathbb{Z}} \varphi(t-k)\,c_k\right) dt \\ &= \sum_{j,k\in\mathbb{Z}} c_j\,c_k \underbrace{\int_{\mathbb{R}} \varphi(t-j)\;\varphi(t-k)\;dt}_{=\delta_{jk}} = \sum_{k\in\mathbb{Z}} c_k^2 = \|c\|_2^2,\end{aligned}$$

wir haben also sogar $A = B = 1$*.*

3. *Die Forderung nach Stabilität befreit uns auch von dem Dilemma translationsinvarianter Räume, das wir in Proposition 6.5 durch die Einschränkung auf Funktionen mit kompaktem Träger zu lösen versuchten. Hier folgt trivialerweise aus der Definition*

> Bilden die Translate von $\varphi \in L_2(\mathbb{R})$ eine Riesz–Basis, dann ist $\mathbb{S}_2(\varphi) \subset L_2(\mathbb{R})$.

[177] Figyes (Frederic) Riesz, 1880–1956, und Marcel Riesz, 1886–1969, ungarisches Brüderpaar von Mathematikern, die genau eine gemeinsame Arbeit verfasst haben, und zwar während des ersten Weltkriegs über das Randverhalten analytischer Funktionen. Marcel Riesz gilt als einer der Väter der Funktionalanalysis und gründete 1922 zusammen mit dem uns auch bereits wohlbekannten Alfred Haar das "János Bolyai" Mathematik–Institut in Szeged (Ungarn).

4. *Noch ein Wort zum Namen "Multiresolution": Die Idee bei der Definition der Räume V_j besteht darin, daß durch die immer feinere Skalierung der Funktionen in V_j für immer größeres j, man Funktionen mit immer feineren Details darstellen kann und daß am jede Funktion $f \in L_2(\mathbb{R})$ durch eine Folge $f_j \in V_j$ (mit immer mehr Details) beliebig gut approximiere kann. Anders gesagt: Mit diesen Funktionen f_j betrachtet man also wegen des (möglicherweise) höheren Detailreichtums immer höher auflösende Näherungen von f.*

Beispiel 7.3 *Die einfachste Multiresolution Analysis, die gleichzeitig auch den "Modellfall" darstellt, wird von der Skalierungsfunktion $\varphi = \chi$ erzeugt. Die Räume $V_j := \sigma_{2^j}\mathbb{S}(\chi)$, $j \in \mathbb{Z}$, sind dann nichts anderes als die* stückweise *konstanten Funktionen, genauer, die Treppenfunktionen, die auf den* dyadischen Intervallen $[2^{-j}k, 2^{-j}(k+1)]$ *konstant sind, siehe Abb. 7.1. Wie sieht es nun mit den Eigenschaften aus? Nun, die Bedingungen (7.2), (7.3) und (7.4) folgen direkt aus der Definition der V_j, die Stabilität (7.5) ergibt sich aus*

$$\|\chi * c\|_2^2 = \int_{\mathbb{R}} \sum_{j\in\mathbb{Z}} |\chi(t-j)\,c(j)|^2 \, dt = \sum_{k\in\mathbb{Z}} \int_0^1 |\underbrace{\chi(t-j+k)}_{=\delta_{jk}\,\chi(t)}\, c(j)|^2 \, dt = \sum_{k\in\mathbb{Z}} |c(k)|^2 = \|c\|_2^2$$

sogar mit[178] $A = B = 1$. *Schließlich ist auch die "Verschachtelung" $V_j \subset V_{j+1}$ klar und daß $\overline{V_j} \to L_2(\mathbb{R})$ für $j \to \infty$ ist die Dichtheit der Treppenfunktionen, wohingegen $V_j \to \{0\}$ für $j \to -\infty$ auf der einfachen Tatsache beruht, daß die einzige konstante Funktion in $L_2(\mathbb{R})$ die Nullfunktion ist.*

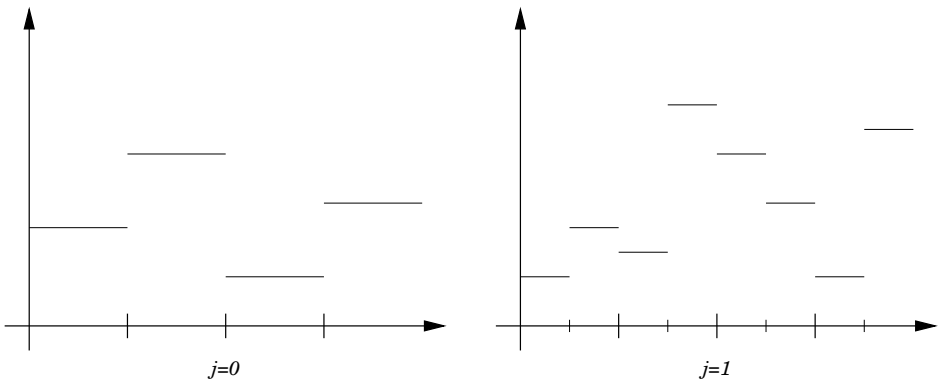


Abbildung 7.1: "Typische" Funktionen aus V_0 und V_1 in Multiresolution Analysis, die von χ erzeugt wird. Je höher der Index j wird, desto mehr wächst natürlich die Fähigkeit der Funktionen, feinere Details darzustellen oder wiederzugeben.

Daß eine Funktion φ eine MRA erzeugt, stellt auch besondere Forderungen an diese Funktion, und zwar, daß sie Lösung einer *Funktionalgleichung* ist.

[178] Na gut, die Translate dieser Funktion sind wegen ihres disjunkten Trägers ja auch noch orthonormal, also würde auch Bemerkung 7.2 reichen, aber es ist doch immer gut, eine zweite Meinung zu hören.

Proposition 7.4 *Erzeugt eine Funktion $\varphi \in L_2(\mathbb{R})$ eine Multiresolution Analysis, dann gibt es eine Folge $a \in \ell_2(\mathbb{Z})$, so daß φ die* Verfeinerungsgleichung[179] *oder* Zweiskalenbeziehung

$$\varphi = \sigma_2 \, (\varphi * a) = \sum_{k \in \mathbb{Z}} a(k) \, \varphi \, (2 \cdot -k) \tag{7.6}$$

erfüllt. Alternativ kann man, die Verfeinerungsgleichung (7.6) auch als

$$\widehat{\varphi}(\xi) = \frac{1}{2} \, \widehat{a} \, (\xi/2) \; \widehat{\varphi} \, (\xi/2) \, , \qquad \xi \in \mathbb{R}, \tag{7.7}$$

schreiben.

Beweis: Wir zeigen zuerst, daß V_1 von den Funktionen $\varphi \, (2 \cdot -k)$, $k \in \mathbb{Z}$, erzeugt wird. Dazu brauchen wir bloß zu bemerken, daß für jedes $f \in V_1$ die Funktion $f(\cdot/2)$ zu $V_0 = \mathbb{S}(\varphi)$ gehört, daß also für ein passendes $a \in \ell(\mathbb{Z})$

$$f(x/2) = \sum_{k \in \mathbb{Z}} a(k) \, \varphi(x - k), \qquad x \in \mathbb{R},$$

und ersetzt man x durch $2x$, dann erhält man, daß $f \in \sigma_2 \mathbb{S}(\varphi)$ liegt. Da insbesondere $\varphi \in V_0 \subset V_1 = \sigma_2 \mathbb{S}(\varphi)$ gibt es also ein $a \in \ell(\mathbb{Z})$, so daß $\varphi = \sigma_2 \, (\varphi * a)$, was nichts anderes als (7.6) ist. Die anderen beiden Eigenschaften folgen mit Satz 6.13.
Daß a zu $\ell_2(\mathbb{Z})$ gehört, folgt schließlich aus der Tatsache, daß

$$\|\varphi\|_2 = \|\sigma_2 \, (\varphi * a)\|_2 = \frac{1}{\sqrt{2}} \, \|(\varphi * a)\|_2 \geq \frac{A}{\sqrt{2}} \, \|a\|_2 \, .$$

□

Übung 7.1 Zeigen Sie, daß $a^* \, (e^{-i \cdot}) = \widehat{a}$ ist und leiten Sie (7.7) aus (7.6) her. ♢

Definition 7.5 *Eine Funktion φ, die eine Verfeinerungsgleichung der Form (7.6) erfüllt, heißt* verfeinerbar.

Übung 7.2 Zeigen Sie: Sind φ, ψ verfeinerbare Funktionen, dann ist $\varphi * \psi$ ebenfalls verfeinerbar. ♢

Beispiel 7.6 *Einfache Beispiele für verfeinerbare Funktionen sind*

1. *die charakteristische Funktion χ von $[0, 1]$. Hier ist ja offensichtlich*

$$\chi = \underbrace{\chi \, (2 \cdot)}_{\sim \left[0, \frac{1}{2}\right]} + \underbrace{\chi \, (2 \cdot -1)}_{\sim \left[\frac{1}{2}, 1\right]} .$$

[179] Englisch: *"refinement equation"*.

2. *die "Hutfunktion"*

$$\varphi(x) = \begin{cases} x+1, & x \in [-1,0], \\ 1-x, & x \in [0,1], \\ 0 & x \in \mathbb{R} \setminus [-1,1], \end{cases}$$

die die Verfeinerungsgleichung

$$\varphi = \frac{1}{2}\underbrace{\varphi\left(2\cdot+1\right)}_{\sim[-1,0]} + \underbrace{\varphi(2\cdot)}_{\sim\left[-\frac{1}{2},\frac{1}{2}\right]} + \frac{1}{2}\underbrace{\varphi\left(2\cdot-1\right)}_{\sim[0,1]}$$

erfüllt, siehe auch Abb 7.2. Als Autokonvolution[180] *der charakteristischen Funktion bleibt der Hutfunktion ja auch gar nichts anderes übrig, als selbst verfeinerbar zu sein.*

Daß alle *B–Splines verfeinerbar sind, ist nun nicht weiter verwunderlich, wenn man die Verfeinerbarkeit der charakteristischen Funktion und Übung7.2 in Betracht zieht.*

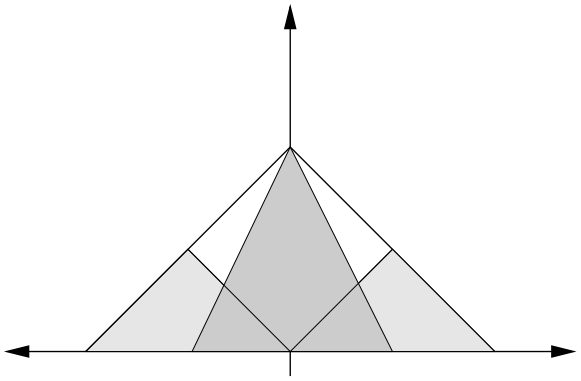


Abbildung 7.2: Verfeinerung der Hutfunktion als "Überlagerung" Ihrer gestauchten und verschobenen Kopien.

Übung 7.3 Zeigen Sie: Der B–Spline N_j ist verfeinerbar mit dem Vektor $a \in \ell(\mathbb{Z})$, der die von Null verschiedenen Einträge

$$a(k) = 2^{-j}\binom{j+1}{k}, \qquad k = 0, \ldots, j+1,$$

hat. ♢

Das führt zu einer netten Beobachtung, wie man (die Fouriertransformation) verfeinerbarer Funktionen berechnen kann: Iteriert man nämlich die Formel (7.7), dann liefert das für $\xi \in \mathbb{R}$

$$\widehat{\varphi}(\xi) = \frac{\widehat{a}(\xi/2)}{2}\,\widehat{\varphi}(\xi/2) = \frac{\widehat{a}(\xi/2)}{2}\,\frac{\widehat{a}(\xi/4)}{2}\,\widehat{\varphi}(\xi/4) = \cdots = \widehat{\varphi}\left(2^{-N}\xi\right)\prod_{j=1}^{N}\frac{\widehat{a}\left(2^{-j}\xi\right)}{2}, \qquad N \in \mathbb{N}.$$

[180] Das "vornehme" Wort für "Faltung mit sich selbst".

Ist nun darüberhinaus[181] $\varphi \in L_1(\mathbb{R})$, ist also φ eine stetige Funktion, dann erhalten wir, daß

$$\widehat{\varphi}(\xi) = \widehat{\varphi}(0) \prod_{j=1}^{\infty} \frac{\widehat{a}\,(2^{-j}\xi)}{2} \tag{7.8}$$

Daraus können wir sofort ein paar Konsequenzen ableiten.

Proposition 7.7 *Sei $0 \neq \varphi \in L_1(\mathbb{R})$ verfeinerbar bezüglich $a \in \ell_1(\mathbb{Z})$, das heißt $\varphi = \sigma_2\,(\varphi * a)$, und es konvergiere das unendliche Produkt in (7.8). Dann*

1. *ist $\widehat{\varphi}(0) \neq 0$.*
2. *ist*
$$2 = \widehat{a}(0) = \sum_{k\in\mathbb{Z}} a(k).$$

Beweis: Wäre $\widehat{\varphi}(0) = 0$, dann liefert (7.8), daß $\widehat{\varphi} = 0$, also $\varphi = 0$; setzt man nun $\xi = 0$ in (7.7), dann erhält man, daß

$$\widehat{\varphi}(0) = \frac{1}{2}\,\widehat{a}(0)\,\widehat{\varphi}(0) \qquad \Longrightarrow \qquad \frac{1}{2}\,\widehat{a}(0) = 1,$$

da $\widehat{\varphi}(0) \neq 0$. □

Wäre übrigens $|\widehat{a}(0)| < 1$, dann konvergiert das unendliche Produkt gegen 0 und es wäre wieder $\varphi = 0$, wäre $|\widehat{a}(0)| > 1$, dann divergiert das Produkt und $\widehat{\varphi}$ wäre überall unendlich.

7.2 Orthogonale Skalierungsfunktionen

In den folgenden Kapiteln werden wir uns nur mit *orthonormalen* Skalierungsfunktionen auseinandersetzen, da für diese die Definition und Berechnung der Wavelets wesentlich einfacher ist. Dazu sehen wir uns erst einmal an, ob und inwiefern Orthonormalität[182] eine Einschränkung bedeutet.
Man kann sowohl Orthogonalität als auch Stabilität von φ relativ schön über die Fouriertransformierte von φ beschreiben. Dazu verwendet man die Plancherel–Identität (6.18) und erhält, daß φ genau dann orthonormale Translate hat, wenn für $j, k \in \mathbb{Z}$

$$\begin{aligned}
\delta_{jk} &= \langle \varphi(\cdot - j), \varphi(\cdot - k)\rangle = \int_{\mathbb{R}} \varphi\,(t-j)\;\varphi\,(t-k)\;dt \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \underbrace{(\tau_{-j}\,\varphi)^{\wedge}\,(\vartheta)}_{\widehat{\varphi}(\vartheta)e^{-ij\vartheta}}\;\underbrace{\overline{(\tau_{-k}\,\varphi)}^{\wedge}(\vartheta)}_{\overline{\widehat{\varphi}(\vartheta)}e^{ik\vartheta}}\;d\vartheta = \frac{1}{2\pi}\int_{\mathbb{R}} |\widehat{\varphi}(\vartheta)|^2\;e^{-i(j-k)\vartheta}d\vartheta
\end{aligned}$$

[181] Zum Beispiel, wenn φ kompakten Träger hat.
[182] Ob wir uns nun mit Ortho*gonal*ität oder Ortho*normal*ität auseinandersetzen, das ist nur eine Frage der Normierung und daher praktisch irrelevant.

$$
\begin{aligned}
&= \frac{1}{2\pi} \sum_{\ell\in\mathbb{Z}} \int_{2\ell\pi}^{2(\ell+1)\pi} |\widehat{\varphi}(\vartheta)|^2 \; e^{-i(j-k)\vartheta} d\vartheta \\
&= \frac{1}{2\pi} \sum_{\ell\in\mathbb{Z}} \int_{0}^{2\pi} |\widehat{\varphi}(\vartheta+2\ell\pi)|^2 \; e^{-i(j-k)\vartheta} \underbrace{e^{-2i(j-k)\ell\pi}}_{=1} d\vartheta \\
&= \frac{1}{2\pi} \int_0^{2\pi} \underbrace{\sum_{\ell\in\mathbb{Z}} |\widehat{\varphi}(\vartheta+2\ell\pi)|^2}_{=:g(\vartheta)} e^{-i(j-k)\vartheta} \, d\vartheta,
\end{aligned}
$$

was der $(j-k)$–te *Fourierkoeffizient* der Funktion g ist. Bezeichnet also wieder g_j den j–ten Fourierkoeffizient von g, so ist, mit $k=0$, Orthonormalität also dazu äquivalent, daß $g_j = \delta_{j0}$, also $g \equiv 1$. Mit anderen Worten:

$$
\langle \varphi(\cdot - j), \varphi(\cdot - k) \rangle = \delta_{jk}, \quad j,k \in \mathbb{Z} \qquad \Longleftrightarrow \qquad \sum_{\ell\in\mathbb{Z}} |\widehat{\varphi}(\cdot + 2\ell\pi)|^2 \equiv 1. \tag{7.9}
$$

Stabilität geht ähnlich; hier sehen wir zuerst mal, daß nach (6.18) und (6.12) für beliebiges $c \in \ell(\mathbb{Z})$

$$
\begin{aligned}
\|\varphi * c\|_2^2 &= \frac{1}{2\pi} \left\| (\varphi * c)^\wedge \right\|_2^2 = \frac{1}{2\pi} \left\| \widehat{c}\,\widehat{\varphi} \right\|_2^2 = \frac{1}{2\pi} \int_{\mathbb{R}} \left| \sum_{k\in\mathbb{Z}} c(k)\, e^{-ik\vartheta} \right|^2 |\widehat{\varphi}(\vartheta)|^2 \; d\vartheta \\
&= \frac{1}{2\pi} \int_0^{2\pi} \left| \sum_{k\in\mathbb{Z}} c(k)\, e^{-ik\vartheta} \right|^2 \underbrace{\sum_{\ell\in\mathbb{Z}} |\widehat{\varphi}(\vartheta+2\ell\pi)|^2}_{=:g(\vartheta)} \; d\vartheta = \frac{1}{2\pi} \int_0^{2\pi} |\widehat{c}(\vartheta)|^2 \; g\,(\vartheta) \; d\vartheta
\end{aligned}
$$

Mittels dieser Identität und der "diskreten Parseval–Identität" (7.12), siehe Übung 7.4, läßt sich dann die Stabilitätsbedingung (7.5) äquivalent in

$$
A^2 \; \|\widehat{c}\|_2^2 \le \int_0^{2\pi} |\widehat{c}(\vartheta)|^2 \; g\,(\vartheta) \; d\vartheta \le B^2 \; \|\widehat{c}\|_2^2, \qquad c \in \ell_2(\mathbb{Z}), \tag{7.10}
$$

umformen, woraus

$$
A^2 \le \sum_{\ell\in\mathbb{Z}} |\widehat{\varphi}(\cdot + 2\ell\pi)|^2 \le B^2 \tag{7.11}
$$

folgt. Denn hätte für ein $\varepsilon > 0$ die Menge

$$
I_{\cdot} = \left\{ \xi \in [0, 2\pi] \; : \; \sum_{\ell\in\mathbb{Z}} |\widehat{\varphi}\,(\xi + 2\ell\pi)|^2 \le A - \epsilon \right\} \subset [0, 2\pi],
$$

positives Maß μ, dann gibt es eine Folge $c_n \in \ell_{00}(\mathbb{Z})$, so daß die *trigonometrischen Polynome*[183] $\widehat{c_n}$, $n \in \mathbb{N}$, in $L_2(\mathbb{T})$ gegen χ_{I_-} konvergieren. Dann ist

$$
\lim_{n\to\infty} \|\widehat{c_n}\|_2 = \left\| \chi_{I_-} \right\|_2, \qquad \text{aber} \qquad \lim_{n\to\infty} \|\widehat{c_n}\,\widehat{\varphi}\|_2 \le A - \varepsilon,
$$

[183] Siehe die komplexe Definition in (1.4)!

was für hinreichend großes n einen Widerspruch zu (7.10) darstellt. Die obere Grenze beweist man entsprechend.

Übung 7.4 Beweisen Sie die *diskrete Parseval–Identität*:

$$\|c\|_2 = \frac{1}{\sqrt{2\pi}} \, \|\widehat{c}\|_{2,\mathbb{T}} \, . \tag{7.12}$$

♢

Aus (7.11) können wir nun mit dem in [16] so genannten[184] "Orthogonalisierungstrick" ersehen, daß das mit der Orthogonalität "eigentlich" gar nicht so wild ist, sondern daß man eigentlich nur die "falsche" Funktion zur Erzeugung des Translationsinvarianten Raums gewählt hat.

Satz 7.8 *Ist $\varphi \in L_2(\mathbb{R})$ eine stabile Funktion[185], dann gibt es eine Funktion $\phi \in \mathbb{S}_2(\varphi)$, die orthonormale Translate hat.*

Beweis: Wir definieren eine Funktion ψ über ihre Fouriertransformierte

$$\widehat{\psi} = \left(\sum_{\ell \in \mathbb{Z}} |\widehat{\varphi}(\cdot + 2\ell\pi)|^2 \right)^{-1/2}, \qquad \frac{1}{B^2} \le \widehat{\psi}^2 \le \frac{1}{A^2},$$

und setzen

$$\phi = \varphi * \psi, \qquad \text{also} \qquad \widehat{\phi} = \widehat{\varphi}\,\widehat{\psi}$$

Dann ist

$$\|\phi\|_2 = \frac{1}{\sqrt{2\pi}} \left\|\widehat{\phi}\right\|_2 = \frac{1}{\sqrt{2\pi}} \left\|\widehat{\varphi}\,\widehat{\psi}\right\|_2 \le \frac{1}{\sqrt{2\pi}} \, \|\widehat{\varphi}\|_2 \left\|\widehat{\psi}\right\|_\infty \le \frac{1}{\sqrt{2\pi}\,A} \, \|\widehat{\varphi}\|_2 = \frac{1}{A} \, \|\varphi\|_2$$

also $\phi \in L_2(\mathbb{R})$ und

$$\sum_{\ell \in \mathbb{Z}} \left|\widehat{\phi}\,(\cdot + 2\ell\pi)\right|^2 \;=\; \sum_{\ell \in \mathbb{Z}} \left| \widehat{\varphi}\,(\cdot + 2\ell\pi) \, \underbrace{\widehat{\psi}\,(\cdot + 2\ell\pi)}_{=\widehat{\psi}(\cdot)} \right|^2 = \left|\widehat{\psi}\right|^2 \underbrace{\sum_{\ell \in \mathbb{Z}} \left|\widehat{\phi}\,(\cdot + 2\ell\pi)\right|^2}_{=\widehat{\psi}^{-2}} = 1,$$

also hat ϕ nach (7.9) orthogonale Translate.
Setzen wir nun

$$c(k) := \frac{1}{2\pi} \int_{\mathbb{T}} \widehat{\psi}(\vartheta) \, e^{-ik\theta} d\vartheta,$$

der Vektor der Fourierkoeffizienten von $\widehat{\psi}$, dann ist

$$\|c\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} \sum_{\ell \in \mathbb{Z}} |\widehat{\varphi}\,(\vartheta + 2\ell\pi)|^2 \, d\vartheta = \frac{1}{2\pi} \int_{\mathbb{R}} |\widehat{\varphi}\,(\vartheta)|^2 \, d\vartheta = \|\varphi\|_2^2 \, ,$$

[184] Dies ist *keine* Rechtschreibreform sondern korrekt! O tempora, o orthographia ...

[185] Wir werden gelegentlich "stabile Funktion" anstelle des eigentlich korrekten "Funktion mit stabilen ganzzahligen Translaten" verwenden – die Verwirrung sollte sich aber trotzdem in Grenzen halten.

und da die Partialsummen der Fourierreihe $\widehat{c}(-\cdot)$ in $L_2(\mathbb{T})$ gegen $\widehat{\psi}$ konvergieren[186], ist $\widehat{c}(-\xi) = \widehat{\psi}(\xi)$, $\xi \in \mathbb{R}$, und somit

$$\widehat{\phi} = \widehat{\psi}\,\widehat{\varphi} = \widehat{c}(-\cdot)\,\widehat{\varphi}, \qquad \text{also} \qquad \phi = -\varphi * \sigma_{-1} c,$$

weswegen wirklich $\phi \in \mathbb{S}_2(\varphi)$ liegt. □

Übung 7.5 Zu $f \in L_2(\mathbb{T})$ sei $c \in \ell_2(\mathbb{Z})$, definiert durch

$$c(k) = \frac{1}{2\pi} \int_{\mathbb{T}} f(t)\, e^{-ikt}\, dt, \qquad k \in \mathbb{Z},$$

sowie $c_n = \chi_{[-N,N]}\, c$, $n \in \mathbb{N}$. Zeigen Sie:

1. $\widehat{c}_n$ ist ein trigonometrisches Polynom der Ordnung n, also $\widehat{c}_n$.
2. $\widehat{c}_n(-\cdot)$ ist *Bestapproximation* in $L_2(\mathbb{T})$:

$$\|f - \widehat{c}_n(-\cdot)\|_2 = \min_{p \in T_n} \|f - p\|_2 .$$

♢

So nett das ganze auch aussieht, es hat durchaus praktische Nachteile: Wir kennen nämlich nicht mehr unsere Skalierungsfunktion selbst, sondern nur noch deren Fouriertransformierte! Natürlich kann man daraus die Funktion über die inverse Fouriertransformation erhalten[187], aber sowas wie geschlossene Ausdrücke können wir uns schenken.

Übung 7.6 Bestimmen Sie die Fouriertransformierte der Orthonormalisierung der zentrierten B–Splines M_k aus Übung 6.5. ♢

7.3 Wavelets für orthonormale Skalierungsfunktionen

Jede Multiresolution Analysis besteht ja aus einer verschachtelten Folge $V_j \subset V_{j+1}$, $j \in \mathbb{Z}$, von linearen Räumen. Dabei braucht man (wegen der "faktischen" halbzahligen Verschiebung) dann, um eine Funktion aus V_{j+1} darzustellen, ziemlich genau doppelt so viel Information wie man für eine Funktion aus V_j benötigen würde. Das ist ja noch angebracht, wenn man es mit $f \in V_{j+1} \setminus V_j$ zu tun hat, aber für ein $f \in V_j$ ist die Darstellung in V_j schlichtweg zu kompliziert.

Beispiel 7.9 *Betrachten wir nur einmal die von χ erzeugte Multiresolution Analysis; dann hat $\chi \in V_0 \subset V_j$, $j \in \mathbb{N}$, in so einem V_j die Darstellung*

$$\chi = \sum_{k=0}^{2^j-1} 1\, \chi\left(2^j \cdot - k\right) =: \chi * c\left(2^j \cdot\right),$$

für die die 2^j Koeffizienten $c(k) = 1$, $k = 0, \ldots, 2^j - 1$, zu speichern sind.

[186] Sie sind sogar *Bestapproximationen* der trigonometrischen Polynome, siehe Übung 7.5.

[187] In L_2 ist das ja alles verhältnismäßig harmlos ...

Schon diese Komplexitätsüberlegungen sind ein guter Grund, sich "irgendwas" einfallen zu lassen, um die komplizierte Darstellung nur für Funktionen zu verwenden, für die man sie auch wirklich braucht. Und hier heißt das Stichwort "Projektionen": Ist nämlich $P_j : L_2(\mathbb{R}) \to V_j$ eine beliebige Projektion[188], dann berechnen wir für ein $f \in V_{j+1}$ zuerst den "V_j–Anteil" als $P_j f$ und stellen dann f als

$$f = \underbrace{P_j f}_{\in V_j} + \underbrace{(f - P_j f)}_{\in (V_{j+1} \setminus V_j) \cup \{0\}}$$

dar. Nun besitzt $L_2(\mathbb{R})$ als Hilbertraum aber eine sehr natürliche Projektion, nämlich die orthogonale Projektion, definiert durch $\langle f - P_j f, V_j \rangle = 0$, die noch dazu den Vorteil hat, eine *Bestapproximation* aus V_j zu sein, siehe (2.17) in Lemma 2.22.

Definition 7.10 *1. Für $j \in \mathbb{Z}$ definieren wir den* Waveletraum $W_j = V_{j+1} \ominus V_j$ *als das* orthogonale Komplement

$$W_j := \{f \in V_{j+1} \ : \ \langle f, V_j \rangle = 0\}\,. \tag{7.13}$$

2. Die Funktion $\psi \in V_1$ heißt Wavelet *zur Skalierungsfunktion φ, wenn ψ orthogonal und $W_0 = \mathbb{S}_2(\psi)$ ist.*

Bemerkung 7.11 *1. Definieren kann man bekanntlich viel! Daher müssen wir natürlich erst mal beweisen, daß es zu jeder Skalierungsfunktion φ tatsächlich auch ein Wavelet gibt – das wird die Hauptaufgabe in diesem Abschnitt sein!*

2. Manche Leute unterscheiden auch zwischen Wavelets und orthogonalen Wavelets*, bei uns gehört Orthogonalität wie in [16] zur Definition des Wavelets.*

3. Ist ψ ein Wavelet zu φ, dann gilt aber wieder

$$W_j := \sigma_{2^j}\, \mathbb{S}_2(\psi), \qquad j \in \mathbb{Z}, \tag{7.14}$$

das heißt, auch die Waveleträume W_j sind Skalenräume.

Übung 7.7 Beweisen Sie (7.14) und zeigen Sie, daß

$$\{\varphi(\cdot - k) \ : \ k \in \mathbb{Z}\} \cup \left\{\psi\left(2^\ell \cdot - k\right) \ : \ \ell = 0, \ldots, j-1,\, k \in \mathbb{Z}\right\}$$

eine Riesz–Basis von V_j ist. ♢

Die Schreibweise "$W_j = V_{j+1} \ominus V_j$" ist eine "saloppe" Umformung von $V_{j+1} = V_j \oplus W_j$, was sich für beliebige $j \in \mathbb{Z}$ und $k \in \mathbb{N}$ als

$$V_{j+k} = V_{j+k-1} \oplus W_{j+k-1} = V_{j+k-2} \oplus W_{j+k-2} \oplus W_{j+k-1} = \cdots = V_j \oplus \bigoplus_{\ell=0}^{k-1} W_{j+\ell} \tag{7.15}$$

[188] Zur Erinnerung: Eine Projektion P ist eine Abbildung mit $P^2 = P$.

schreiben läßt. Bezüglich der Skalierungsfunktionen und Wavelets heißt dies, daß jedes $f \in V_{j+k}$ eine Darstellung

$$f = (\varphi * c_j)\left(2^j \cdot\right) + \sum_{\ell=0}^{k-1} (\psi * d_{j+\ell})\left(2^{j+\ell}\cdot\right), \qquad c_j, d_{j+\ell} \in \ell_2(\mathbb{Z}), \tag{7.16}$$

besitzt. Mit $j = 0$ und der Dichtheitsbedingung (7.1) liefert dann (7.16) für jede Funktion $f \in L_2(\mathbb{R})$ die *Wavelet–Darstellung*

$$f = \varphi * c + \sum_{j=0}^{\infty} (\psi * d_j)\left(2^j \cdot\right), \qquad c, d_j \in \ell_2(\mathbb{Z}). \tag{7.17}$$

Unser Hauptresultat in diesem Abschnitt ist der folgende Satz.

Satz 7.12 *Zu jeder Skalierungsfunktion φ einer Multiresolution Analysis existiert ein zugehöriges Wavelet $\psi \in V_1$.*

Der Beweis hat zwei wesentliche Zutaten: Zuerst einmal können wir, Satz 7.8 sei Dank, annehmen, daß φ eine *orthonormale* Multiresolution Analysis erzeugt[189] denn die Existenz des Wavelets ist ja unabhängig davon, welche Basis von V_0 wir wählen. Außerdem wissen wir aus (7.4), daß es eine Folge $a \in \ell(\mathbb{Z})$ gibt, so daß $\varphi = \sigma_2\,(\varphi * a)$. Für orthonormale Skalierungsfunktionen haben diese Verfeinerungskoeffizienten noch eine interessante Eigenschaft.

Lemma 7.13 *Ist $\varphi \in L_2(\mathbb{R})$ verfeinerbar bezüglich $a \in \ell_2(\mathbb{R})$ und orthonormal, dann ist*

$$\sum_{k\in\mathbb{Z}} a\,(k-2j)\;a(k) = 2\delta_{j0}, \qquad j \in \mathbb{Z}. \tag{7.18}$$

Beweis: Für $j \in \mathbb{Z}$

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} a\,(k-2j)\;a(k) &= \sum_{k,\ell\in\mathbb{Z}} a(\ell)\,a(k)\,\delta_{\ell,k-2j} \\
&= 2\sum_{k,\ell\in\mathbb{Z}} a(\ell)\,a(k)\int_{\mathbb{R}} \varphi\,(2t-\ell)\;\varphi\,(2t-k+2j)\;dt = 2\langle (\varphi * a)\,(2\cdot)\,, (\varphi * k)\,(2\cdot+2j)\rangle \\
&= 2\langle \varphi, \varphi\,(\cdot + j)\rangle = 2\delta_{0j}
\end{aligned}$$

□

Jetzt können wir nämlich den Kandidaten für unser Wavelet sogar explizit angeben.

Definition 7.14 *Zu einer orthonormalen Skalierungsfunktion $\varphi \in L_2(\mathbb{R})$ mit Verfeinerungsfolge $a \in \ell_2(\mathbb{Z})$ definiert man das* Wavelet[190] *$\psi \in L_2(\mathbb{R})$ als*

$$\psi := \sigma_2\left(\varphi * a^{\perp}\right), \qquad a^{\perp} := (-1)^{(\cdot)} a\,(1-\cdot)\,. \tag{7.19}$$

[189] Was wiederum nichts anderes bedeutet, als daß die Translate von φ orthonormal sind.

[190] Wäre die Skalierungsfunktion nicht orthonormal, dann spricht man hier von einem *Prewavelet*. Diese Funktionen sind nicht unbedingt orthogonal zu φ, haben dafür aber, im Gegensatz zu den wirklichen Wavelets immer noch kompakten Träger.

Wir werden nun Satz 7.12 zur Abwechslung mal "fourierfrei" beweisen, erstens wird der Beweis auch nicht viel länger, zweitens ist er elementarer und drittens – warum nicht? Wir beginnen mit zwei einfachen Beobachtungen über ψ.

Lemma 7.15 *Ist φ orthonormal, dann gehört die Funktion ψ zu W_1:*

$$\langle \psi, V_0 \rangle = 0. \tag{7.20}$$

Beweis: Für $j \in \mathbb{Z}$ ergibt sich, unter Verwendung von (7.19) und der Verfeinerungsgleichung (7.6) die Rechnung

$$\begin{aligned}
\langle \psi, \varphi(\cdot - j) \rangle &= \left\langle \sigma_2 \left(\varphi * a^\perp \right), \sigma_2 \left(\tau_j \varphi * a \right) \right\rangle \\
&= \sum_{k,\ell \in \mathbb{Z}} a^\perp(k)\, a(\ell) \underbrace{\int_{\mathbb{R}} \varphi\left(2t-k\right)\, \varphi\left(2t-2j-\ell\right)\, dt}_{= \frac{1}{2}\delta_{k,\ell+2j}} \\
&= \sum_{\ell \in \mathbb{Z}} a^\perp(\ell+2j)\, a(\ell) = \sum_{\ell \in \mathbb{Z}} (-1)^\ell a(1-\ell-2j)\, a(\ell) = \sum_{\ell \in \mathbb{Z}} (-1)^{1-\ell-2j} a(\ell)\, a(1-\ell-2j) \\
&= -\sum_{\ell \in \mathbb{Z}} (-1)^\ell a(1-\ell-2j)\, a(\ell) = -\langle \psi, \varphi(\cdot - j) \rangle,
\end{aligned}$$

also $\langle \psi, \varphi(\cdot - j) \rangle = 0$ und somit (7.20). □

Lemma 7.16 *Ist $\varphi \in L_2(\mathbb{R})$ eine* orthonormale *Skalierungsfunktion, dann hat das Wavelet ψ orthonormale Translate:*

$$\langle \psi(\cdot - j), \psi(\cdot - k) \rangle = \delta_{jk}, \qquad j,k \in \mathbb{Z}. \tag{7.21}$$

Beweis: Für (7.21) genügt es wegen der Translationsinvarianz des Integrals, $k = 0$ anzunehmen und man erhält dann ganz entsprechend mittels Lemma 7.13 für $j \in \mathbb{Z}$, daß

$$\begin{aligned}
\langle \psi, \psi(\cdot - j) \rangle &= \sum_{k,\ell \in \mathbb{Z}} a^\perp(k)\, a^\perp(\ell) \underbrace{\int_{\mathbb{R}} \varphi\left(2t-k\right)\, \varphi\left(2t-2j-\ell\right)\, dt}_{= \frac{1}{2}\,\delta_{k,2j+\ell}} \\
&= \frac{1}{2} \sum_{\ell \in \mathbb{Z}} a^\perp\left(2j+\ell\right)\, a^\perp(\ell) = \frac{1}{2} \sum_{\ell \in \mathbb{Z}} (-1)^\ell\, a\left(1-2j-\ell\right)\, \left(-1\right)^\ell a(1-\ell) \\
&= \frac{1}{2} \sum_{\ell \in \mathbb{Z}} a\left(\ell-2j\right)\, a(\ell) = \delta_{j0},
\end{aligned}$$

was uns (7.21) liefert. □

Wir sind also schon ganz schön weit, denn wir wissen nun, daß $\mathbb{S}_2(\psi)$ einen Teilraum von W_1 aufspannt und daß die Translate von ψ sogar eine Orthonormalbasis von $\mathbb{S}_2(\psi)$ sind. Was noch fehlt, das ist die Inklusion $W_1 \subseteq \mathbb{S}_2(\psi)$, also, daß wir auch wirklich *alle* Funktionen von W_1 durch Linearkombinationen von Translaten von ψ darstellen können. Das erledigt uns das

folgende Lemma 7.17, das zeigt, daß man jede der Funktionen $\varphi(2\cdot -j)$, $j\in\mathbb{Z}$, die ja V_1 aufspannen, mittels $\mathbb{S}_2(\varphi)\oplus\mathbb{S}_2(\psi)$ darstellen kann. Das heißt dann, daß

$$V_1\subseteq\mathbb{S}_2(\varphi)\oplus\mathbb{S}_2(\psi)\subseteq V_1 \qquad\Longrightarrow\qquad \mathbb{S}_2(\psi)=V_1\ominus\underbrace{\mathbb{S}_2(\varphi)}_{=V_0}=V_1\ominus V_0=W_0,$$

was den Beweis von Satz 7.12 komplettiert.
Die explizite Darstellung dieser Koeffizienten in (7.17) mag unnötig kompliziert und geschäftig wirken, aber wird noch ein wichtiges Hilfsmittel für den algorithmischen Umgang mit Wavelets sein.

Lemma 7.17 *Ist φ orthonormal, dann ist für $\epsilon\in\{0,1\}$*

$$\varphi\left(2\cdot-\varepsilon\right)=\frac{1}{2}\left(\varphi*\sigma_{-2}\,\tau_\epsilon a+(-1)^\epsilon\,\psi*\sigma_2\,\tau_{1-\epsilon}a\right). \tag{7.22}$$

Beweis: Wir bezeichnen mit g_ϵ die orthogonale Projektion von $f_\epsilon:=\varphi\left(2\cdot-\varepsilon\right)$ auf $\mathbb{S}_2(\varphi)\oplus\mathbb{S}_2(\psi)$. Da die Translate von φ und ψ nach Lemma 7.16 eine *Orthonormalbasis* dieses Raums bilden, ist, für $\epsilon\in\{0,1\}$,

$$\begin{aligned}
g_\epsilon &= \sum_{k\in\mathbb{Z}}\left\langle f_\epsilon,\varphi\left(\cdot-k\right)\right\rangle\varphi\left(\cdot-k\right)+\sum_{k\in\mathbb{Z}}\left\langle f_\epsilon,\psi\left(\cdot-k\right)\right\rangle\psi\left(\cdot-k\right)\\
&= \sum_{k\in\mathbb{Z}}\left\langle f_\epsilon,\left(\varphi*a\right)\,\left(2\cdot-2k\right)\right\rangle\varphi\left(\cdot-k\right)+\sum_{k\in\mathbb{Z}}\left\langle f_\epsilon,\left(\varphi*a^\perp\right)\,\left(2\cdot-2k\right)\right\rangle\psi\left(\cdot-k\right)\\
&= \sum_{k,\ell\in\mathbb{Z}}a(\ell)\,\varphi(\cdot-k)\underbrace{\int_{\mathbb{R}}\varphi\left(2t-\epsilon\right)\,\varphi\left(2t-2k-\ell\right)\,dt}_{=\frac{1}{2}\,\delta\epsilon,2k+\ell}\\
&\qquad+\sum_{k,\ell\in\mathbb{Z}}a^\perp(\ell)\,\psi(\cdot-k)\underbrace{\int_{\mathbb{R}}\varphi\left(2t-\epsilon\right)\,\varphi\left(2t-2k-\ell\right)\,dt}_{=\frac{1}{2}\,\delta\epsilon,2k+\ell}\\
&= \frac{1}{2}\left(\sum_{k\in\mathbb{Z}}a\left(\epsilon-2k\right)\,\varphi(\cdot-k)+\sum_{k\in\mathbb{Z}}(-1)^\epsilon\,a\left(1-\epsilon+2k\right)\,\psi(\cdot-k)\right),
\end{aligned}$$

was nichts anderes als die explizite Schreibweise von (7.22) ist, die Projektion auf $\mathbb{S}_2(\varphi)\oplus\mathbb{S}_2(\psi)$ hat also schon einmal die gewünschte Gestalt. Was wir noch zeigen müssen ist, daß tatsächlich $f_\epsilon=g_\epsilon$ ist. Dazu berechnen schreiben wir $f_\epsilon=g_\epsilon\oplus h_\epsilon$, $\langle g_\epsilon,h_\epsilon\rangle=0$ und erhalten, da die Translate von φ und ψ eine Orthonormalbasis des von ihnen aufgespannten Raumes bilden, daß

$$\begin{aligned}
\|h_\epsilon\|_2^2 &= \langle h_\epsilon,h_\epsilon\rangle=\langle g_\epsilon+h_\epsilon,g_\epsilon+h_\epsilon\rangle-\langle g_\epsilon,g_\epsilon\rangle=\|f_\epsilon\|_2^2-\|g_\epsilon\|_2^2=\frac{1}{2}-\|g_\epsilon\|_2^2\\
&= \frac{1}{2}-\frac{1}{4}\left(\sum_{k\in\mathbb{Z}}\left|a\left(\epsilon-2k\right)\right|^2+\sum_{k\in\mathbb{Z}}\left|a\left(1-\epsilon+2k\right)\right|^2\right)
\end{aligned}$$

$$= \frac{1}{2} - \frac{1}{4}\left(\sum_{k\in\mathbb{Z}} |a(2k+\epsilon)|^2 + \sum_{k\in\mathbb{Z}} |a(2k+1-\epsilon)|^2\right) = \frac{1}{2} - \frac{1}{4}\sum_{k\in\mathbb{Z}} |a(k)|^2$$
$$= \frac{1}{2} - \frac{1}{4}\|a\|_2^2 = \frac{1}{2} - \frac{1}{4}\|\varphi * a\|_2^2 = \frac{1}{2} - \frac{1}{4}\underbrace{\left\|\sigma_{1/2}\varphi\right\|_2^2}_{=2} = 0,$$

also $h_\epsilon = 0$, das heißt, $f_\epsilon = g_\epsilon$ □

Beispiel 7.18 *Das einfachste Wavelet ist sicherlich das* "Haar–Wavelet" *zur (trivialerweise orthogonalen) Skalierungsfunktion $\varphi = \chi$. Da hier der Koeffizientenvektor der Verfeinerungsgleichung nur*[191] *$a(0) = a(1) = 1$. Dann ist also $a^\perp(-1) = 1$ und $a^\perp(1) = -1$ und wir erhalten das Wavelet $\chi(2\cdot) - \chi(2\cdot -1)$, eine einfache "Rechteckswelle", siehe Abb 7.3.*

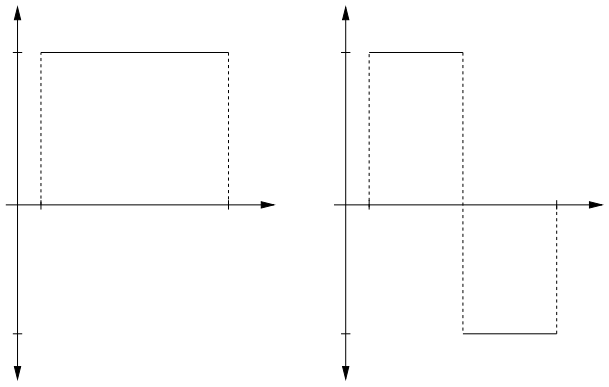


Abbildung 7.3: Die charakteristische Funktion alias "Rechteckspuls" und ihr Wavelet, das "Haar–Wavelet". Der einfachste und Modellfall von Skalierungsfunktion und Wavelet, manchmal allerdings etwas zu einfach.

7.4 Approximation mit Wavelets

Um Aussagen über die Approximationsgüte unsrer Skalenräume machen zu können, werden wir natürlich auf die Theorie der translationsinvarianten Räume zurückgreifen, insbesondere auf Satz 6.29. Um den aber anwenden zu können, müssen wir natürlich annehmen, daß die Skalierungsfunktion, die ja den translationsinvarianten Raum aufspannt, zu $C_{00}(\mathbb{R})$ gehört, also *stetig* ist und *kompakten Träger* hat.

Bemerkung 7.19 *Es ist klar, daß es stetige Funktionen gibt, die Skalierungsfunktionen mit kompaktem Träger sind, nämlich die B–Splines der Ordnung 1 und höher, aber die sind nicht*

[191] Wenn wir Koeffizienten so eines Vektors nicht erwähnen, so sei dies gleichbedeutend damit, daß diese Koeffizienten den Wert 0 haben.

orthonormal. Dafür gibt es stetige Funktionen, die orthonormale Skalierungsfunktionen sind, nämlich beispielsweise die Orthonormalisierung der B–Splines nach Satz 7.8. Und es gibt eine orthonormale Skalierungsfunktion mit kompaktem Träger, nämlich χ, aber die ist nun wieder nicht stetig! Es scheint wie verhext! Gibt es also überhaupt Skalierungsfunktionen, die

1. *stetig sind,*
2. *orthonormal sind,*
3. *kompakten Träger haben?*

Die Antwort ist nicht nur "ja", es gibt sogar Skalierungsfunktionen, die beliebig oft differenzierbar und orthogonal sind und obendrein kompakten Träger haben! Ein Konstruktionsverfahren für solche Wavelets wurde von Ingrid Daubechies erstmals in [15] angegeben. Beispiele für solche Funktionen können in Abb. 7.4 und Abb. 7.5 bewundert werden.

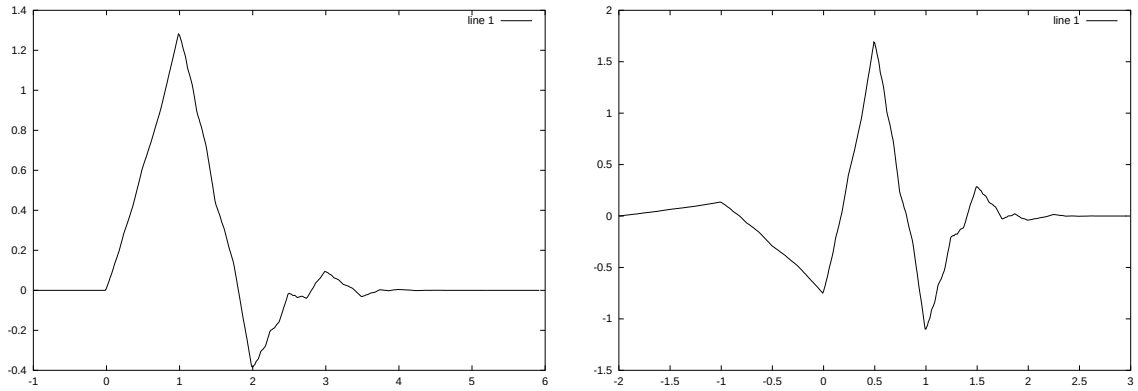


Abbildung 7.4: Die Skalierungsfunktion für das Daubechies–Wavelet D_3 (was auch immer das ist und das zugehörige Wavelet. Beide Funktionen haben orthogonale Translate und kompakten Träger. Stetigkeit sieht man ja ganz gut.

Der folgende Satz über die Approximationsordnung der Räume V_j, $j \in \mathbb{Z}$, ist nun eine unmittelbare Konsequenz aus Satz 6.29.

Satz 7.20 *Erfüllt die Skalierungsfunktion $\varphi \in C_{00}(\mathbb{R})$ die Strang–Fix–Bedingungen der Ordnung n, dann ist für jedes $f \in L_2^{n+1}(\mathbb{R})$*

$$d\left(f, V_j\right) := \inf_{c \in \ell_2(\mathbb{Z})} \|f - \sigma_{2^j}\left(\varphi * c\right)\| \leq C 2^{-j(n+1)} \left\|f^{(n+1)}\right\|_2 \tag{7.23}$$

Bemerkung 7.21 *1. Es gibt eine Vielzahl von Varianten dieser Abschätzung der Approximationsordnung von Skalierungsräumen, in denen beispielsweise die Annahme $\varphi \in C_{00}(\mathbb{R})$ abgeschwächt wird. Außerdem kann man natürlich auch "gebrochene" Regularitätseigenschaften von f und entsprechend Approximationsordnungen der Form $2^{-j\alpha}$ für $\alpha > 0$ betrachten.*

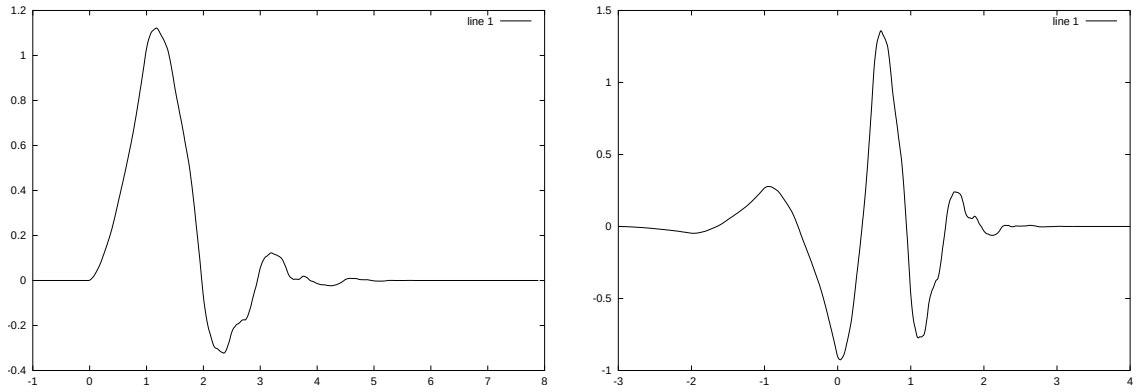


Abbildung 7.5: Skalierungsfunktion und Wavelet zu D_4.

2. *Da φ kompakten Träger hat, liese sich die Aussage auch "lokalisieren" und erhält so auch Aussagen für Funktionen, die beispielsweise nur stückweise differenzierbar sind.*

3. *In Falle einer* Skalierungsfunktion*, die ja per Definitionem stabil ist, werden die Strang–Fix–Bedingungen an φ zu dem einfacheren*

$$\widehat{\varphi}^{(j)}(2k\pi) = 0, \qquad j = 0, \dots, n, \quad k \in \mathbb{Z} \setminus \{0\},$$

denn unter der Annahme der Stabilität ist dann

$$A \le \sum_{k\in\mathbb{Z}} |\widehat{\varphi}(0 + 2k\pi)|^2 = |\widehat{\varphi}(0)|^2 + \sum_{k\in\mathbb{Z}\setminus\{0\}} |\underbrace{\widehat{\varphi}(2k\pi)}_{=0}|^2 = |\widehat{\varphi}(0)|^2 .$$

Aber das ist nur die halbe Wahrheit. Viel besser ist die Tatsache, daß man die Regularität der Funktion, also ihre Differenzierbarkeit, auch von der Abfallrate der Waveletkoeffizienten ablesen kann. Dazu bemerken wir zuerst einmal, daß sich jede Funktion $f \in L_2(\mathbb{R})$ für beliebiges $j \in \mathbb{Z}$ als

$$f = \varphi * c\left(2^j \cdot\right) + \sum_{k=j}^{\infty} \psi * d_k\left(2^k \cdot\right)$$

mit

$$c(\ell) = 2^j \big\langle f, \varphi\left(2^j \cdot - \ell\right)\big\rangle, \qquad d_k(\ell) = 2^k \big\langle f, \psi\left(2^k \cdot - \ell\right)\big\rangle, \qquad \ell \in \mathbb{Z},\ k \ge j,$$

für eine orthonormale Skalierungsfunktion[192] schreiben läßt, wobei man den Vektor $d_k \in \ell_2(\mathbb{Z})$ als Vektor der *Waveletkoeffizienten der Ordnung k* bezeichnet. Um eine "Auswirkung" der Strang–Fix–Bedingung auf die Wavelets zu sehen, erst noch ein bißchen Notation.

[192] Die zweite Identität, die für die Waveletkoeffizienten, gilt aber immer, denn ein Wavelet ist ja als *orthonormaler* Erzeuger von W_0 definiert.

Definition 7.22 *Zu $f \in L_2(\mathbb{R})$ bezeichne*[193]

$$\mu(f) = \left\{ \mu(f)(k) = \int_{\mathbb{R}} t^k \, f(t)\, dt \;:\; k \in \mathbb{N}_0 \right\} \quad \in \ell\,(\mathbb{N}_0)$$

die Momentenfolge *zu f. Wir sagen eine Funktion f habe $n+1$* verschwindende Momente*, wenn*

$$\mu(f)(0) = \cdots = \mu(f)(n) = 0.$$

Proposition 7.23 *Erfüllt die orthonormale Skalierungsfunktion $\varphi \in C_{00}(\mathbb{R})$ die Strang–Fix–Bedingungen der Ordnung n, dann hat das Wavelet ψ mindestens $n+1$ verschwindende Momente.*

Beweis: Seien wieder p_j so, daß $x^j = \varphi * p_j(x)$, dann ist

$$\left\langle (\cdot)^j, \psi \right\rangle \;=\; \int_{\mathbb{R}} t^j \, \psi(t)\, dt = \int_{\mathbb{R}} \varphi * p_j(t)\, \psi(t)\, dt = \sum_{k\in\mathbb{Z}} p_j(k) \underbrace{\int_{\mathbb{R}} \varphi(t-j)\, \psi(t)\, dt}_{=0} = 0.$$

Hierbei ist die Vertauschung von Summe und Integral problemlos, da sowohl φ als auch ψ kompakten Träger haben. □

Zumindest für Wavelets mit kompaktem Träger können wir dann eine recht schöne Aussage machen.

Satz 7.24 *Sei $\psi \in C_{00}(\mathbb{R})$ das Wavelet zu einer Skalierungsfunktion $\varphi \in C_{00}(\mathbb{R})$, die die Strang–Fix–Bedingungen der Ordnung n erfüllt und sei $f \in L_2^{n+1}(\mathbb{R})$. Dann gibt es eine Konstante $C > 0$, so daß*

$$|d_k(\ell)| = 2^k \left| \left\langle f, \psi\left(2^k \cdot - \ell\right) \right\rangle \right| \le C 2^{-k(n+1)} \, \left\| f^{(n+1)} \right\|_2 \tag{7.24}$$

Beweis: Wir bezeichnen mit Ψ_j, $j = 0, \ldots, n$, die j–te Stammfunktion zu ψ, definiert durch

$$\Psi_0 := \psi \qquad \Psi_{j+1}(x) := \int_{-\infty}^{x} \Psi_j(t)\, dt, \quad x \in \mathbb{R}, \quad j = 0, \ldots, n.$$

Wir zeigen zuerst, daß Ψ_j kompakten Träger hat und daß[194]

$$\int_{\mathbb{R}} t^k \, \Psi_j(t)\, dt = 0, \qquad k = 0, \ldots, n-j, \quad j = 0, \ldots, n, \tag{7.25}$$

[193] Eigentlich sollten wir zuerst einmal voraussetzen, daß f tatsächlich so "gebaut" ist, daß diese Folge tatsächlich existiert! Kompakter Träger von f wäre beispielsweise eine schöne Eigenschaft. Wir wollen aber hier etwas großzügig mit den Voraussetzungen sein, sollten uns aber darüber im Klaren sein, daß "eigentlich" Vorsicht geboten ist!

[194] Für $j = n+1$ ist dies trivialerweise erfüllt!

was wir durch Induktion über j nachweisen werden. Da ψ nach Proposition 7.23 $n+1$ verschwindende Momente hat, folgt (7.25) für $j=0$ aus

$$0 = \int_{\mathbb{R}} t^n \, \psi(t) \, dt;$$

der kompakte Träger von ψ war außerdem sogar ein Stück der Annahme. Für den Induktionsschritt $j \to j+1$, $j+1 \le n$, nehmen wir an, daß $\Psi_j(x) = 0$ für $x \notin [a,b]$ und dann ist

$$\Psi_{j+1}(x) = \int_{-\infty}^{x} \Psi_j(t) = \begin{cases} 0, & x \le a, \\ \Psi_{j+1}(b), & x \ge b, \end{cases}$$

und da

$$\lim_{x\to\infty} \Psi_{j+1}(x) = \int_{\mathbb{R}} \Psi_j(t) \, dt = \int_{\mathbb{R}} 1 \, \Psi_j(t) \, dt = 0$$

nach (7.25), ist auch $\Psi_{j+1}(x) = 0$ für $x \notin [a,b]$. Für $k = 0, \ldots, n-j-1$ ist dann, mit partieller Integration,

$$\int_{\mathbb{R}} t^k \, \Psi_{j+1}(t) \, dt = \frac{1}{k+1} \left(t^{k+1} \, \Psi_{j+1}(t) \big|_{t=-\infty}^{\infty} - \int_{\mathbb{R}} t^{k+1} \, \Psi_j(t) \, dt \right)$$

und der erste Term verschwindet, da Ψ_{j+1} kompakten Träger hat, der zweite hingegen nach Induktionsannahme. Damit ist (7.25) bewiesen.
Nun verwenden wir wieder partielle Integration, um

$$\begin{aligned} \left\langle f, \psi\left(2^k \cdot - \ell\right) \right\rangle &= \int_{\mathbb{R}} f(t) \, \psi\left(2^k t - \ell\right) \, dt \\ &= 2^{-k} \underbrace{f(t) \, \Psi_1\left(2^k t - \ell\right) \big|_{t=-\infty}^{\infty}}_{=0} - 2^{-k} \int_{\mathbb{R}} f'(t) \, \Psi_1\left(2^k t - \ell\right) \, dt \\ &\vdots \\ &= (-1)^{n+1} 2^{-k(n+1)} \int_{\mathbb{R}} f^{(n+1)}(t) \, \Psi_{n+1}\left(2^k t - \ell\right) \, dt, \end{aligned}$$

also ist

$$\begin{aligned} |d_k(\ell)| &\le \|\Psi_{n+1}\|_\infty \, 2^{-k(n+1)} \int_{\mathbb{R}} \chi_{[a,b]}\left(2^k t - \ell\right) \left| f^{(n+1)}(t) \right| \, dt \le \|\Psi_{n+1}\|_\infty \, 2^{-k(n+1)} \underbrace{\left\|\chi_{[a,b]}\right\|_2}_{=\sqrt{b-a}} \left\| f^{(n+1)} \right\|_2 \\ &= \underbrace{\|\Psi_{n+1}\|_\infty \sqrt{|b-a|}}_{=:C} \, 2^{-kn} \left\| f^{(n+1)} \right\|_2 . \end{aligned}$$

□

Bemerkung 7.25 *Die Forderung in Satz 7.24, daß das Wavelet kompakten Träger haben soll, ist natürlich eine Voraussetzung, die erst einmal erfüllt sein will. Trotzdem kann man einiges dazu sagen.*

1. *In (7.24) findet sich nur eine Approximationsordnung von 2^{-kn} während wir für die Approximationsordnung des Raums V_k in Satz 7.20 ja die Ordnung $2^{-k(n+1)}$ bekommen haben. Es ist aber nichts faul mit den Wavelets, wie man auch aus dem Beweis von Satz 7.24 ersehen kann, sondern es ist lediglich eine Normierungsfrage: Die skalierten Wavelets $\psi\left(2^k \cdot -\ell\right)$ erfüllen schließlich "nur"*

$$\left\langle \psi\left(2^k \cdot -\ell\right), \psi\left(2^k \cdot -\ell'\right)\right\rangle = 2^{-k}\,\delta_{\ell,\ell'}, \qquad \ell,\ell' \in \mathbb{Z},$$

weswegen sie eigentlich zu

$$\widetilde{\psi}_{k,\ell} := 2^{k/2}\,\psi\left(2^k \cdot -\ell\right), \qquad k,\ell \in \mathbb{Z},$$

*umskaliert werden müssten, um eine Orth*normal*basis zu bilden. Verwendet man diese Funktionen, so ergibt sich dann auch die "erwartete" Approximationsordung $2^{-k(n+1)}$.*

2. *Es gibt Wavelets mit kompaktem Träger, nämlich die bereits erwähnten Daubechies–Wavelets, die sogar verhältnismäßig einfach zu konstruieren sind*[195]*, ihre Verfeinerungskoeffizienten, um genau zu sein, denn man kennt weder geschlossene Formeln für diese Funktionen, noch für ihre Fouriertransformierte! Das macht aber auch nichts, denn diese Koeffizienten reichen für das praktische Arbeiten mit Wavelets vollkommen aus. Außerdem sind die Verfeinerungskoeffizienten zu Daubechies–Wavelets kleiner Ordnung sogar in [16, 54] aufgelistet.*

3. *Wir können den kompakten Träger der Wavelets sogar als Stärke sehen: Hat ψ den Träger $[a,b]$, so hat $\psi\left(2^k \cdot -\ell\right)$ den Träger $2^{-k}\left(\ell + [a,b]\right)$ und berücksichtigt so nur das* lokale *Glattheitsverhalten der Funktion f um den Punkt $2^{-k}\ell$. Hat also beispielsweise eine Funktion einen "Knick" an einer Stelle x^* und ist ansonsten glatt, dann werden die Waveletkoeffizienten $d_k\left(\lfloor 2^k x \rfloor\right)$ deutlich langsamer abfallen. Das kann man beispielsweise zur automatischen Erkennung von Singularitäten nutzen, siehe Abb. 7.6.*

4. *Aber auch bei der Kompression hilft das: Ist nämlich die Funktion f "weitestgehend" glatt, dann werden die meisten Wavelet–Koeffizienten ziemlich schnell ziemlich klein werden und wir können sie einfach vergessen und nur die betragsmäßig größten Koffizienten oder alle, deren Absolutbetrag oberhalb eines bestimmten Wertes liegt*[196] *behalten. Der Vorteil solcher Kompressionsverfahren liegt darin, daß die Rekonstruktion oftmals kaum sichtbare Unterschiede zum Original aufweist.*

5. *Was aber tun mit Skalierungsfunktionen wie Splines, bei denen wir entweder nur ein Prewavelet mit kompaktem Träger oder aber ein Wavelet mit unendlichem Träger bekommen? Nun, abgesehen davon, daß man wegen des Abfalls von Funktion und Koeffizienten numerisch einfach "abschneiden" könnte, gibt es auch noch* biorthogonale *Wavelets, siehe z.B. [54] oder [49].*

[195] Aber trotzdem gehört dieser Punkt in eine Spezialvorlesung über Wavelets.
[196] Dieses Verfahren bezeichnet man als *"Threshholding"*.

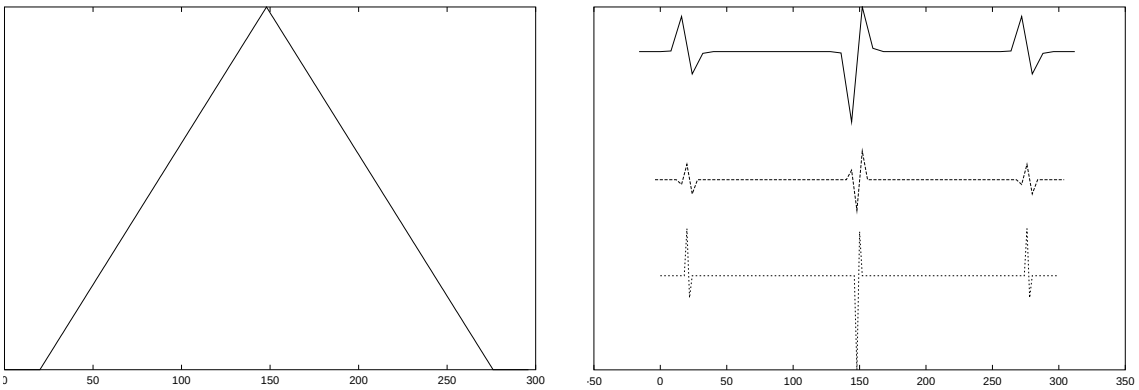


Abbildung 7.6: Die Funktion auf der linken Seite hat offensichtlich zwei Knicke, ansonsten ist sie linear. Auf der rechten Seite sind die Wavelet–Koeffizienten der “ersten drei” Ordnungen geplottet (mit nach unten zunehmender Auflösung), die bereits mit einem Faktor 2^k nachskaliert wurden. Sie weisen schon ziemlich deutlich auf die drei kritischen Stellen hin. Ach ja: Verwendet wurde hierbei das Daubechies–Wavelet D_3.

Bleibt noch die Frage: Wie rechnet man eigentlich mit Wavelets, wie implementiert man sie numerisch am Computer. Das Stichwort hier heisst *Filterbänke* und Informationen zu diesem Thema finden sich beispielsweise in [16, 54, 81, 83], aber wegen der ansonsten zu großen Überlappung mit [71] wollen wie hier nicht weiter darauf eingehen – das ist zwar schade, aber was will man machen?

Dicet quis: "enuclea!
quid est hoc, quod ais?"
Wirft einer ein: "Erläutere das! Was
meinst Du damit?"

Carmina Burana, 226, *Über den Zustand der Welt*

8 Der Satz von Kolmogoroff

Wir wollen uns nochmal einem eher "theoretischen" Thema zuwenden, nämlich der Frage, ob es überhaupt wirklich multivariate Funktionen gibt. Die erste, eher verblüffte Antwort ist: "Natürlich", denn bereits $f(x, y) = x + y$ ist natürlich eine Funktion in zwei Variablen. Aber eigentlich auch wieder nicht wirklich, denn schließlich ist f ja "nur" Summe von zwei einfachen Funktionen in einer Variablen und Addieren ist ja nun nicht gerade die hohe Schule. Die kompliziertere Funktion

$$f(x, y) = xy = e^{\log x + \log y}$$

ist auch "nur" eine Summe, die allerdings noch durch eine **univariate** Funktion geschickt werden muss, um die gewünschte Funktion zu ergeben. Gibt es also "richtige" Funktionen in mehreren Variablen oder ist alles auf so eine Art darstellbar? Die Antwort ist, überraschenderweise, letzteres. Na gut, aus Sicht der stetigen Funktionen gibt es sowieso keine Dimension.

8.1 Nomographie, Hilberts 13. Problem und Kolmogoroffs Lösung

Beginnen wir mit etwas ganz anderem, nämlich mit der *Nomographie*, also der *zeichnerischen* bzw. *grafischen* Lösung mathematischer Probleme. Die einfachste Anwendung der Nomographie ist logarithmisches Papier, also Papier, das in einer Richtung linear, in einer logarithmisch liniert ist und mit dem man exponentielle Gesetzmäßigkeiten darstellen kann, siehe Abb. 8.1. Um diese Gerade zu bestimmen, braucht man nur zwei Werte $y(x_1)$ und $y(x_2)$ zu kennen und kann dann für jedes x den zugehörigen y–Wert auf der logarithmischen Skalenachse ablesen. Diese grafische "Lösungsmethode" hat, vor allem in der Prätaschenrechnerperiode, zwei große Vorteile:

- Man muss nicht rechnen können, zumal die Berechnung komplexer Funktionen wie dem Logarithmus ohnehin Tabellenwerke plus Interpolation fordern würde, siehe [1].
- Das Ergebnis hat automatisch die richtige *relative* Genauigkeit, die man braucht, es geht also wieder mal um gültige Stellen, ganz genau so, wie man es aus der "normalen" Numerik kennt [31, 69].

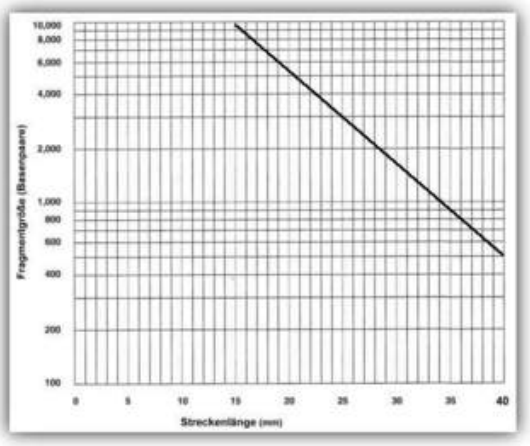


Abbildung 8.1: Logarithmisches Papier mit einem Zusammenhang der Form $y(x) = C_0 + C_1 e^{-x}$, wie sie beispielsweise bei Ladezeitberechnungen von Kondensatoren auftreten.

Man sollte allerdings Nomographie nicht unterschätzen – solche grafischen Verfahren gibt es von fast beliebiger Komplexität, denn die Kurve mit deren Hilfe man das Problem zu lösen versucht, muß beileibe keine Gerade sein, siehe Abb. 8.2.

Kurvenbasiert kann Nomographie[197] natürlich immer nur **bivariat** funktionieren, also Funktionen $y = f(x)$ bestimmen. Wie ist es aber nun mit $y = f(x_1, \ldots, x_2)$? Um da Nomographie anwenden zu können, müssen wir alles auf die Kaskadierung von *univariaten* Funktionen und *einfachen* Rechenoperationen wie beispielsweise die Addition zurückführen.

Und genau das bringt uns nun zu David Hilbert[198], der auf dem internationalen Mathematikerkongress 1900 in Paris eine Rede mit den 23 seiner Meinung nach bedeutendsten offenen Problemen der Mathematik hielt. Das 13te darunter lautet wie folgt[199]:

> Wir kommen nun zur Algebra; ich nenne im Folgenden ein Problem aus der Gleichungstheorie und eines, auf welches mich die Theorie der algebraischen Invarianten geführt hat.
>
> **13. Unmöglichkeit der Lösung der allgemeinen Gleichung 7ten Grades mittelst Functionen von nur 2 Argumenten.**

[197] Wir wollen uns jetzt nicht auf holographische 3D–Nomographie kaprizieren, denn inzwischen gibt es ja bessere numerische Methoden, um Sachen auszurechnen.

[198] David Hilbert, 23.1.1861–14.2.1943, der wohl profilierteste deutsche Mathematiker nach Gauß, oder um [51], die beste Internetquelle zur Geschichte der Mathematik zu zitieren: *"Hilbert's work in geometry had the greatest influence in that area after Euclid. A systematic study of the axioms of Euclidean geometry led Hilbert to propose 21 such axioms and he analysed their significance. He made contributions in many areas of mathematics and physics"*.

[199] Das ist dann auch der Originaltext.

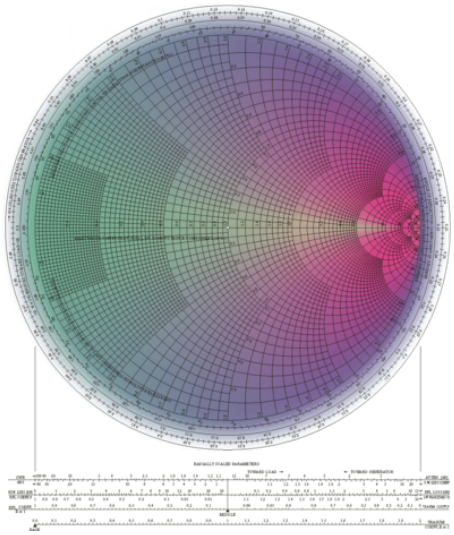


Abbildung 8.2: Ein sogenanntes *Smith–Chart* zur Berechnung der komplexen Impedanz einer Transmissionsleitung. Quelle: Wikipedia.

Die Nomographie M. d'Ocagne, Trait de Nomographie, Paris 1899 hat die Aufgabe Gleichungen mittelst gezeichneter Curvenschaaren zu lösen, die von einem willkürlichen Parameter abhängen. Man sieht sofort, da jede Wurzel einer Gleichung, deren Coefficienten nur von zwei Parametern abhängen, d. h. jede Function von zwei unabhängigen Veränderlichen auf mannigfache Weise durch das der Nomographie zu Grunde liegende Princip darstellbar ist. Ferner sind durch dieses Princip offenbar auch eine große Klasse von Functionen von drei und mehr Veränderlichen darstellbar, nämlich alle diejenigen Functionen, die man dadurch erzeugen kann, daß man zunchst eine Function von zwei Argumenten bildet, dann jedes dieser Argumente wieder gleich Functionen von zwei Argumenten einsetzt, an deren Stelle wiederum Functionen von zwei Argumenten treten u. s. f., wobei eine beliebige endliche Anzahl von Einschachtelungen der Functionen zweier Argumente gestattet ist. So gehört beispielsweise jede rationale Function von beliebig vielen Argumenten zur Klasse dieser durch nomographische Tafeln construirbaren Functionen; denn sie kann durch die Prozesse der Addition, Subtraction, Multiplikation und Division erzeugt werden, und jeder dieser Prozesse repräsentirt eine Function von nur zwei Argumenten. Man sieht leicht ein, daß auch die Wurzeln aller Gleichungen, die in einem natürlichen Rationalitätsbereiche durch Wurzelziehen auflösbar sind, zu der genannten Klasse von Functionen gehören; denn hier kommt zu den vier elementaren Rechnungsoperationen nur noch der Prozeß des Wurzelziehens hinzu, der ja lediglich eine Function eines Argumentes repräsentirt. Desgleichen sind die allgemeinen Gleichungen 5ten und 6ten Grades durch geeignete nomographische Tafeln auflösbar; denn diese können durch solche Tschirn-

> hausentransformationen, die ihrerseits nur Ausziehen von Wurzeln verlangen, in eine Form gebracht werden, deren Coefficienten nur von zwei Parametern abhngig sind.
>
> Wahrscheinlich ist nun die Wurzel der Gleichung 7ten Grades eine solche Function ihrer Coefficienten, die nicht zu der genannten Klasse nomographisch construirbarer Functionen gehrt, d. h. die sich nicht durch eine endliche Anzahl von Einschachtelungen von Functionen zweier Argumente erzeugen lt. Um dieses einzusehen, wre der Nachweis dafr nötig, daß die Gleichung 7ten Grades
>
> $$f^7 + x\, f^3 + y\, f^2 + z\, f + 1 = 0$$
>
> nicht, mit Hülfe beliebiger stetiger Functionen von nur zwei Argumenten lösbar ist. Daß es überhaupt analytische Functionen von drei Argumenten x, y, z giebt, die nicht durch endlich-malige Verkettung von Functionen von nur zwei Argumenten erhalten werden können, davon habe ich mich, wie ich noch bemerken möchte, durch eine strenge Ueberlegung überzeugt.

Es mag ein wenig seltsam erscheinen, daß Hilbert hier ein Problem angibt, das in *zwei Variablen* zu lösen zu sein scheint, in *drei Variablen* hingegen unlösbar sein soll. Sowas ist aber bei multivariaten Polynomen gar nicht einmal so außergewöhnlich, es gibt da viele Situationen, wo sich der trivariate Fall noch mal **signifikant** vom bivariaten unterscheidet, sei es geometrisch oder algebraisch. Dennoch – hier hatte Hilbert unrecht, denn Kolmogoroff zeigte 1957 in[200] [38], bzw. sein Schüler Arnol'd in [4], daß sich *jede* stetige Funktion als Überlagerung univariater Funktionen unter ausschließlicher Verwendung der Addition (und Konkatenation natürlich) darstellen läßt.

Satz 8.1 (Satz von Kolmogoroff – Originalversion) *Es gibt $s(2s+1)$ stetige Funktionen ϕ_{jk}, $j = 0, \ldots, 2s$, $k = 1, \ldots, s$, dergestalt, daß man zu jeder stetigen Funktion $f \in C\left([0,1]^s\right)$ stetige Funktionen $g_0, \ldots, g_{2s}$ finden kann, für die*

$$f\left(x_1, \ldots, x_s\right) = \sum_{j=0}^{2s} g_j \left(\sum_{k=1}^{s} \phi_{jk}\left(x_k\right) \right) \tag{8.1}$$

gilt.

Dieses Resultat wurde dann im Laufe der Zeit noch verfeinert, im wesentlichen von Lorentz[201] [46] und David Sprecher [77, 76, 78]. In der Tat stellte sich heraus, daß man bereits mit einer Funktion g auskommt, die von f abhängt und daß man die die Funktionen ϕ_{jk}, $k = 1, \ldots, s$ als passende Vielfache von Funktionen ϕ_j, $j = 0, \ldots, 2s$, wählen kann. Außerdem lassen sich diese Funktion sogar "kontrolliert" stetig, genauer gesagt Lipschitz–stetig konstruieren. Die "endgültige" Version des Satzes von Kolmogoroff, wie sie auch in [47] zu finden ist, sieht dann folgendermaßen aus.

[200] Diese Publikationsform verdient eine eigene Fußnote! Die guten alten *Doklady* der UDSSR waren ein mathematisches Veröffentlichungsmedium, in denen Resultate *angekündigt* wurden, zumeist **ohne** Beweis, der dann später in einer langen Veröffentlichung anderswo "nachgereicht" wurde – oder eben auch nicht.

[201] Den wir ja auch schon kennen

Satz 8.2 (Satz von Kolmogoroff) *Es gibt Konstanten $\lambda_k \in (0,1]$, $k = 1, \ldots, s$, und Funktionen ϕ_j, $j = 0, \ldots, 2s$, mit den folgenden Eigenschaften:*

1. *Die Funktionen ϕ_j sind Lipschitz–stetig für einen passenden Exponenten $\alpha > 0$ und strikt monoton steigend.*

2. *Zu jedem $f \in C([0,1]^s)$ gibt es ein $g \in C([0,s])$, so daß*

$$f(x_1, \ldots, x_s) = \sum_{j=0}^{2s} g\left(\lambda_1 \phi_j(x_1) + \cdots + \lambda_s \phi_j(x_s)\right). \tag{8.2}$$

Bemerkung 8.3 *Bevor wir uns an den Beweis dieser Aussage machen, sehen wir uns nochmal genau an, was wirklich wovon abhängt.*

1. *Sowohl die Funktionen $\phi_0, \ldots, \phi_{2s}$ wie auch die Werte $\lambda_0, \ldots, \lambda_{2s}$ sind **universell**, hängen also nicht von der darzustellenden Funktion f ab, sondern nur von der Dimensionalität des Problems. Sie lassen sich unabhängig vom Kolmogoroffschen Satz auch als Lösung gewisser* Einbettungsprobleme *geometrisch interpretieren, siehe [47, S. 169].*

2. *Das Einzige, was wirklich von f abhängt, ist die Funktion g. Wir werden sehen, daß hier auch ein bisschen die Crux des* Darstellungs*satzes 8.2 liegt, denn g wird Grenzwert einer Folge von Approximationsfunktionen sein.*

3. *"Praktisch" kann man also auch den Satz von Kolmogoroff weniger als exakte Darstellung realisieren, sondern eben auch nur als **Approximation** der multivariaten Funktion f durch Superposition.*

8.2 Von Würfeln und Intervallen

Dimension ist eine interessante Sache. Einerseits ist der Würfel $[0,1]^s$ ja ein s–dimensionales Objekt, andererseits lässt er sich aber sehr einfach bijektiv auf das Intervall $[0,1]$ abbilden. Dazu betrachten wir zu $x \in [0,1]$ und $B \in \mathbb{N}$, $B > 1$, die *B–adische Entwicklung*

$$x = .\xi_1 \xi_2 \ldots = \sum_{j=1}^{\infty} \xi_j B^{-j}, \qquad \xi_j \in \{0, \ldots, B-1\}.$$

Mit deren Hilfe können wir nun jedem $x = (x_1, \ldots, x_s) \in [0,1]^s$ mit dyadischen Entwicklungen

$$x_j = .\xi_{j,1} \xi_{j,2} \ldots, \qquad j = 1, \ldots, s,$$

den Punkt

$$[0,1] \ni x^* = \theta(x) = .\xi_{1,1} \ldots \xi_{s,1} \xi_{1,2} \ldots \xi_{s,2} \ldots = \sum_{j=1}^{\infty} B^{-s(j-1)} \sum_{k=1}^{s} \xi_{j,k} B^{-k}$$

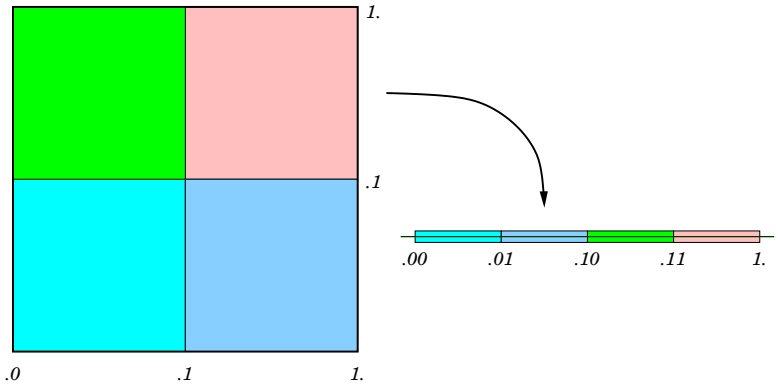


Abbildung 8.3: Die Abbildung θ für $s = 2$ und $B = 2$. Da man in zwei übereinanderliegenden Teilquadraten Punkte wählen kann, die beliebig dicht beisammen liegen, deren Bilder aber in deutlich getrennten Intervallteilen landen, kann θ nicht stetig sein.

zuordnen. Damit wäre $g = f \circ \theta^{-1}$ eine univariate Funktion, die f darstellt, nur leider ist θ nicht stetig, siehe Abb. 8.3 und damit g entsprechend kompliziert. Wir müssen uns also etwas besseres einfallen lassen.

Die Konstruktion der "magischen" Funktionen ϕ_j aus Satz 8.2 basiert auf der Konstruktion eines leicht pathologischen Objekts, nämlich einer *monoton steigenden Funktion, deren Ableitung fast überall Null ist*[202]. Und in der Tat ist die Konstruktion solch einer Funktion gar nicht mal schwer. Zu einer Basis $B \geq 3$ und $k \in \mathbb{N}$ definieren wir die *Intervalle der Stufe* k

$$I_j^k := j\, B^{1-k} + B^{-k}\,[1, B-1]\,, \qquad j = 0, \ldots, B^{k-1}. \tag{8.3}$$

Abgesehen von $I_{B^k}^k$, das wir aber der Vollständigkeit halber mitdefinieren wollen, sind das alles Teilintervalle von $[0, 1]$, die an den Abtastpunkten mit Schrittweite $1/B^{k-1}$ beginnen und die Breite $(B-2)/B^k < 1/B^{k-1}$ haben – damit sind alle Intervalle der Stufe k disjunkt und überdecken $[0, 1]$ **nicht** vollständig.

Beispiel 8.4 *Für* $B = 10$ *haben wir auf den Stufen* 1 *bis* 3 *die folgenden Intervalle in* $[0, 1]$

$$\begin{gathered} I_0^1 = [.1, .9] \\ I_0^2 = [.01, .09],\ \ldots,\ I_9^2 = [.91, .99] \\ I_0^3 = [.001, .009],\ \ldots,\ I_{99}^2 = [.991, .999] \end{gathered}$$

Am einfachsten beschreibt man diese Intervalle über Ziffern! Sei[203] $\mathbb{Z}_B := \mathbb{Z}/B\mathbb{Z} \simeq \{0, \ldots, B-$

[202] Dies Funktion kann natürlich nicht mehr stetig differenzierbar sein, außer sie wäre konstant

[203] Wie immer sei hier auf den Unterschied hingewiesen: Eigentlich ist – ganz analog wie beim Torus – $\mathbb{Z}_B$, sprich "$\mathbb{Z}$ modulo B", nicht nur die Menge $\{0, \ldots, B-1\}$, sondern diese Menge zusammen mit den Rechenregeln modulo B, also einer wohldefinierten Addition und Multiplikation.

$1\}$ und $\beta \in \mathbb{Z}_B^{k-1}$ eine *B–adische Zifferndarstellung*. Dann ist, mit der Abkürzung[204] $\rho = B-1$,

$$I(\beta) = [.\beta_1 \cdots \beta_{k-1}1, .\beta_1 \cdots \beta_{k-1}\rho]\,, \qquad \ell(\beta) := k-1,$$

ein derartiges Intervall der Länge $(B-2)\,B^{-(\ell(\beta)+1)}$. Mit Hilfe dieser Intervalle konstruieren wir nun eine Funktion ψ wie folgt:

1. Wir setzen $\psi(0) = 0$ und $\psi(1) = 1$.
2. Auf dem Intervall der Stufe 1, also $I() = I_0^1$ setzen wir
$$\psi(x) = \frac{1}{2}, \qquad x \in I(0).$$
3. Vor und hinter dieses Intervall passen genau die beiden Intervalle $I_0^2 = I(0)$ und $I_{B-1}^2 = I(\rho)$, auf denen wir ψ den Wert $\frac{1}{4}$ bzw. $\frac{3}{4}$ geben.
4. Die nun verbleibenden Lücken füllen wir dann mit
$$\psi(x) = \begin{cases} \frac{1}{8}, & x \in I\,(0,0)\,, \\ \frac{3}{8}, & x \in I\,(0,\rho)\,, \\ \frac{5}{8}, & x \in I\,(\rho,0)\,, \\ \frac{7}{8}, & x \in I\,(\rho,\rho)\,, \end{cases}$$
auf.
5. Jetzt ist es nicht mehr schwer die allgemeine Formel für $k \in \mathbb{N}$ zu raten, wir setzen nämlich für $k \in \mathbb{N}$ und $\beta \in \{0,\rho\}^{k-1}$
$$\psi(x) = 2^{-k} + \sum_{j=1}^{k-1} \underbrace{\frac{\beta_j}{\rho}}_{=:\epsilon_j \in \{0,1\}} 2^{-j} = .\epsilon_1 \ldots \epsilon_{k-1}1, \qquad x \in I(\beta). \tag{8.4}$$

Damit ist ψ auf den disjunkten Intervallen $I(\beta)$, $\beta \in \{0,\rho\}^{k-1}$, $k \in \mathbb{N}$, definiert. Auf Stufe k gibt es 2^{k-1} dieser Intervalle der Länge $(B-2)B^{-k}$, das heißt, die Intervalle auf Stufe k überdecken einen Bereich der Länge $\frac{B-2}{B}\left(\frac{2}{B}\right)^{k-1}$ und die Länge des Gesamtbereichs ist

$$\frac{B-2}{B}\sum_{k=1}^{\infty}\left(\frac{2}{B}\right)^{k-1} = \frac{B-2}{B}\sum_{k=0}^{\infty}\left(\frac{2}{B}\right)^{k} = \frac{B-2}{B}\frac{1}{1-2/B} = 1,$$

die Funktion ψ ist also *fast überall* definiert. Nun erweitern wir sie stetig zu der offensichtlich monoton steigenden Funktion ϕ, indem wir

$$\phi(x) := \max_{0\le y\le x} \psi(y), \qquad x \in [0,1] \tag{8.5}$$

setzen.

[204] Wie schon in [69] soll diese Abkürzung auf den Fall $B = 10$ anspielen, da der Buchstabe ρ der Ziffer 9 am ähnlichsten sieht.

Lemma 8.5 *Die Funktion ϕ aus (8.5) ist stetig, monoton steigend und erfüllt fast überall $\phi'(x) = 0$.*

Beweis: Auf den Intervallen $I(\beta)$, $\beta \in \{0, \rho\}^{k-1}$, $k \in \mathbb{N}$, ist ϕ ebenso wie ψ konstant, also im Inneren dieser Intervalle differenzierbar mit Ableitung 0. Da diese Intervalle zusammen aber Maß 1 haben, gilt diese Eigenschaft damit überall.

Bleibt die Stetigkeit. Dazu betrachten wir die B–adische Entwicklung

$$x = .\xi_1\xi_2\ldots$$

eines Punktes $x \in [0, 1]$. Ist nun $\xi_1 \in \{1, \ldots, B-2\}$, dann gehört x zu $I()$, in dessen Innerem ϕ konstant, also insbesondere stetig ist. "Problematisch" sind nur die beiden Randpunkte

$$.10\ldots = .0\rho\ldots \qquad \text{und} \qquad .(\rho-1)\rho\ldots = .\rho0\ldots.$$

Mittels der Invervalle $I(0)$ und $I(\rho)$ erhalten wir dann auch Stetigkeit an allen Punkten der Form $.0\xi_2\ldots$ bzw. $.\rho\xi_2\ldots$, $\xi_2 \in \{1, \ldots, B-2\}$ wieder bis auf die Randpunkte

$$.010\ldots = 0.00\rho\ldots, \qquad \text{und} \qquad .0(\rho-1)\rho\ldots = .0\rho0\ldots$$

sowie

$$.\rho10\ldots = 0.\rho0\rho\ldots, \qquad \text{und} \qquad .\rho(\rho-1)\rho\ldots = .\rho\rho0\ldots.$$

Setzen wir das weiter fort, so stellen wir fest, daß Stetigkeit nur an den Punkten der Form $x = .\xi_1\xi_2\ldots$ mit $\xi_j \in \{0, \rho\}$ nachzuprüfen ist. Diese Punkte sind entweder linker Randpunkt eines Intervalls, wenn ρ unendlich oft wiederholt wird oder rechter Randpunkt, wenn 0 unendlich oft auftaucht[205]. Sehen wir uns mal die Definition von $\psi(x)$ nach (8.4) für $k \in \mathbb{N}$ so einen linken Randpunkt an:

$$x = .\xi_1\ldots\xi_{k-1}0\rho\ldots \qquad \Rightarrow \qquad \psi(x) = .\epsilon_1\ldots\epsilon_{k-1}1, \qquad \epsilon_j = \frac{\xi_j}{\rho}, \quad j = 1, \ldots, k-1.$$

Nun wird x aber durch die Intervalle $I(\xi_1, \ldots, \xi_{k-1}, 0, \rho, \ldots, \rho)$ angenähert, auf denen ψ und damit dann auch ϕ den Wert $.\xi_1, \ldots, \xi_{k-1}, 0, 1, \ldots, 1$ hat, was gegen $\psi(x) = \epsilon_1\ldots\epsilon_{k-1}1$ konvergiert. Für die rechten Randpunkte argumentiert man entsprechend. □

Bemerkung 8.6 *Die Funktion ϕ, die wir hier kennengelernt haben, hat eine recht faszinierende Eigenschaft: Alle Intervalle der Konstruktion werden auf* dyadische *Zahlen abgebildet. Die Gesamtheit der Intervalle ist eine Menge vom Maß 1, die dyadischen Zahlen hingegen sind eine Menge vom Maß 0. Und damit:*

> *Die Funktion ϕ ist eine monoton steigend Bijektion von $[0, 1]$ auf sich selbst, die eine Menge vom Maß 0 auf eine Menge vom Maß 1 abbildet und umgekehrt.*

[205] In beiden Fällen bestehen die Zahlen genaugenommen aus einem endlichen Teil und ab einem bestimmten Punkt kommen dann nur noch Nullen oder Einsen.

Wie schon gesagt – so viel ist Stetigkeit eben auch nicht.

Man kann für die Funktion ϕ übrigens auch recht einfach einen Auswertungsalgorithmus angeben, indem man einfach die dyadische Entwicklung berücksichtigt. Bei geeigneter Wahl[206] von B hat ja jede auf dem Rechner darstellbare Zahl x eine **endliche** B–adische Entwicklung $.\xi_1 \dots \xi_N$, die sich mit der rekursiven Vorschrift

$$\phi(x) = \begin{cases} \frac{1}{2}, & \xi_1 \in \{1, \dots, B-2\}, \\ \frac{1}{2}\phi\left(.\xi_2 \dots \xi_N\right), & \xi_1 = 0, \\ \frac{1}{2} + \frac{1}{2}\phi\left(.\xi_2 \dots \xi_N\right), & \xi_1 = B-1, \end{cases}$$

die sich durchaus praktisch implementieren lässt, siehe Abb 8.4.

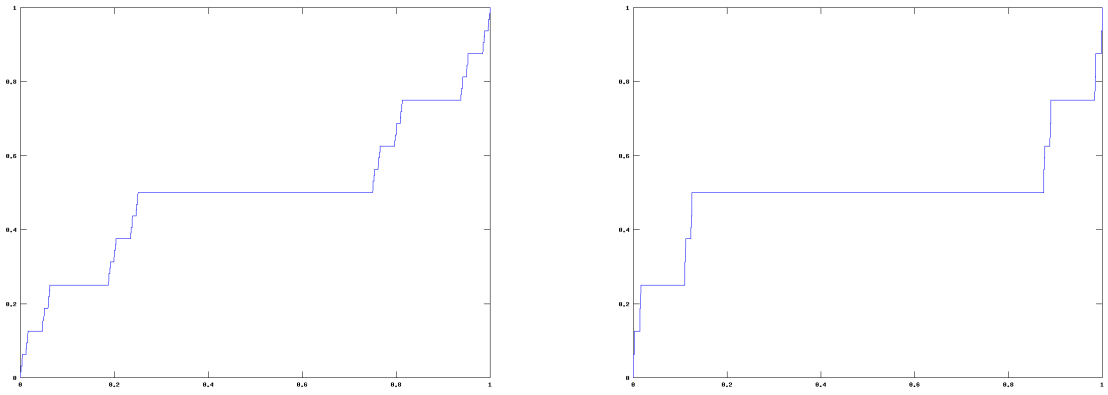


Abbildung 8.4: Die Funktion ϕ zu den Basen $B = 4$ *(links)* und $B = 8$ *(rechts)*. In beiden Fällen kann man die fraktale Natur der Funktion gut erkennen.

Die Funktionen ϕ_j aus (8.2) konstruiert man fast genauso[207], allerdings mittels leicht verschobener Intervalle. Für $k \in \mathbb{N}$ und $\beta \in \mathbb{Z}_B^{k-1}$ definieren wir nämlich[208]

$$I_j(\beta) = I(\beta) - j\,\frac{B-2}{2s}B^{-k}, \qquad j = 0, \dots, 2s, \tag{8.6}$$

und konstruieren ϕ_j aus wie oben aus den Intervallen $I_j(\beta) \cap [0,1]$. Ausserdem fixieren wir jetzt auch noch die Basis B, und zwar als

$$B = 4s + 2, \tag{8.7}$$

so daß

$$I_j(\beta) = I(\beta) - 2jB^{-k}, \qquad j = 0, \dots, 2s, \tag{8.8}$$

[206] Also beispielsweise nicht $B = 10$, wenn man es mit einem "normalen" Computer zu tun hat, der dual rechnet.
[207] Warum sonst hätten wir uns mit dieser Konstruktion so lange aufhalten sollen?
[208] Wir verschieben also die Intervalle geleichmäßig um das $j/2s$–fache der Intervalllänge $(B-2)/B^k$.

ist. Der Grund hierfür ist einfach: Für $\beta \in \mathbb{Z}_B^{k-1}$ sind die Randpunkte des Intervalls $I_j(\beta)$ von der Form

$$\sum_{\ell=1}^{k-1} \beta_\ell B^{-\ell} + (\beta_k - 2j)\ B^{-k} = \left(\sum_{\ell=1}^{k} B^{k-\ell}\beta_\ell - 2j \right) B^{-k} = NB^{-k}$$

für ein $N \in \mathbb{N}$ und damit ist kein Randpunkt der Stufe k auch Randpunkt der Ordnung k' mit $k' > k$, denn letztere verteilen sich ja auf einem feineren Gitter.

8.3 Der Beweis

So, jetzt wird es aber Zeit, sich an den Beweis von Satz 8.2 zu machen. Zuerst die Bestimmung der λ_j, die noch recht einfach ist: Im Geiste unserer rational/irrationalen Unterscheidungen sollen $\lambda_0, \ldots, \lambda_{2s}$ *irrationale* Zahlen sein, die über $\mathbb{Q}$ linear unabhängig sind, so daß es also keine rationalen Zahlen $q_0, \ldots, q_{2s}$ gibt, die

$$\sum_{j=0}^{2s} \lambda_j\, q_j = 0$$

erfüllen. Sowas kann man ganz fein und abstrakt mit algebraischen Köpererweiterungen machen, es geht aber auch einfacher.

Lemma 8.7 *Für jedes $n \in \mathbb{N}$ gibt es irrationale Zahlen $\lambda_1, \ldots, \lambda_n \in (0,1)$, die über $\mathbb{Q}$ linear unabhängig sind.*

Beweis: Wir wählen n *transzendente* Zahlen $\mu_1, \ldots, \mu_n \in (1,2)$, deren Produkt ebenfalls transzendent ist[209] und setzen $\lambda_j = \log_2 \mu_j \in (0,1)$, $j = 1, \ldots, n$. Wären diese λ_j nun linear abhängig über $\mathbb{Q}$, dann gäbe es q_j, $j = 1, \ldots, n$, so daß

$$0 = \sum_{j=1}^{n} q_j\, \lambda_j = \sum_{j=1}^{n} k_j\, \lambda_j, \qquad k_j \in \mathbb{Z},$$

wobei die k_j durch einfache Hauptnennerbildung generiert werden. Also ist

$$1 = 2^{\sum_{j=1}^n k_j\, \lambda_j} = \prod_{j=1}^{n} \left(2^{\log_2 \mu_j}\right)^{k_j} = 2^k \prod_{j=1}^{n} \mu_j^{k_j},$$

im deutlichen Widerspruch zur Transzendenz des Produkts der der μ_j. □

[209] Das geht, weil die rationalen und algebraischen Zahlen beide abzählbar sind – jede algebraische Zahl ist eine von endlich vielen Nullstellen eines Polynoms mit rationalen Koeffizienten und die sind abzählbar, wie man sich leicht überlegen kann: Man "sortiert" nach Grad und zählt dann die einzelnen Koeffizienten ab, was man mischt wir beim Cantorschen Diagonalverfahren für die rationalen Zahlen.

Jetzt kommen wir zum entscheidenden Lemma, das uns ein bisschen beschäftigen wird. Aber vorher noch ein klein wenig Notation. Wir fixieren immer noch[210] gemäß (8.7) $B = 4s + 2$ und betrachten die Intervalle

$$[a_{j,\beta}, b_{j,\beta}] = I(\beta) - 2j\, B^{-\ell(\beta)}, \qquad \beta \in \mathscr{B} = \bigcup_{k=0}^{\infty} \mathbb{Z}_B^{k-1}, \tag{8.9}$$

mit den *rationalen* Endpunkten $a_{j,\beta}, b_{j,\beta}$. Ausserdem setzen wir

$$\mathscr{B}_n = \bigcup_{k=0}^{n} \mathbb{Z}_B^{k-1}, \qquad n \in \mathbb{N}_0.$$

Lemma 8.8 *Es gibt* ***strikt*** *monoton steigende Funktionen $\phi_j : [0,1] \to [0,1]$, $j = 0, \ldots, 2s$, Lipschitz–stetig zu einem passenden $\alpha > 0$, so daß für jede feste Stufe n die Intervalle*

$$J_j\left(\beta^1, \ldots, \beta^s\right) := \left[\sum_{k=1}^{s} \lambda_k\, \phi_j\left(a_{j,\beta^k}\right), \sum_{k=1}^{s} \lambda_k\, \phi_j\left(b_{j,\beta^k}\right)\right], \qquad \begin{array}{l} \ell\left(\beta^1\right) = \cdots = \ell\left(\beta^s\right) = n, \\ j = 0, \ldots, 2s+1, \end{array} \tag{8.10}$$

disjunkt sind.

Beweis: Wir konstruieren die Funktionen ϕ_j simultan und induktiv nach n. Für $n = 0$ setzen wir $a_{j,()} = 0$, $b_{j,()} = 1$, und definieren die Funktionen an diesen Endpunkten als

$$\phi_j(0), \phi_j(1) \in \mathbb{Q}, \qquad 0 \le \phi_j(0) < \phi_j(1) \le 1, \qquad j = 0, \ldots, 2s,$$

so, daß für verschiedene j all diese Werte verschieden sind.

Im $n$ten Schritt legen wir dann

$$\phi_j\left(a_{j,\beta}\right) \text{ und } \phi_j\left(a_{j,\beta}\right), \qquad j = 0, \ldots, 2s, \quad \ell(\beta) = n,$$

fest, was, nach unserer Bemerkung über die Disjunktheit dieser Randpunkte, konsistent durchführbar ist. Darüberhinaus definieren wir die Erweiterung, so daß auch die folgenden Bedingungen erfüllt sind:

1. Jedes ϕ_j ist *strikt monoton steigend* auf den Punkten $a_{j,\beta}, b_{j,\beta}$, $\ell(\beta) \le n$.
2. Für jedes β mit $\ell(\beta) \le n$ und jedes $j = 0, \ldots, 2s$ hat die stückweise lineare Funktion f mit
$$f\left(a_{j,\beta}\right) = \phi_j\left(a_{j,\beta}\right), \qquad f\left(b_{j,\beta}\right) = \phi_j\left(b_{j,\beta}\right),$$
eine **strikt** kleinere Steigung als $(B/2)^{\ell(\beta)}$.
3. Alle Werte
$$\phi_j\left(a_{j,\beta}\right), \phi_j\left(a_{j,\beta}\right), \qquad j = 0, \ldots, 2s, \quad \beta \in \mathscr{B}_n,$$
sind verschieden.

[210] Das dürfte nicht wirklich notwendig sein, macht die Argumente aber einfacher

4. Die Intervalle $J\left(\beta^1,\ldots,\beta^s\right)$ aus (8.10) sind alle disjunkt.

Nehmen wir also an, wir hätten die Werte der Funktionen an allen Stellen $a_{j,\beta}, b_{j,\beta}$, $j=0,\ldots,2s$, $\beta\in\mathscr{B}_{n-1}$ passend bestimmt und sehen uns an, wie die Funktionen nun auf $a_{j,\beta}, b_{j,\beta}$, $j=0,\ldots,2s$, $\ell(\beta)=n$, fortzusetzen sind.

Zuerst setzen wir jedes ϕ_j individuell so fort, daß Eigenschaft 2) erfüllt ist. Auf Stufe n haben wir es bei der Zerlegung mit Intervallen der Form

$$\left[NB^{-n+1},(N+1)B^{-n+1}\right]\supset[a_{j,\beta},b_{j,\beta}]=I_j(\beta),$$

die zu einem Teil $(B-2)/B$ von $I_j(\beta)$ ausgefüllt werden, wobei zwischen zwei derartigen Intervallen Lücken der Länge $2/B$ auftreten. Daher kann so ein Intervall $I_j(\beta)$ entweder einen der früheren Punkte enthalten[211], oder keinen[212], siehe Abb. 8.5. Auf jedem Intervall der Stufe

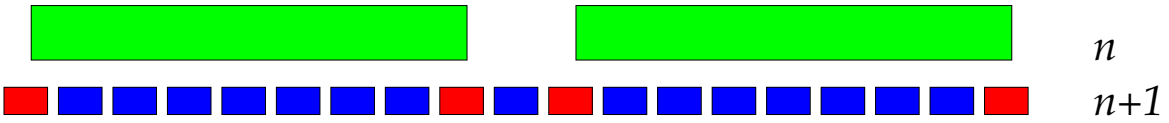


Abbildung 8.5: Die Intervalle auf Stufe n und $n+1$. Diejenigen Intervalle, die einen Randpunkt der darüberliegenden Stufe enthalten, sind rot eingefärbt.

n legen wir nun die Werte an den Endpunkten als

$$\phi_j\left(a_{j,\beta}\right)=\phi_j\left(b_{j,\beta}\right)=\phi_{j,\beta} \tag{8.11}$$

fest; das ist natürlich **nicht** streng monoton steigend, aber dieses Manko lässt später beheben. Enthält das Intervall $I_j(\beta)$ einen "übergeordeten" Endpunkt

$$a\in A_{j,n-1}:=\left\{a_{j,\beta}\ :\ \beta\in\mathscr{B}_{n-1}\right\}\cup\left\{b_{j,\beta}\ :\ \beta\in\mathscr{B}_{n-1}\right\},$$

dann setzen wir[213] $\phi_{j,\beta}=\phi_j(a)$ so daß unsere Definition von ϕ_j konsistent bleibt. Für die anderen Intervalle betrachten wir zwei beliebige benachbarte Punkte $a<a'\in A_{j,n-1}$, für die, nach Induktionsvoraussetzung, $\phi_j\left(a'\right)-\phi_j(a)\le(B/2)^{n-1}\left(a'-a\right)$ gilt, und die auf dem Gitter $\mathbb{N}_0B^{-n+1}$ liegen[214]. Daher verteilen sich die Intervalle und Lücken auf Stufe n gleichmäßig zwischen diesen Punkten, siehe Abb. 8.6, und die Lücken machen eine Gesamtlänge von $\frac{2(a'-a)}{B}$ aus. In jeder dieser Lücken erhöhen wir dann $\phi_{j,\beta}$ um einen Faktor, der proportional zu dieser Länge ist. Ist nun $x=b_{j,\beta}$ der rechte Randpunkt eines Intervalls und $x'=a_{j,\beta'}$ der linke

[211] Nämlich dann, wenn sich die Intervalle überlappen. Dieses Überlappen ist durch die Verschiebung bedingt.
[212] Nämlich dann, wenn das Intervall in eines niedrigerer Stufe "eingepasst" ist.
[213] Eine andere Wahl haben wir ja auch gar nicht.
[214] Hier wirkt sich die Kopplung von B und s in (8.7) positiv aus.

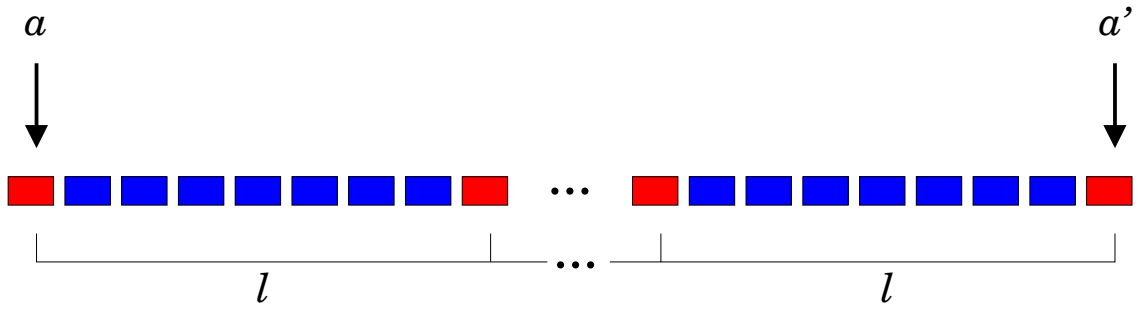


Abbildung 8.6: Die Punkte a, a' und die zugehörigen Intervalle auf Stufe n. Nachdem $a = k\,B^{-n+1}$ und $a' = k'\,B^{-n+1}$ ist, liegen zwischen den beiden endlich viele Stücke der Länge $l = B^{-n+1}$, die von den Intervallen auf Stufe n *gleichmäßig* überdeckt werden, so daß sich Intervalle und Lücken exakt gleich verteilen.

Randpunkt des rechen Nachbarn dazu, dann ist

$$\begin{aligned}\phi_j(x') - \phi_j(x) &= \frac{B}{2(a'-a)}(x'-x)(\phi_j(a') - \phi_j(a)) = \frac{B}{2}(x'-x)\underbrace{\frac{\phi_j(a') - \phi_j(a)}{(a'-a)}}_{<(B/2)^{n-1}} \\ &< \left(\frac{B}{2}\right)^n,\end{aligned}$$

und da die stückweise lineare Funktion auf den Intervallen ja erst einmal konstant ist, gilt 2) auch für n.

Von nun an betrachten wir die Funktionen ϕ_j nicht mehr separat. Da 2) unter hinreichend kleinen Störungen erhalten bleibt[215] und wir sonst noch nichts erreicht haben, können wir die neuen Werte so modifizieren, daß alle $\phi_{j,\beta}$ rational und voneinander verschieden sind, $j = 0, \ldots, 2s$, $\ell(\beta) = n$, ohne irgendwelche Abstriche machen zu müssen. Und die Funktionswerte, die wir aus früheren Stufen übernommen haben, haben diese Eigenschaft per Induktion. Nachdem die λ_k linear unabhängig über $\mathbb{Q}$ sind, sind dann auch die Werte

$$\sum_{k=1}^{s} \lambda_k\,\phi_{j,\beta^k}, \qquad j = 0, \ldots, 2s, \quad \ell\left(\beta^k\right) = n, \tag{8.12}$$

alle voneinander verschieden – identische Werte für verschiedene j oder verschiedene β würden ja zu einer linearen Abhängigkeit der λ_k führen.

Nun modifizieren die Werte von ϕ_j an den Intervallendpunkten $a_{j,\beta}, b_{j,\beta}$, die im Inneren $(0,1)$ des Intervalls $[0,1]$ liegen. Da die Werte in (8.12) alle verschieden sind, gibt es auch ein

[215] Die Freuden des "<".

$\varepsilon > 0$, so daß die Umgebungen

$$U_j\left(\beta^1,\ldots,\beta^k\right) := \sum_{k=1}^{s} \lambda_k\, \phi_{j,\beta^k} + \underbrace{\left(\sum_{k=1}^{s} \lambda_k\right)}_{=:\lambda} (-\varepsilon,\varepsilon)\,, \qquad j=0,\ldots,2s, \quad \ell\left(\beta^k\right) = n, \tag{8.13}$$

allesamt disjunkt sind, und wir wählen nun die festzulegenden Werte von ϕ_j jeweils so, daß

$$\phi_{j,\beta} - \varepsilon < \phi_j\left(a_{j,\beta}\right) < \phi_{j,\beta} < \phi_j\left(a_{j,\beta}\right) < \phi_{j,\beta} + \varepsilon, \qquad j=0,\ldots,2s, \quad \ell(\beta) = n,$$

was für hinreichend kleines ε auch nach wie vor ohne Verletzung von 2) möglich ist. Damit sind aber auch 1) und 3) erfüllt und für beliebige $\beta^1,\ldots,\beta^k$ mit $\ell\left(\beta^j\right) = n$ gilt außerdem

$$\sum_{k=1}^{s} \lambda_k\, \phi_j\left(a_{j,\beta^k}\right) > \sum_{k=1}^{s} \lambda_k\, \left(\phi_{j,\beta^k} - \varepsilon\right) > \sum_{k=1}^{s} \lambda_k\, \left(\phi_{j,\beta^k}\right) - \lambda\,\varepsilon \tag{8.14}$$

und ganz analog

$$\sum_{k=1}^{s} \lambda_k\, \phi_j\left(b_{j,\beta^k}\right) < \sum_{k=1}^{s} \lambda_k\, \left(\phi_{j,\beta^k}\right) + \lambda\,\varepsilon,$$

also insgesamt

$$\underbrace{\sum_{k=1}^{s} \lambda_k\, \left(\phi_{j,\beta^k}\right) - \lambda\,\varepsilon}_{\in U_j\left(\beta^1,\ldots,\beta^k\right)} < \sum_{k=1}^{s} \lambda_k\, \phi_j\left(a_{j,\beta^k}\right) < \sum_{k=1}^{s} \lambda_k\, \phi_j\left(b_{j,\beta^k}\right) < \underbrace{\sum_{k=1}^{s} \lambda_k\, \left(\phi_{j,\beta^k}\right) + \lambda\,\varepsilon}_{\in U_j\left(\beta^1,\ldots,\beta^k\right)}. \tag{8.15}$$

Mit anderen Worten,

$$\left[\sum_{k=1}^{s} \lambda_k\, \phi_j\left(a_{j,\beta^k}\right), \sum_{k=1}^{s} \lambda_k\, \phi_j\left(b_{j,\beta^k}\right)\right] \subset U_j\left(\beta^1,\ldots,\beta^k\right)$$

und da diese Intervalle alle disjunkt waren, gilt 4) bzw. (8.10).

Bleibt noch die Lipschitz–Stetigkeit und die ist natürlich der Grund, warum wir auf Eigenschaft 2) geachtet haben. Wir fixieren $j \in \{0,\ldots,2s\}$ und betrachten alle zu ϕ_j konstruierten Punkte, nämlich

$$A_j = \bigcup_{n\in\mathbb{N}_0} A_{j,n},$$

die eine dichte Teilmenge von $[0,1]$ bilden, von der aus wir ϕ_j natürlich wie vorher wieder passend fortsetzen können. Wählen wir nun $x \in A_j$ und[216] $0 < h < 2/B$ so, daß auch $x + h \in A_j$, dann gibt es $n > 0$, so daß $2\cdot B^{-n-1} \le h < 2\cdot B^{-n}$ ist[217] und das offene Intervall $(x, x+h)$ kann damit höchstens einen Punkt aus Stufe n, also einen der Punkte $a_{j,\beta}, b_{j,\beta}$, $\ell(\beta) = n$, enthalten. Es gibt drei Möglichkeiten:

[216] Für Lipschitzsstetigkeit ist ja nur der Fall kleiner h relevant, alles anderes lässt sich immer in die Konstante packen.

[217] Der größere der beiden Werte ist die Länge einer Lücke auf Stufe n.

- $a_{j,\beta} \in (x, x+h)$, dann ist $(x, x+h) \subset [b_{j,\beta'}, b_{j,\beta}]$, wobei β' der Index des linken Nachbarintervalls von $I_j(\beta)$ ist.

- $b_{j,\beta} \in (x, x+h)$, dann ist $(x, x+h) \subset [a_{j,\beta}, a_{j,\beta'}]$, nur steht eben jetzt β' für den Index des rechten Nachbarn.

- Keiner Punkte liegt in $(x, x+h)$, das dann entweder in einem Intervall oder einer Lücke komplett enthalten ist, also sowohl in einem Intervall der Form $[a_{j,\beta}, a_{j,\beta'}]$ als auch in einem der Form $[b_{j,\beta'}, b_{j,\beta}]$ liegt.

In allen drei Fällen gibt es also (mindestens) ein Intervall der Länge B^{-n+1} mit Randpunkten a, a' der Stufe n und wegen der strikten Monotonie[218] von ϕ_j und 2) ist daher

$$\begin{aligned} \phi_j(x+h) - \phi_j(x) &< \phi_j(a') - \phi_j(a) < \left(\frac{B}{2}\right)^n \underbrace{(a'-a)}_{=B^{-n+1}} < B2^{-n} = 2B\,2^{-n-1} \\ &= 2B\,B^{-\log_B 2\,(n+1)} = 2\left(2B^{-n-1}\right)^{\log_B 2} \leq 2h^{\log_B 2}, \end{aligned}$$

der in Lemma 8.8 angegebene Lipschitz–Exponent ist also ganz explizit $\alpha = \log_B 2$ und wird für wachsendes B, insbesondere also wegen (8.7) für wachsendes s, immer kleiner. □

Nachdem der Beweis durch die unverzichtbaren technischen Details etwas kompliziert geworden ist, fassen wir nochmal schnell die wesentliche Idee zusammen:

1. Zuerst definieren wir ϕ_j auf den Intervallen $I_j(\beta)$ einer festen Stufe n konstant mit rationalem Wert. Damit ist dann $\phi_j(I_j(\beta))$ einpunktig.

2. Wegen der rationalen linearen Unabhängigkeit der λ_k sind diese einpunktigen Mengen dann alle disjunkt.

3. Weil es nur endlich viele einpunktige Mengen sind, gibt es auch ε–Umgebungen dieser Mengen mit einem festen $\varepsilon > 0$ für alle, die ebenfalls immer noch disjunkt sind.

4. Wir verändern nun die Werte von ϕ_j an den Intervallendpunkten so, daß wir "richtige" Intervalle in diesen ε–Umgebungen erhalten, die immer noch disjunkt sind.

5. Den restlichen Aufwand betreibt man, um letztendlich Konsistenz zwischen den einzelnen Stufen, strikte Monotonie und Lipschitz–Stetigkeit zu bekommen.

So, den "schlimmsten" Teil des Beweises haben wir schon hinter uns gebracht. Jetzt setzen wir unsere Intervalle $I_j(\beta)$ zu s–dimensionalen Würfeln[219]

$$I_j\left(\beta^1, \ldots, \beta^s\right) := I_j\left(\beta^1\right) \times \cdots \times I_j\left(\beta^s\right) \subset \mathbb{R}^s, \qquad \ell\left(\beta^1\right) = \cdots = \ell\left(\beta^s\right),$$

[218] Nicht einschlafen, wir haben's gleich! Das kommt einem alles nur so monoton vor.
[219] Da wir alle Indizes β^k von gleicher Länge wählen, sind das auch **wirklich** Würfel.

zusammen, die durch die Funktionen

$$\theta_j\left(x_1,\ldots,x_s\right) := \sum_{k=1}^{s} \lambda_k\, \phi_j\left(x_k\right)$$

auf die entsprechenden J_j abgebildet werden:

$$\theta_j\left(I_j\left(\beta^1,\ldots,\beta^s\right)\right) = J_j\left(\beta^1,\ldots,\beta^s\right), \qquad j=0,\ldots,2s, \quad \ell\left(\beta^1\right)=\cdots=\ell\left(\beta^s\right). \quad (8.16)$$

Nach Lemma 8.8 werden damit diese disjunkten Würfel auf disjunkte Intervalle abgebildet.

Alles, was wir bisher gemacht haben, betraf eigentlich immer die ϕ_j separat. Warum aber brauchen wir eigentlich so viele davon? Nun, wenn wir uns die Intervalle $I_j(\beta)$, $\ell(\beta)=n$ aus (8.6), dann überdecken diese für festes j und n nicht das gesamte Intervall $[0,1]$, zusammen hingegen schon:

$$[0,1] = \bigcup_{\ell(\beta)=n} \bigcup_{j=0}^{2s} I_j(\beta).$$

Lemma 8.9 *Jeder Punkt $x\in[0,1]$ ist für jedes $n\in\mathbb{N}_0$ in mindestens $2s$ der Intervalle $I_j(\beta)$, $j=0,\ldots,s$, $\ell(\beta)=n$, enthalten.*

Beweis: Sei $x\in(0,1)$, dann gibt es $0\le N<B^{n-1}$ so daß $NB^{-n+1}<x\le(N+1)\,B^{-n+1}$. Dieses x liegt in einem Intervall der Form

$$NB^{-n+1} + B^{-n}\left([1,B-1]-2j\right)$$

wenn

$$x - NB^{-n+1} =: y \in [1,B-1]-2j,$$

also $j\in y/2-\left[\frac{1}{2},\frac{B-1}{2}\right]$. Dieses Intervall hat die Länge $\frac{B-2}{2}$ und enthält deswegen auch $\frac{B-2}{2}=2s$ ganzzahlige Werte, die einer passenden Verschiebung entsprechen[220]. Damit ist x in mindestens $2s$ dieser Intervalle enthalten und nachde die Intervalle für ein festes j disjunkt sind, braucht man auch wirklich diese Variation über j. □

Lemma 8.10 *Jeder Punkt $x\in[0,1]^s$ ist in mindestens $s+1$ der Würfel*

$$I_j\left(\beta^1,\ldots,\beta^s\right), \qquad j=0,\ldots,2s, \quad \ell\left(\beta^k\right)=n,$$

enthalten.

Beweis: Wiederholte Anwendung von Lemma 8.9. Es gibt höchstens einen Index j, so daß x_1 nicht von den Intervallen $I_j\left(\beta^1\right)$, $\ell\left(\beta^1\right)=n$, getroffen wird, unter den verbleibenden $2s$ kann wieder höchstens ein Index nicht bei x_2 erfolgreich sein und so weiter. Nach s Schritten bleiben so mindestens $s+1$ Indizes übrig. □

Nach diesen vorbereitenden Zählbemerkungen nun zum Abschluss des Beweises von Satz 8.2, nämlich zur Konstruktion einer "nicht schlecht" approximierenden nomographischen Funktion.

[220] **Achtung:** Genau hier wird die Kopplung zwischen s und B aus (8.7) wichtig!

Lemma 8.11 *Sei* $\frac{s}{s+1} < \mu < 1$. *Zu jeder Funktion* $f \in C\left([0,1]^s\right)$ *gibt es eine Funktion* $g \in C\left[0,\lambda\right]$, *so daß*

$$\left\| f - \sum_{j=0}^{2s} g \circ \theta_j \right\| \leq \mu \left\| f \right\|, \qquad \|g\| \leq \frac{1}{s+1}\|f\|. \tag{8.17}$$

Beweis: Wir wählen $\varepsilon \in \left(0, \mu - \frac{s}{s+1}\right)$ und dann n so groß, daß

$$\max_{j=0,\ldots,s} \max_{\ell(\beta^1)=\cdots=\ell(\beta^s)=n} \max_{x,x' \in I_j\left(\beta^1,\ldots,\beta^k\right)} |f(x) - f\left(x'\right)| < \varepsilon \|f\|$$

ist, was wegen der gleichmäßigen Stetigkeit von f und der gleichmäßig immer kleiner werdenden Größe der Würfelchen immer möglich ist. Bezeichnen wir mit $f_j\left(\beta^1,\ldots,\beta^s\right)$ den Wert am Mittelpunkt von $I_j\left(\beta^1,\ldots,\beta^s\right)$, dann ist

$$\max_{x \in I_j\left(\beta^1,\ldots,\beta^k\right)} \left| f(x) - f_j\left(\beta^1,\ldots,\beta^s\right)\right| < \varepsilon \|f\|.$$

Nun definieren wir

$$g(t) = \frac{1}{s+1} f_j\left(\beta^1,\ldots,\beta^s\right), \qquad t \in J_j\left(\beta^1,\ldots,\beta^s\right) = \theta\left(I_j\left(\beta^1,\ldots,\beta^s\right)\right),$$

was nach Lemma 8.8 widerspruchsfrei möglich ist und setzen g außerhalb dieser Intervalle linear fort, so daß g auf alle Fälle stetig ist und die zweite Bedingung in (8.16) erfüllt.

Sei nun $x = (x_1,\ldots,x_s) \in [0,1]^s$, dann gibt es nach Lemma 8.10 mindestens $s+1$ Indizes, sagen wir[221] $j = 0,\ldots,s$, so daß $x \in I_j\left(\beta^1,\ldots,\beta^s\right)$ mit möglicherweise von j abhängigen Indizes[222] und so ist

$$\begin{aligned}
\left| f(x) - \sum_{j=0}^{2s} g\left(\theta_j(x)\right)\right| &\leq \left| f(x) - \sum_{j=0}^{s} g\left(\theta_j(x)\right)\right| + \sum_{j=s+1}^{2s} \underbrace{\left|g\left(\theta_j(x)\right)\right|}_{\leq \|g\| \leq \|f\|/(s+1)} \\
&\leq \frac{s}{s+1}\|f\| + \left| f(x) - \sum_{j=0}^{s} \frac{f_j\left(\beta^1,\ldots,\beta^s\right)}{s+1}\right| \\
&\leq \frac{s}{s+1}\|f\| + \sum_{j=0}^{s} \frac{1}{s+1} \underbrace{\left| f(x) - f_j\left(\beta^1,\ldots,\beta^s\right)\right|}_{\leq \varepsilon\|f\|} \\
&\leq \underbrace{\left(\frac{s}{s+1} + \varepsilon\right)}_{<\mu} \|f\| < \mu \left\|f\right\|,
\end{aligned}$$

wie in (8.16) behauptet. □

[221] Die Reihenfolge, in der die ϕ_j indiziert werden, ist ja schließlich irrelvant.

[222] Aber berechtigt das die Einführung eines weiteren Index, also sowas wie $\beta^{j,1},\ldots,\beta^{j,s}$? Ich denke nicht. es reicht, darauf hinzuweisen.

Bemerkung 8.12 *Dieses Argument zeigt auch, daß die Anzahl der Funktionen $\phi_0, \ldots, \phi_{2s}$ im Satz 8.2 nicht ganz zufällig war. Damit das obige Argument mit m Funktionen, das heisst mit m Verschiebungen von Intervallen funktioniert*[223] *müssen wir für die m Verschiebungen und $m-s$ Überdeckungen von Würfeln die Beziehung $\frac{s}{m-s} < 1$ erhalten, denn sonst funktioniert die letzte Abschätzung von oben nicht. Das liefert aber $m - s < s$ oder eben $m > 2s$, also $m \geq 2s+1$.*

Und jetzt haben wir es so gut wie geschafft.
Beweis von Satz 8.2: Wir bestimmen uns die Darstellung durch iterierte Approximation vermittels Lemma 8.11. Sei g_1 die in Lemma (8.17) gefundene Approximation und wenden wir das Lemma erneut an, diesmal auf

$$r_1 = f - \sum_{j=0}^{2s} g_1 \circ \theta_j,$$

und erhalten g_2, so daß

$$\left\| f - \sum_{k=1}^{2} \sum_{j=0}^{2s} g_1 \circ \theta_j \right\| = \left\| r_1 - \sum_{j=0}^{2s} g_2 \circ \theta_j \right\| \leq \mu \left\| r_1 \right\| = \mu \left\| f - \sum_{j=0}^{2s} g_1 \circ \theta_j \right\| \leq \mu^2 \|f\|$$

sowie

$$\|g_2\| \leq \frac{1}{s+1} \|r_1\| \leq \frac{\mu}{s+1} \|f\|.$$

Generell setzen wir

$$r_n = f - \sum_{k=1}^{n} \sum_{j=0}^{2s} g_1 \circ \theta_j = f - \sum_{j=0}^{2s} \left(\sum_{k=1}^{n} g_k \right) \circ \theta_j, \qquad n \in \mathbb{N},$$

erhalten damit nach Lemma 8.11 ein g_{n+1} mit

$$\left\| f - \sum_{j=0}^{2s} \left(\sum_{k=1}^{n} g_k \right) \circ \theta_j \right\| = \left\| r_n - \sum_{j=0}^{2s} g_2 \circ \theta_j \right\| \leq \mu^{n+1} \left\| f \right\|, \qquad \|g_{n+1}\| \leq \frac{\mu^n}{s+1} \|f\|$$

und damit konvergiert die Reihe der g_k gleichmäßig gegen eine Grenzfunktion mit der Eigenschaft

$$f = \sum_{j=0}^{2s} g \circ \theta_j = \sum_{j=0}^{2s} g \left(\sum_{k=1}^{s} \lambda_k \, \phi_j \left(x_k \right) \right)$$

was nun endlich (8.2) ist. □

Bemerkung 8.13 *Auch wenn der Beweis des Satzes von Kolmogoroff eigentlich konstruktiv ist, gäbe es bei einer wirklichen Realisierung auf dem Computer doch noch einige Probleme zu lösen:*

[223] Natürlich ist dies nur eine **technische** Einschränkung, die dafür sorgt, daß **dieser** Beweis funktioniert und es schließt erst einmal nicht aus, daß es einen ganz anderen Beweis für das Resultat geben könnte, der mit weniger Funktionen auskommt. Bekannt ist allerdings bis heute meines Wissens keiner!

1. *Bei der Konstruktion der ϕ_j müssen "kleine Änderungen" an den Werten der Funktionen vorgenommen werden. In der Praxis heisst dies, daß derartige Änderungen klein genug sein müssten, um schwierige zu kontrollierenden Anforderungen zu genügen, gleichzeitig aber dann doch groß genug, um vom Rechner überhaupt wahrgenommmen zu werden.*

2. *Die numerische Realisierung von reellen Zahlen, die über $\mathbb{Q}$ linear unabhängig sind, ist eigentlich unmöglich, schließlich zahlen sich derartige Zahlen ja einerseits dadurch aus, daß sie bezüglich jeder Basis eine unendliche und nichtperiodische Entwicklung haben, andererseits sind sie aber ohnehin eher schlecht durch rationale Zahlen zu approximieren, siehe [36, 72].*

3. *Die Funktion g liegt nicht wirklich exakt vor, sondern ist Grenzwert einer Funktionenreihe. Das ist aber für praktische Zwecke die geringste Einschränkung, denn mehr als eine Approximation von f durch nomographische Funktionen brauchen wir uns sowieso nicht einzubilden und die Partialsummen sind ja gute Approximationen, schließlich ist der Fehler von der Form*

$$\left\| g - \sum_{k=1}^{n} g_j \right\| \leq \sum_{k=n+1}^{\infty} \|g_k\| \leq \|f\| \sum_{k=n+1}^{\infty} \frac{\mu^{k-1}}{s+1} = \frac{\mu^n}{(s+1)(1-\mu)} \|f\|,$$

was sogar exponentiell gegen Null geht – lineare Konvergenzordnung *nennt man sowas, siehe [69].*

8.4 Neuronale Netze

Der Satz von Kolmogoroff hat aufgrund seiner mathematischen Natur allein schon durchaus das Zeug zum Klassiker aber es gibt darüber hinaus auch noch eine Interpretation im Sinne moderner numerischer Anwendungen, nämlich der *neuronalen Netze*. Die Idee des neuronalen Netzes ist es, die Arbeitsweise der Gehirns zu simulieren und Funktionen aus einfachen Elementen, eben *Neuronen* zusammenzusetzen. Ein derartiges Neuron ist eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ der Form

$$f(x) = \varphi_{w,w_0}(x) := \varphi\left(w^T x + w_0\right), \qquad w \in \mathbb{R}^n,\, w_0 \in \mathbb{R}, \tag{8.18}$$

wobei φ die *Anregungsfunktion* des Neurons ist. Eine Funktion wie in (8.18) bezeichnet man auch als *Ridge–Funktion* und solche Funktionen sind sehr einfach, denn sie sind entlang der affinenen Hyperebenen

$$y + w^{\perp}, \qquad w^{\perp} = \left\{ x \in \mathbb{R}^n \;:\; w^T x = 0 \right\}$$

konstant: Für $x^{\perp} \in w^{\perp}$ ist

$$f\left(y + x^{\perp}\right) = \varphi\left(w^T\left(y + x^{\perp}\right) + w_0\right) = \varphi\left(w^T y + w^T x^{\perp} + w_0\right) = \varphi\left(w^T y + w_0\right) = f(y).$$

Das Konzept des Neurons ist nun, daß die Aktivierungsfunktion als fest angesehen wird und man nur die Parameter w_0 und w variiert. Um die Notation zu vereinfachen setzen wir nun

$w = (w_0, \ldots, w_n)$ und[224] $\widehat{x} = (1, x)$, denn dann wird (8.18) zu

$$f(x) = \varphi_w(x) = \varphi\left(w^T \widehat{x}\right), \qquad x \in \mathbb{R}^n. \tag{8.19}$$

Beispiel 8.14 (Aktivierungsfunktionen) *Typische Aktivierungsfunktionen für neuronale Netze sind*

1. $\varphi(x) = \operatorname{sgn} x$*, also*

$$\varphi_w(x) = 1 \quad \Leftrightarrow \quad 0 < w^T \widehat{x} = w_0 + \sum_{j=1}^{n} w_j\, x_j \quad \Leftrightarrow \quad \sum_{j=1}^{n} w_j\, x_j > -w_0,$$

   das Neuron feuert *also, wenn das innere Produkt größer als eine* Aktivierungsschwelle*, auf englisch* Threshold .

2. $\varphi : \mathbb{R} \to [0, 1]$ *strikt monoton steigend, also sozusagen eine kontinuierliche Aktivierung im Gegensatz zur Sprunfunktion* $x \mapsto \operatorname{sgn} x$*.*

3. *Ein konkrete Aktivierungsfunktion von dieser Art ist die* Sigmoidfunktion

$$\sigma(x) = \frac{1}{1 + e^{-x}}, \tag{8.20}$$

   siehe [3] bzw Abb. 8.7.

Ein neuronales Netzwerk besteht nun aus einer beliebigen Anzahl von *Lagen*, auf Englisch als *Layer* bezeichnet, die man erst einmal als Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ bezeichnen könnte. Soweit ist das nicht spektakulär, aber jeder dieser Layer soll eine *einfache Struktur* haben und nur auf einem Neuron beruhen. Und zwar werden alle Eingabedaten *gewichtet* in Kopien *desselben Neurons* geführt, siehe Abb. und die Resultate als Ausgabewerte genommen. Mathematisch ist so ein Layer also als

$$\ell(x) = \left[\varphi_{w(j)}(x) \;:\; j = 1, \ldots, m\right] = \begin{bmatrix} \varphi\left(w(1)^T \widehat{x}\right) \\ \vdots \\ \varphi\left(w(m)^T \widehat{x}\right) \end{bmatrix}, \qquad w(j) \in \mathbb{R}^{n+1}, \quad j = 1, \ldots, m, \tag{8.21}$$

dargestellt. Es gibt zwei spezielle Typen von Layern, nämlich

Eingabelayer, bei denen $m = n$ und $w(j)_k = \delta_{jk}$ ist, wo also alle Eingabekanäle einmal durchs Neuron gejagt werden und

Ausgabelayer, bei denen $m = 1$ ist, also einfach die Eingangskanäle gewichtet aufsummiert und dann durch das Neuron "gefiltert" werden.

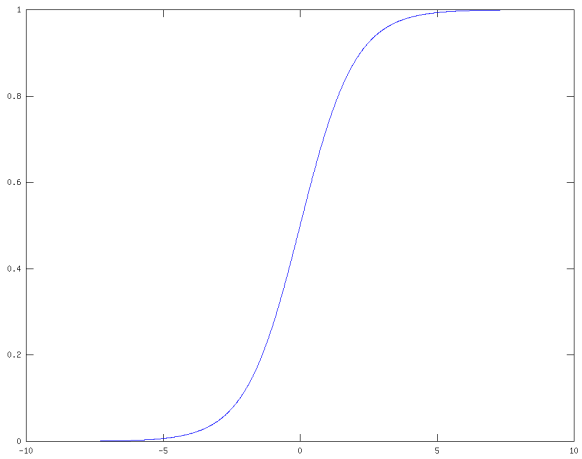


Abbildung 8.7: Die Sigmoidfunktion aus (8.20). Durch Umskalierung $\sigma(\lambda\cdot)$, $\lambda > 0$, kann man den Steigungsteil natürlich beliebig steil oder flach machen. Obwohl die Funktion nie exakt die Werte 0 und 1 annimmt, erreicht sie diese aber "praktisch" schon relativ bald.

Wie der Name sagt, stehen Eingabelayer immer am Anfang, Ausgabelayer immer am Ende des Prozesses. Derartige Layer können nun beliebig baumartig ineinander verschachtelt werden und führen dann eben zu einem *neuronalen Netzwerk*[225].

Die "moderne" numerische Bedeutung des Satzes von Kolmogoroff liegt nun darin, daß man ihn als einen Darstellungssatz für neuronale Netzwerke interpretieren kann, wobei für $j = 0, \ldots, 2s$ zuerst je einen Eingabelayer basierend auf ϕ_j verwendet[226], diese Ausgaben dann durch jeweils einen einfachen Layer mit Anregungsfunktion g, Gewichten $\lambda_1, \ldots, \lambda_s$ und nur einem Ausgabkanal schickt und die Resultate durch ein triviales Ausgabenetzwerk kombinieren lässt. Trivial bedeutet in diesem Fall, daß alle Gewichte den Wert 1 und die Anregungsfunktion die Identität ist. Toll – neuronale Netze können also jede Funktion darstellen, **aber** (zumindest, wenn man Satz 8.2 verwenden will) in diesem Fall muss eine der Anregungsfunktionen sehr massiv von der darzustellenden Funktion f abhängen und müsste dann selbst wieder durch ein geeignetes Teilnetzwerk approximiert werden.

Bleibt noch eine Frage zum Schluss: Wie erstellt man eigentlich generell so ein Netzwerk in einer Anwendung? Das ist erstaunlich einfach! Man verwendet sogenannte *Trainingsdaten* $(x^j, y_j) \in \mathbb{R}^{s+1}$, belegt die freien Parameter $\boldsymbol{w}$, also die Gewichte in den einzelnen Layern,

[224] Wer will kann dies als eine Einführung von projektiven Koordinaten sehen, auch wenn wir hier beim besten Willen keine projektive Geometrie betreiben.

[225] Enthält dieses Netzwerk als Graph keine geschlossenen Kreise, dann spricht man von einem *Feedworward–*Netzwerk, mit Kreisen kann es noch lustiger werden, denn dann kann sich das Netzwerk ja auch rekursiv selbst anregen.

[226] Die ϕ_j sind monoton steigende Funktionen von $[0, 1]$ nach $[0, 1]$, wenn man die noch für $x < 0$ zu 0 und für $x > 1$ zu 1 fortsetzt, dann hat man es durchaus mit einer Art Sigmoidfunktion zu tun.

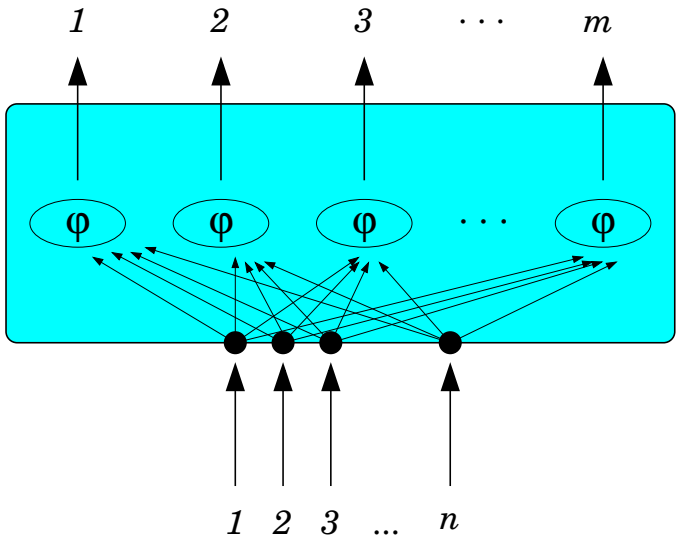


Abbildung 8.8: Ein "Layer" eines neuronalen Netzwerkes: Die n Eingangskanäle werden, jeweils mit zu wählenden Gewichten, in alle Neuronen geführt. Diese Neuronen sind alle Kopien voneinander, haben also dieselbe Anregungsfunktion

mit meist zufälligen Werten vor und minimiert dann die Abweichung des Netwerkes von den vorgegebenen Daten:

$$\min_{\boldsymbol{w}} \sum_{j=1}^{N} \left(f_{\boldsymbol{w}}\left(x_j\right) - y_j\right)^2 .$$

Das ist ein nichtlineares Optimierungsproblem und für sowas gibt es Methoden, normalerweise sogenannte Abstiegsverfahren, siehe beispielsweise [58, 70]. Was ein paar ganz interessante Bemerkungen hervorruft:

1. Solche nichtlinearen Optimierungsverfahren finden normalerweise nur *lokale* Minima und es kann somit nicht garantiert werden, daß das Netzwerk die Parameter wirklich optimal einstellt.

2. Durch die zufällige Vorbelegung kann es passieren, daß bei verschiedenen "Trainingssitzungen" dieselben Eingaben zu verschiedenen Resultaten führen.

3. Generell haben neuronale Netzwerke relativ wenige wirklich beweisbare Eigenschaften und man kann nie wirklich garantieren, daß das Netzwerk für alle Eingabewerte gesicherte Ergebnisse liefert.

Man kann noch vieles über neuronale Netze erzählen, aber das alles wäre wieder eine ganz andere Geschichte für sich.

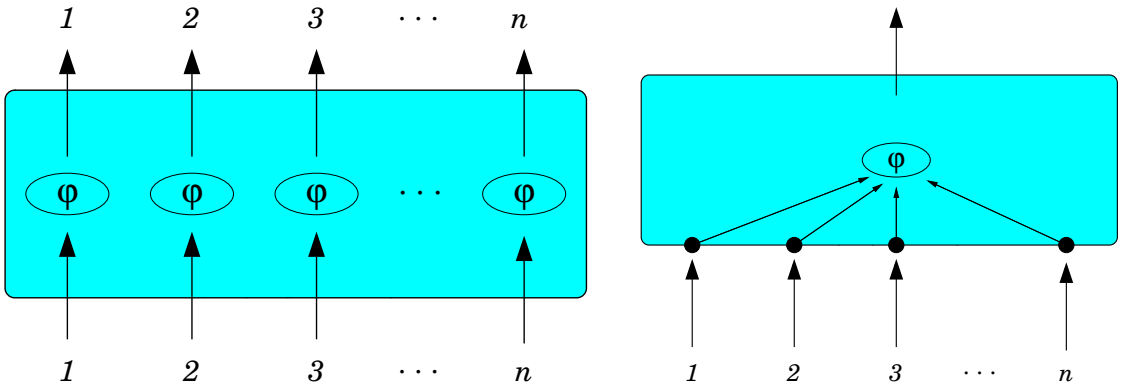


Abbildung 8.9: Ein Eingabe- *(links)* und ein Ausgabelayer *(rechts)*. Man sieht in beiden Fällen die besonders einfache Struktur.

Uns ist in alten mæren
wunders viel geseit
von Helden lobebæren
von grôzer arebeit

Das Nibelungenlied

8 Literatur

[1] A. G. Aitken. On interpolation by iteration of proportional parts, without the use of differences. *Proc. Edinburgh Math. Soc.*, **3** (1932), 56–76.

[2] V. G. Amel'kovič. A theorem converse to a theorem of Voronovskaya type. *Teor. Funkčiĭ, Funkcional Anal. i Prilozen, Vyp*, **2** (1966), 67–74.

[3] M. Anthony. *Discrete Mathematics of Neural Networks. Selected Topics*. Monographs on Discrete Mathematics and Applications. SIAM, 2001.

[4] V. I. Arnol'd. The representation of functions of several variables. *Mat. Prosvešč*, **3** (1958), 41–61.

[5] B. Bajšanski and R. Bojanić. A note on approximation by Bernstein polynomials. *Bull. Amer. Math. Soc.*, **70** (1964), 675–677.

[6] H. Berens and R. DeVore. A characterization of Bernstein polynomials. In E. W. Cheney, editor, *Approximation Theory III, Proc. Conf. Hon. G. G. Lorentz, Austin/Tex.*, pages 213–219. Academic Press, 1980.

[7] S. Bernstein. Sur l'ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné. *Memoires couronnés de l'Academie de Belgique*, (1912), 78–85.

[8] S. N. Bernstein. Démonstration du théorème de Weierstrass, fondée su le calcul des probabilitiés. *Commun. Soc. Math. Kharkov*, **13** (1912), 1–2.

[9] E. Bishop. A generalization of the Stone–Weierstrass theorem. *Pacific J. Math.*, **11** (1961), 777–783.

[10] B. Brosowski and F. Deutsch. An elementary proof of the Stone–Weierstrass theorem. *Proc. Amer. Math. Soc.*, **81** (1981), 89–92.

[11] P. L. Butzer and H. Berens. *Semi–Groups of Operators and Approximation*. Grundlehren der mathematischen Wissenschaften. Springer–Verlag, 1967.

[12] G. Z. Chang and J. Z. Zhang. Converse theorems of convexity for Bernstein polynomials over triangles. *J. Approx. Theory*, **61** (1990), 265–278.

[13] W. Dahmen. Convexity and Bernstein–Bézier polynomials. In A. Le Méhauté and L. L. Schumaker, editors, *Curves and Surfaces*, pages 107–134. Academic Press, 1991.

[14] W. Dahmen and C. A. Micchelli. Convexity and Bernstein polynomials on k-simploids. *Acta Math. Appl. Sinica*, **6** (1990), 50–66.

[15] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Commun. on Pure and Appl. Math.*, **41** (1988), 909–996.

[16] I. Daubechies. *Ten Lectures on Wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, 1992.

[17] P. J. Davis. *Interpolation and Approximation*. Dover Books on Advanced Mathematics. Dover Publications, 1975.

[18] R. A. DeVore and G. G. Lorentz. *Constructive Approximation*, volume 303 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1993.

[19] A. Dinghas. Über einige Identitäten vom Bernsteinschen Typus. *Det Koneglige Norske Videnskabers Selskab*, **24** (1951).

[20] E. Doblhofer. *Zeichen und Wunder. Geschichte und Entzifferung verschollener Schriften und Sprachen*. Paul Neff Verlag, Wien. Lizenzausgabe Weltbild Verlag, 1990.

[21] L. Fejér. Untersuchungen über Fouriersche Reihen. *Math. Annalen*, **58** (1904), 51–60.

[22] L. Fejér. Beispiele stetiger Funktionen mit divergenter Fourierreihe. *Journal Reine Angew. Math*, **137** (1910), 1–5.

[23] O. Forster. *Analysis 3. Integralrechung im $\mathbb{R}^n$ mit Anwendungen*. Vieweg, 3. edition, 1984.

[24] C. F. Gauss. Methodus nova integralium valores per approximationem inveniendi. *Commentationes societate regiae scientiarum Gottingensis recentiores*, **III** (1816).

[25] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Grundlehren der mathematischen Wissenschaften. Springer–Verlag, 2. edition, 1983.

[26] D. Ch. von Grüningen. *Digitale Signalverarbeitung*. VDE Verlag, AT Verlag, 1993.

[27] A. Haar. Die Minkowskische Geometrie und die Annäherung stetiger funktionen. *Math. Ann.*, **78** (1918), 294–311.

[28] G. H. Hardy and W. W. Rogosinsky. *Fourier series*. Cambridge University Press, 1950.

[29] H. Heuser. *Lehrbuch der Analysis. Teil 2*. B. G. Teubner, 2. edition, 1983.

[30] H. Heuser. *Lehrbuch der Analysis. Teil 1*. B. G. Teubner, 3. edition, 1984.

[31] N. J. Higham. *Accuracy and stability of numerical algorithms*. SIAM, 2nd edition, 2002.

[32] D. Jackson. On the approximation by trigonometric sums and polynomials. *Trans. Amer. Math. Soc.*, **12** (1912), 491–515.

[33] D. Jackson. *The Theory of Approximation*, volume XI of *AMS Colloquium Publications*. Amer. Math. Soc., 1930.

[34] R. Q. Jia and J. Lei. Approximation by multiinteger translates of functions having global support. *J. Approx. Theor.*, **72** (1993), 2–23.

[35] Y. Katznelson. *An Introduction to Harmonic Analysis*. Dover Books on advanced Mathematics. Dover Publications, 2. edition, 1976.

[36] A. Ya. Khinchin. *Continued fractions*. University of Chicago Press, 3rd edition, 1964. Reprinted by Dover 1997.

[37] A. N. Kolmogoroff. A remark on the polynomials fo Chebyshev, deviating at least from a given function. *Ushepi*, **3** (1948), 216–221. Probably in Russian.

[38] A. N. Kolmogoroff. On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition. *Dokl. Akad. Nauk. SSSR*, **114** (1957), 369–373.

[39] N. P. Korneĭčuk. The best uniform approximation of differentiable functions. *Dokl. Akad. Nauk. SSSR*, **141** (1961), 304–307. Probably in Russian.

[40] N. P. Korneĭčuk. The exact constant in the theorem of D. Jackson on the best uniform approximation of continuous periodic functions. *Dokl. Akad. Nauk. SSSR*, **145** (1962), 514–515. Probably in Russian.

[41] N. P. Korneĭčuk. On the existence of a linear polynomial operator which gives best approximation on a class of functions. *Dokl. Akad. Nauk. SSSR*, **143** (1962), 25–27. Probably in Russian.

[42] N. P. Korneĭčuk. The best approximation of continuous functions. *Izv. Akad. Nauk. SSSR*, **27** (1963), 29–44. Probably in Russian.

[43] E. Kreyszig. *Introductionary Functional Analysis with Applications*. John Wiley & Sons, 1978.

[44] H. Kuhn. Ein elementarer Beweis des Weierstraßschen Approximationssatzes. *Arch. Math.*, **15** (1964), 316–317.

[45] G. G. Lorentz. *Bernstein Polynomials*. University of Toronto Press, 1953.

[46] G. G. Lorentz. Metric entropy, widths, and superpositions of functions. *Amer. Math. Monthly*, **69** (1962), 469–485.

[47] G. G. Lorentz. *Approximation of functions*. Chelsea Publishing Company, 1966.

[48] G. G. Lorentz, M. v. Golitschek, and Y. Makovoz. *Constructive Approximation. Advanced Problems*, volume 304 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1996.

[49] A. K. Louis, P. Maaß, and A. Rieder. *Wavelets*. B. G. Teubner, 2. edition, 1998.

[50] S. Machado. On Bishop's generalization of the Stone–Weierstrass theorem. *Indag. Math.*, **39** (1977), 218–224.

[51] MacTutor. The MacTutor History of Mathematics archive. `http://www-groups.dcs.st-and.ac.uk/~history`, 2003. University of St. Andrews.

[52] J. C. Mairhuber. On Haar's theorem concerning Chebychev approximation problems having unique solutions. *Proc. Am. Math. Soc.*, **7** (1956), 609–615.

[53] J. C. Mairhuber, I. J. Schoenberg, and R. E. Williamson. On variation diminishing transformations on the circle. *Rend. Circ. Mat. Palermo, II. Ser.*, **8** (1959), 241–270.

[54] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 2. edition, 1999.

[55] H. M. Mhaskar and D. V. Pai. *Fundamentals of Approximation Theory*. Narosa Publishing House, 2000.

[56] C. A. Micchelli. *Mathematical Aspects of Geometric Modeling*, volume 65 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, 1995.

[57] Ch. H. Müntz. Über den Approximationssatz von Weierstrass. In *Schwarz–Festschrift*, pages 302–312. ???, Berlin, 1914.

[58] J. Nocedal and S. J Wright. *Numerical Optimization*. Springer Series in Operations Research. Springer, 1999.

[59] Chr. Rabut. On Pierre Bézier's life and motivations. Technical report, INSA Toulouse, 2001.

[60] H. Rademacher and I. J. Schoenberg. Helly's theorems on convex domains and Tchbycheff's approximation problem. *Canad. J. Math.*, **2** (1950), 245–256.

[61] T. Radó. *Subharmonic Functions*. Ergebnisse der Mathematik und ihrer Grenzgebiete. Julius Springer, Berlin, 1937.

[62] T. J. Ransford. A short elementary proof of the Bishop–Stone–Weierstrass theorem. *Math. Proc. Camb. Phil. Soc.*, **96** (1984), 309–311.

[63] T. J. Rivlin and H. S. Shapiro. A unified approach to certain problems of approximation and minimization. *J. Soc. Indust. and Appl. Math.*, **9** (1961), 670–699.

[64] T. Sauer. Multivariate Bernstein polynomials and convexity. *Comp. Aided Geom. Design*, **8** (1991), 465–478.

[65] T. Sauer. Note: On a maximum principle for Bernstein polynomials. *J. Approx. Theory*, **71** (1992), 121–122.

[66] T. Sauer. Axial convexity – a well shaped shape property. In P.-J. Laurent, A. Le Méhauté, and L. L. Schumaker, editors, *Curves and Surfaces in Geometric Design*, pages 419–425. A K Peters, 1994.

[67] T. Sauer. Axially parallel subsimplices and convexity. *Comp. Aided Geom. Design*, **12** (1995), 491–505.

[68] T. Sauer. Multivariate Bernstein polynomials, convexity and related shape properties. In J. M. Pena, editor, *Shape preserving representations in Computer Aided Design*. Nova Science Publishers, 1999.

[69] T. Sauer. Numerische Mathematik I. Vorlesungsskript, Friedrich–Alexander–Universität Erlangen–Nürnberg, Justus–Liebig–Universität Gießen, 2000. `http://www.math.uni-giessen.de/tomas.sauer`.

[70] T. Sauer. Optimierung. Vorlesungsskript, Justus–Liebig–Universität Gießen, 2002. `http://www.math.uni-giessen.de/tomas.sauer`.

[71] T. Sauer. Digitale Signalverarbeitung. Vorlesungsskript, Justus–Liebig–Universität Gießen, 2003. `http://www.math.uni-giessen.de/tomas.sauer`.

[72] T. Sauer. Kettenbrüche und Approximation. Vorlesungsskript, Justus–Liebig–Universität Gießen, 2005. `http://www.math.uni-giessen.de/tomas.sauer`.

[73] H. J. Schmid. Bernsteinpolynome. Manuscript, 1975.

[74] I. J. Schoenberg. Contributions to the problem of approximation of equidistant data by analytic functions. part B. – on the second problem of osculatory interpolation. a second class of analytic approximation formulae. *Quart. Appl. Math.*, **4** (1949), 112–141.

[75] L. G. Šnirel'man. On uniform approximations. *Izv. Akad. Nauk. SSSR, Ser. Mat.*, **2** (1938), 53–59. In Russian, French Abstract, pp. 59–60.

[76] D. A. Sprecher. On the structure of continuous functions of several variables. *Trans. Amer. Math. Soc.*, **115** (1965), 340–355.

[77] D. A. Sprecher. A representation theorem for continuous functions of several variables. *Proc. Amer. Math. Soc.*, **16** (1965), 200–203.

[78] D. A. Sprecher. An improvement in the superposition theorem of kolmogorov. *J. Math. Anal. Appl.*, **38** (1972), 208–213.

[79] M. H. Stone. The generalized Weierstrass approximation theorem. *Math. Magazine*, **21** (1948), 167–184, 237–254.

[80] G. Strang and G. Fix. A Fourier analysis of the finite element variational method. In *Constructive aspects of functional analysis*, pages 793–840. C.I.M.E, Il Ciclo 1971, 1973.

[81] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley–Cambridge Press, 1996.

[82] A. F. Timan. A strengthening of Jackson's theorem on the best approximation of continuous functions by polynomials on a finite interval of the real axis. *Dokl. Akad. Nauk. SSSR*, **78** (1951), 17–20.

[83] M. Vetterli and J. Kovačević. *Wavelets and Subband Coding*. Prentice Hall, 1995.

[84] E. Voronowskaja. Détermination de la forme asymptotique d'approximation de la forme de S. Bernstein I,II. *C. R. Acad. Sci. U.S.S.R.*, (1930), 563–568,595–600.

[85] K. Weierstraß. Über die analytische Darstellbarkeit sogenannter willkührlicher Funktionen reller Argumente. *Sitzungsber. Kgl. Preuss. Akad. Wiss. Berlin*, (1885), 633–639,789–805.

[86] K. Yosida. *Functional Analysis*. Grundlehren der mathematischen Wissenschaften. Springer–Verlag, 1965.

[87] A. Zygmund. Smooth functions. *Duke Math. J.*, (1945), 47–76.

8 Index

Abschluß, 19
Abstand, 37
Algebra, 18
  punktetrennende, 19, 22
Algorithmus
  Remez-, 65, *65*
    Austausch, 61
    Beispiel, 66
Alternante, 55, *55*, 59, 61, 65
  Eindeutigkeit, 58
Alternantensatz, 55
AMEL'KOVIČ, V., 78
Analysis
  harmonische, 114, 138
  Multiresolution-, *siehe* MRA 138
Approximation
  beste, *siehe* Bestapproximation 37
  nichtlineare, 36
  Shape preserving, 67
  Simultan-, 67, 69–72
Approximationsgute
  von Skalenräumen, 152
Approximationsgüte, *92*, 93, 96, 100, 102, 104–106, 108, 133
  lokale, 111
  translationsinvarianter Räume, 133
  von Wavelets, 155
Approximationsordnung, 93
  ganzzahlige, 107
Approximationssatz
  Bernstein–Satz, 100, 102, 111
  Bishop, 22
  Jackson–Satz, 96, 104, 109, 133
  Korovkin, 82
  Müntz, 27
  Stone, 19
  Weierstraß, 7, 15
Auflösung, 140

B–Spline
  kardinaler, 118, 131
BAJŠANSKI, B., 78
Banachraum, 13
Basis
  Riesz, 139
  Riesz-, 139
Bedingungen
  Strang–Fix, 129, *129*, 131, 133, 134, 152–155
BERENS, H., 81
BERNSTEIN, S., 16, 26, 93, 100
Bernsteinpolynom
  Ableitung, 67, *68*, 72
  achsenkonvexes, 87
  Graderhöhung, 79, *85*
  Konvexität, 73, *78*, 86
  monotone Konvergenz, *78*, 86, 90
  multivariates, 82–91
  Optimalität, 81
  Randverhalten, 84
  Saturation, 76, 78
  Variationsverminderung, 74
  Voronovskaja–Formel, 76, 91
Bestapproximation, *37*, 55, 59, 93
  Beispiele, 57
  Bestimmung, 59, *65*
  Charakterisierung, 41, 44, 45
  diskrete, 59, *63*, 65
  Eindeutigkeit, 40, 49
  Existenz, 37
  polynomiale, 67
BOJANIĆ, R., 78
BOLYAI, J., 139

BONAPARTE, N., 5
BÉZIER, P., 16, 78

CHAMPOLLION, J. F., 5
CRAMER, G., 31

DAHMEN, W., 90
DAUBECHIES, I., 152
DE LA VALLÉE–POUSSIN, CH., 58
Determinante
    Gramsche, 29, 31
    Vandermonde-, 50
DEVORE, R., 81
Dichtheit
    trigonometrischer Polynome, 7
Differentialoperator
    elliptischer, 91
Differenz
    Vorwärts-, 68–70
Differenzierbarkeit, 105
    Sobolev-, 127
DINGHAS, A., 82
DIRICHLET, L., 7
Dirichlet–Problem, 91
DU BOIS–REYMOND, P., 6

Einbettungsproblem, 162
Einheitskugel, 39
Einheitssimplex, *44*
EINSTEIN, A., 26
Entwicklung
    B–adische, 162
Extremalpunkte, 41

Faltung, *7*, 97, 115
    semidiskrete, 116
FEJÉR, L., 9
Fejérsche Mittel, 10
FIX, G., 129
Formel
    Taylor-, 136
FOURIER, J. B., 5
Fourierkoeffizienten, 5, 144
Fourierreihe, *6*, 125
    Divergenz, *6*, 9
    Du Bois–Reymond, 6
    Partialsumme, *6*, 9, 146
Fouriertransformation, 119
    inverse, 120, 127, 146
Fouriertransformierte, *119*, 120, 124
Funktion
    achsenkonvexe, 88
    Aktivierungs-, 176
    charakteristische, 117, 140, 151
    konvexe, 88
    Ridge-, 176
    Sigmoid-, 177
    stabile, 153
Funktionen
    gerade, 108
    glatte, 107
    orthogonale, 139
    orthonormale, 145
    periodische, *5*
    stabile, 139, 144, 145
    stetige, 5
    verfeinerbare, 141, *141*

Gleichstetigkeit, 95
Glättemodul, 101, 102
GRAM, J., 29
Gruppe
    additive, 114
    multiplikative, 114

HAAR, A., 49, 139
Haar–Raum, *49*, 55, 58, 59, 65
    Beispiele, 52
    Charakterisierung, 49
    reeller, 56
HAUSDORFF, F., 18
HILBERT, D., 159
homöomorph, 54
Hutfunktion, 142
Hülle
    konvexe, *44*, 45
Hülle, konvexe, 82

Identität
approximative, 11, *11*, 12, 17
Index
kritischer, 127
Interpolation, 19, 49
stückweise lineare, 51
Intervall
der Stufe k, 163

JACKSON, D., 96

Kern, *7*, 11
Dirichlet, 8
Dirichlet-, *7*
Fejér-, 10, *10*, 12, 97, 122
Jackson-, *97*, 104
positiver, 11
KOLMOGOROFF, A. N., 41, 161
Kolomogoroff
Satz von, 162
Komplement
orthogonales, 147
Kompression, 157
Konvexität, 37, 72, 85
Achsen-, 87, 90
Richtungs-, 85
Umkehrsatz, 90
Konvexkombination, *44*, 88
strikte, 45
Konvextät, 86
Koordinaten
baryzentrische, 82
KORNEĬČUK, N. P., 99
KOROVKIN, P., 81
Kovergenzordnung
lineare, 176
Kriterium
Kolmogoroff-, 41, 45, 50
Kurven
Bézier-, 78

LANDAU, CH., 129
Layer, 177
LEBESGUE, H., 114
LIPSCHITZ, R., 96
Lipschitz–Klasse, 107
LORENTZ, G. G., 82, 133, 161

MAIRHUBER, J. C., 54
Majorante
konkave, 99
MALLAT, S., 138
Matrix
bedingt positiv definite, 85
Gramsche, 29, 31, 149
Hesse-, 86
Maximum
Bestimmung, 62
Maß
Haar-, 114
Maß
Haar-, 119
Lebesgue-, 114
Menge
antisymmetrische, 22, *22*
konvexe, 37
MICCHELLI, C. A., 90
Momente, *154*
verschwindende, 154
Monotonie, 72
MRA, *138*, 148
Modellfall, 140
orthonormale, 148
Multiindex, 83
Länge, 83
Multiresolution Analysis, *siehe* MRA 138
MÜNTZ, C. H., 26

Netz
neuronales, 176
Neuron, 176
Nomographie, 158
Norm
strikt konvexe, 39, 40
Supremums-, 5
Normalengleichungen, 30
Nullstellen, 49

Operator
Bernstein–Durrmeyer, 91
Faltungs-, 7, 97
positiver, 11, 82
Orthogonalisierungstrick, 145

Plancherel, 124
POISSON, S., 125
Polynom
algebraisches, *15*, 18, 36, 92
Bernstein-, *16*
Bernstein–Bézier-, 16, 83
Reproduktion, 127, 129
Taylor-, 133, 135
trigonometrisches, 6, *7*, 36, 92, 100, 146
Positivität, 72
Produkt
inneres, 29
Projektion, 147
metrische, 37
orthogonale, 147

Quasiinterpolant, 133

RADEMACHER, H., 129
Raum
$C(X)$, 14
$C_u(X)$, 14
$L_2(\mathbb{R})$, 124
L_p, 13
$\ell_p(\mathbb{Z})$, 115
Banach-, 13, 133
Folgen-, 115
FSI, 115
Haar-, *siehe* Haar–Raum 49
Hausdorff-, 18
Hilbert-, 147
PSI, 115
Skalen-, 139, 147
translationsinvarianter, *115*, 116, 133, 139
Tschebyscheff-, *siehe* Haar–Raum 49
Wavelet-, 147
Rechtschreibreform, 145
Regel
Cramersche, 31
Reihe
Fourier-, 5, *siehe* Fourierreihe 6
trigonometrische, 119
Reihenfolge
alphabetische, 129
REMEZ E. Y., 59
Restglied
Integral-, 136
Richtung, 85
RIESZ, F., 139
RIESZ, M., 139
RIVLIN, T., 44

Saturation, 67, *76*, 82
Satz
von Kolmogoroff, 161, 162
von Mairhuber, 54
von Rivlin und Shapiro, 44
von Stone, 19
von Weierstrass, 15
SCHMID, H. J., 86
SCHOENBERG, I. J., 129
SCHWARZ, H. A, 80
Shape properties, 72
SHAPIRO, H. S., 44
Signatur, *43*, 58
extremale, *43*, 44, 48, 56
Simplex
achsenparalleles, 88
Simultanapproximation, 67
Singularität
Erkennung, 156
Skalenraum, 139, 152
Skalierung, 118, 120
Skalierungsfunktion, *139*, 148, 153, 154
mit kompaktem Träger, 151
orthonormale, 143, 148, 149, 151
stetige, 151
SOBOLEV, S. L., 127
Spline, 118
SPRECHER, D., 161
Stabilität, *139*, 144

Stammfunktion, 155
Standardsimplex, *83*
    Seiten, 84
Stetigkeit
    gleichgradige, 95
    gleichmäßige, 12, 17
    Lipschitz, 107
    Lipschitz-, 96, 104, 161, 168, 172
Stetigkeitsmodul, 70, 96, 98, 100, *111*
    höherer Ordnung, *siehe* Glättemodul 101
STIELTJES, TH. J., 74
STONE, M., 18
STRANG, G., 129
Subalgebra, 18
Subharmonizität, 91
Summenformel
    Poisson-, 125, 130

Teilräume
    verschachtelte, 138
Threshold, 176
Thresholding, 157
TIMAN, A. F., 111
Torus, 5, *5*, 13, 54
Trainingsdaten, 179
Translation, 120
Translationsinvarianz, 114
Träger
    kompakter, 153
TSCHEBYSCHEFF, P., 55
Tschebyscheff–System, 49

Ungleichung
    Bernsteinsche, 100
    Hölder-, 28, 117
    Jensensche, 85

Variation
    eines Vektors, 73
    totale, 73, 74
Vektorraum, 37
Verfeinerungsgleichung, 141
    Fourier–transformierte, 141
    Koeffizienten, 148
Verfeinerungsgleichungen
    Koeffizienten, 156
VORONOVSKAJA, E., 76
Vorzeichenwechsel, 58

Wavelet, *147*, 148
    -Darstellung, 148
    -Koeffizienten, 155, 156
    biorthogonales, 157
    Daubechies-, 152–154, 156
    Existenz, 148
    Haar-, 151
    orthogonales, 147
WEIERSTRASS, K., 7

Zahl
    irrationale, 167
Zweiskalenbeziehung, 141
ZYGMUND, A., 107
Zygmund–Klasse, 107, 108