Der Approximationssatz von Stone-Weierstraß

Wir haben in der Vorlesung bereits gesehen, dass unendlich oft differenzierbare Funktionen $f : [a, b] \to \mathbb{R}$ unter gewissen Bedingungen uniform durch ihre Taylorpolynome approximiert werden können. Hier wollen wir einen deutlich allgemeineren Sachverhalt beweisen, die *Möglichkeit der Approximation stetiger Funktionen*.
Im folgenden sei X stets ein kompakter metrischer Raum. Wie in der Vorlesung bezeichnen wir mit $C(X, \mathbb{R})$ den Raum der stetigen Funktionen auf X. Im Kapitel 5 der Vorlesung haben wir gesehen, dass dieser Raum mit der Metrik

$$d(f, g) := \sup\{|f(x) - g(x)| \,:\, x \in X\}$$

ein vollständiger metrischer Raum ist, und das Konvergenz in $C(X, \mathbb{R})$ der gleichmäßigen Konvergenz der Funktionen entspricht. Inhaltlich schließen die folgenden Betrachtungen an die Behandlung der Funktionenräume $C(X, \mathbb{R})$ und $B(X, \mathbb{R})$ im Kapitel 5 der Vorlesung an, sind also insbesondere völlig unabhängig von der Differentialrechnung.

Definition 1. Ein *Funktionenring auf* X ist eine Teilmenge $R \subset C(X, \mathbb{R})$ mit folgenden Eigenschaften:

- $\forall\, f, g \in R \,:\, f + g \in R$, und
- $\forall\, f, g \in R \,:\, f \cdot g \in R$.

Im folgenden bezeichnen wir eine Funktion $f : [a, b] \to \mathbb{R}$ als *polynomiale Funktion*, falls ein Polynom $P(t)$ existiert, so dass $f(t) = P(t)$ für alle $t \in [a, b]$.

Beispiele. Die Teilmenge der polynomialen Funktionen $\mathcal{P} \subset C([0, 1], \mathbb{R})$ bildet einen Funktionenring, ebenso wie die Teilmenge $\mathcal{P}_{\geq n}$ der polynomialen Funktionen, welche nur Monome vom Grad mindestens n enthalten. Auch die konstanten Funktionen $f(x) = c$ mit $c \in \mathbb{R}$ bilden einen Funktionenring in $C(X, \mathbb{R})$ für jedes X.

Bemerkung. Genau genommen haben wir soeben *reelle* Funktionenringe definiert. Ersetzt man in der Definition $C(X, \mathbb{R})$ durch $C(X, \mathbb{C})$, so erhält man *komplexe* Funktionenringe.

Wir sagen, dass ein Funktionenring *die Punkte von* X *trennt*, wenn es zu je zwei Punkten $x_1 \neq x_2 \in X$ eine Funktion $f \in R$ mit $f(x_1) \neq f(x_2)$ gibt. Die Polynome aus unserem Beispiel trennen Punkte in $[0, 1]$, während die konstanten Funktionen offenbar die Punkte von X nicht trennen. Wir können nun Stone's Fassung des Satzes von Weierstraß formulieren:

Theorem 1 (Approximationssatz für reellwertige Funtionen). *Sei X ein kompakter metrischer Raum und $R \subset C(X, \mathbb{R})$ ein Funktionenring, so dass*

- *R alle konstanten Funktionen $f(x) = c$ mit $c \in \mathbb{R}$ enthält, und*
- *R Punkte trennt.*

Dann ist $R \subset C(X, \mathbb{R})$ dicht, d.h. zu jeder stetigen Funktion $f : X \to \mathbb{R}$ existiert eine Folge $\{f_n\}_{n\geq 1}$ mit $f_n \in R$ und $f_n \rightrightarrows f$.

Wir notieren auch noch die

Folgerung 1 (klassischer Approximationssatz von Weierstraß).
Ist $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, so existiert zu jedem $\varepsilon > 0$ eine polynomiale Funktion $Q_\varepsilon : [a, b] \to \mathbb{R}$ mit $|Q_\varepsilon(x) - f(x)| < \varepsilon$ für alle $x \in [a, b]$.
□

Bevor wir den Beweis des Approximationsatzes beginnen, benötigen wir einige Vorbereitungen. Zunächst beginnen wir mit einem

Lemma 2. *Es existiert eine Folge $\{P_n\}_{n\geq 1}$ von polynomialen Funktionen $P_n : [0, 1] \to \mathbb{R}$, die gleichmäßig gegen die Funktion $f(t) = \sqrt{t}$ konvergieren.*

Beweis. Wir definieren die Funkionen $P_n(t)$ rekursiv durch

$$P_0(t) \equiv 0\,, \quad P_{n+1}(t) = \frac{1}{2}(t - P_n^2(t)) + P_n(t) \text{ für } n \geq 0.$$

Wir erhalten also $P_1(t) = \frac{1}{2}t$, $P_2(t) = \frac{1}{2}(t - \frac{1}{4}t^2) + \frac{1}{2}t$ usw. Die so definierten Funktionen P_n sind für $n \geq 1$ jeweils Polynome vom Grad 2^{n-1}.
Wir überzeugen uns nun, dass

$$\forall\, n \geq 0\, \forall\, t \in [0, 1] \,:\, 0 \leq P_n(t) \leq \sqrt{t}.$$

Der Beweis dazu wird durch vollständige Induktion geführt. Für $n = 0$ ist die Behauptung nach Definition von P_0 wahr. Gilt aber die Aussage für P_n, so folgt

$$\begin{aligned}
\sqrt{t} - P_{n+1}(t) &= \sqrt{t} - \frac{1}{2}(t - P_n^2(t)) - P_n(t) \\
&= (\sqrt{t} - P_n(t))(1 - \frac{1}{2}(\sqrt{t} + P_n(t)) \\
&\geq \underbrace{(\sqrt{t} - P_n(t))}_{\geq 0 \text{ nach IV}} \underbrace{(1 - \frac{1}{2}2\sqrt{t})}_{\geq 0 \text{ für } t \in [0,1]} \geq 0,
\end{aligned}$$

wobei wir bei der Abschätzung die Induktionsvoraussetzung $P_n(t) \leq \sqrt{t}$ benutzt haben.

Außerdem gilt für alle $t \in [0,1]$

$$P_{n+1}(t) = P_n(t) + \frac{1}{2}\underbrace{(t - P_n^2(t))}_{\geq 0} \geq P_n(t),$$

d.h. für festes $t \in [0,1]$ ist die Folge $\{P_n(t)\}_{n\geq 0}$ monoton wachsend und von oben durch $\sqrt{t}$ beschränkt, konvergiert also gegen ein $P(t) \leq \sqrt{t}$. Der Grenzwert muss aber die Gleichung

$$P(t) = \frac{1}{2}(t - P^2(t)) + P(t)$$

erfüllen, d.h. es gilt $P(t)^2 = t$, woraus $P(t) = \sqrt{t}$ folgt. Wir haben also gezeigt, dass die Folge von Polynomen $P_n(t)$ punktweise und monoton gegen $f(t) = \sqrt{t}$ konvergiert. Nach dem Satz von Dini folgt daraus aber die gleichmäßige Konvergenz $P_n(t) \rightrightarrows \sqrt{t}$. $\square$

Für jede stetige Funktion $f : X \to \mathbb{R}$ ist die Funktion $|f| : X \to \mathbb{R}$ mit $|f|(x) := |f(x)|$ ebenfalls stetig, denn sie ist eine Verknüpfung von stetigen Funktionen. Es folgt, dass für beliebige Funktionen $f, g \in C(X, \mathbb{R})$ die beiden Funktionen

$$\begin{aligned} \min(f,g) &= \frac{1}{2}\left(f(x) + g(x) - |f(x) - g(x)|\right) \quad \text{und} \\ \max(f,g) &= \frac{1}{2}\left(f(x) + g(x) + |f(x) - g(x)|\right) \end{aligned}$$

ebenfalls stetig, also Elemente in $C(X, \mathbb{R})$ sind.

Lemma 3. *Ist $R \subset C(X, \mathbb{R})$ ein Funktionenring, der alle Konstanten enthält, und sind $f, g \in R$, so folgt $|f| \in \overline{R}$, $\min(f,g) \in \overline{R}$ und $\max(f,g) \in \overline{R}$, wobei $\overline{R} \subset C(X, \mathbb{R})$ die abgeschlossene Hülle von R im metrischen Raum $C(X, \mathbb{R})$ bezeichnet.*

Beweis. Jedes $f \in R$ ist stetig, also (weil X kompakt ist) beschränkt. Es existiert also ein $M \in \mathbb{R}$, so dass für alle $x \in X$

$$|f(x)| \leq M$$

gilt. Da R alle Konstanten enthält, ist $\frac{1}{M} \in R$, also auch $\frac{1}{M} f \in R$, und es gilt $\left|\frac{f(x)}{M}\right| \leq 1$ und somit

$$0 \leq \frac{f(x)^2}{M^2} \leq 1$$

für alle $x \in X$.
Sein nun P_n die Folge von Polynomen auf $[0,1]$ aus Lemma 2 mit $P_n(t) \rightrightarrows \sqrt{t}$. Dann konvergiert die Folge von Funktionen $P_n(\frac{f(x)^2}{M^2})$ gleichmäßig gegen $\frac{|f(x)|}{M}$.

Da die P_n Polynome sind und $\frac{f(x)^2}{M^2} \in R$, folgt auch $P_n(\frac{f(x)^2}{M^2}) \in R$ für alle $n \in \mathbb{N}$. Dann folgt aber auch $M \cdot P_n(\frac{f(x)^2}{M^2}) \in R$. Da

$$M \cdot P_n(\frac{f(x)^2}{M^2}) \rightrightarrows |f(x)|,$$

haben wir $|f| \in \overline{R}$ bewiesen. Da R alle Konstanten enthält, enthält es mit jeder Funktion g auch $-g$, also mit f und g auch $f-g$. Da sich $\max(f,g)$ und $\min(f,g)$ als Linearkombination von $f, g \in R$ und $|f-g| \in \overline{R}$ schreiben lassen, sind diese auch in $\overline{R}$ enthalten. $\square$

Wir kommen nun zum

Beweis des Approximationssatzes.
Wir werden zeigen: Für alle $f \in C(X,\mathbb{R})$ und alle $\varepsilon > 0$ existiert ein $f_\varepsilon \in R$ mit

$$d(f, f_\varepsilon) = \sup\{\,|f(x) - f_\varepsilon(x)| \,:\, x \in X\} < 2\varepsilon.$$

1. Schritt: Wir zeigen zunächst, dass zu $x_1 \neq x_2 \in X$ und $r_1, r_2 \in \mathbb{R}$ eine Funktion $g \in R$ mit $g(x_1) = r_1$ und $g(x_2) = r_2$ existiert.
Nach Voraussetzung existiert $h : X \to \mathbb{R}$ mit $h(x_1) \neq h(x_2)$. Da R alle Konstanten enthält, enthält es auch die Funktion $g^* : X \to \mathbb{R}$ mit

$$g^*(x) := \frac{h(x) - h(x_1)}{h(x_2) - h(x_1)}.$$

Man beachte, dass $g^*(x_1) = 0$ und $g^*(x_2) = 1$, so dass $g(x) := r_1 + (r_2 - r_1)g^*(x)$ die verlangten Eigenschaften hat.

2. Schritt: Seien nun $f \in C(X,\mathbb{R})$ und $\varepsilon > 0$ gegeben. Wir zeigen: Zu jedem Punkt $b \in X$ existiert eine offene Umgebung $V_b \subset X$ von b und eine stetige Funktion $f_b \in \overline{R}$ mit

$$f_b(x) < f(x) + \varepsilon \quad \text{für } x \in X \quad \text{und } f_b(x) > f(x) - \varepsilon \quad \text{für } x \in V_b. \tag{1}$$

Wir fixieren $b \in X$ und finden zu beliebigem $a \in X$ mit der Konstruktion in Schritt 1 eine Funktion $f_{a,b} \in R$ mit $f_{a,b}(a) = f(a)$ und $f_{a,b}(b) = f(b)$. Nun definieren wir die offenen Mengen

$$U_{a,b} := \{x \in X \mid f_{a,b}(x) < f(x) + \varepsilon\} \quad \text{und} \quad V_{a,b} := \{x \in X \mid f_{a,b}(x) > f(x) - \varepsilon\}.$$

Da a und b in $U_{a,b} \cap V_{a,b}$ enthalten sind, sind beide Mengen nicht leer. Da X kompakt ist und die Mengen $\{U_{a,b}\}_{a \in X}$ eine offene Überdeckung von X bilden, existieren endliche viele Punkte $a_1, \ldots, a_k \in X$, so dass

$$X = U_{a_1,b} \cup U_{a_2,b} \cup \cdots \cup U_{a_k,b}.$$

Nach Lemma 3 ist die Funktion

$$f_b := \min(f_{a_1,b}, f_{a_2,b}, \ldots, f_{a_k,b}) \in \overline{R},$$

und nach Konstruktion haben wir $f_b(x) \leq f_{a_i,b}(x) < f(x) + \varepsilon$ für $x \in U_{a_i,b}$, also

$$f_b(x) < f(x) + \varepsilon \quad \text{für alle } x \in X.$$

Definieren wir

$$V_b := V_{a_1,b} \cap V_{a_2,b} \cap \cdots \cap V_{a_k,b},$$

so gilt ebenfalls nach Konstruktion $b \in V_b$ und

$$f_b(x) > f(x) - \varepsilon \quad \text{für alle } x \in V_b.$$

3. Schritt: Wir konstruieren nun eine Funktion $f_\varepsilon^* \in \overline{R}$ mit

$$|f_\varepsilon^*(x) - f(x)| < \varepsilon \quad \text{für alle } x \in X.$$

Im Schritt 2 haben wir zu jedem $b \in X$ eine Funktion $f_b \in \overline{R}$ und eine offene Umgebung V_b konstruiert, so dass (1) gilt. Die Mengen $\{V_b\}_{b \in X}$ bilden eine offene Überdeckung von X, also finden wir wiederum endlich viele Punkte $b_1, \ldots, b_r \in X$ mit

$$X = V_{b_1} \cup V_{b_2} \cup \cdots \cup V_{b_k}.$$

Nach Lemma 3 ist die Funktion

$$f_\varepsilon^* := \max(f_{b_1}, f_{b_2}, \ldots, f_{b_r}) \in \overline{\overline{R}} = \overline{R},$$

und es gilt

$$f_\varepsilon^*(x) < f(x) + \varepsilon \quad \text{für alle } x \in X.$$

Andererseits gilt für $x \in V_{b_i}$ auch $f_\varepsilon^*(x) \geq f_{b_i}(x) > f(x) - \varepsilon$, und damit

$$f_\varepsilon^*(x) > f(x) - \varepsilon \quad \text{für alle } x \in X.$$

Zusammen ergeben diese beiden Aussagen die Behauptung von Schritt 3.

4.Schritt: Wir zeigen die eingangs formulierte Behauptung, dass es zu jeden $\varepsilon > 0$ eine Funktion $f_\varepsilon \in R$ mit $d(f, f_\varepsilon) < 2\varepsilon$ gibt.
In Schritt 3 haben wir bereits $f_\varepsilon^* \in \overline{R}$ mit $d(f, f_\varepsilon^*) < \varepsilon$ konstruiert. Da $f_\varepsilon^* \in \overline{R}$, finden wir aber nach der Definition der abgeschlossenen Hülle ein $f_\varepsilon \in R$ mit $d(f_\varepsilon, f_\varepsilon^*) < \varepsilon$, so dass die Behauptung aus der Dreiecksungleichung folgt.
Dies beendet den Beweis des Approximationsatzes. □

Der Approximationssatz besitzt auch eine komplexe Formulierung, welche sich leicht aus der reellen Version herleiten lässt.

Theorem 4 (Approximationssatz für komplexwertige Funktionen). *Sei X ein kompakter metrischer Raum und $R \subset C(X, \mathbb{C})$ ein Funktionenring, so dass*

- *R alle konstanten Funktionen $f(x) = c$ mit $c \in \mathbb{C}$ enthält,*

- *R Punkte trennt, und*
- *R abgeschlossen unter Konjugation ist, d.h. aus $f = u + iv \in R$ folgt auch $\bar{f} = u - iv \in R$.*

Dann ist R dicht in $C(X, \mathbb{C})$.

Beweis. Wir betrachten den Teilring $R^* \subset R$ der reellwertigen Funktionen in R, und wollen zunächst beweisen, dass R^* die Voraussetzungen des Approximationssatzes für reellwertige Funktionen erfüllt. Zunächst stellen wir fest, dass R^* alle reellen Konstanten enthält.
Seien nun $x_1 \neq x_2 \in X$. Dann existiert nach Voraussetzung ein $h \in R$ mit $h(x_1) \neq h(x_2)$. Wie im ersten Schritt des Beweises des reellen Approximationssatzes finden wir nun eine Funktion $g \in R$, nämlich

$$g(x) := \frac{h(x) - h(x_1)}{h(x_2) - h(x_1)},$$

mit $g(x_1) = 0$ und $g(x_2) = 1$. Nach Voraussetzung ist $\overline{g} \in R$, also auch $f := g + \overline{g} \in R$. Nun ist aber nach Konstruktion $f = 2\,\mathrm{Re}(g)$ eine reellwertige Funktion, d.h. $f \in R^*$. Außerdem gilt $f(x_1) = 0$ und $f(x_2) = 2$. Damit haben wir gezeigt, dass R^* auch Punkte trennt.
Nach dem reellen Approximationssatz ist also $R^* \subset C(X, \mathbb{R})$ dicht. Ist also $f = u + iv \in C(X, \mathbb{C})$ und $\varepsilon > 0$ gegeben, so existieren $f_1, f_2 \in R^*$ mit $d(f_1, u) < \varepsilon$ und $d(f_2, v) < \varepsilon$. Da R Konstanten (also auch i) enthält, ist die Funktion $f_\varepsilon := f_1 + if_2 \in R$, und aus der Dreiecksungleichung folgt

$$d(f, f_\varepsilon) < 2\varepsilon.$$

□

Folgerung 2. *Ist $f : [a, b] \to \mathbb{C}$ stetig und $\varepsilon > 0$, so existiert eine polynomiale Funktion $P(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ mit $a_i \in \mathbb{C}$, so das $|f(x) - P(x)| < \varepsilon$ für alle $x \in [a, b]$.* □

Approximation periodischer Funktionen

Um dem Eindruck vorzubeugen, dass der allgemeine Approximationssatz eine Nutzlose Verallgemeinerung des als Folgerung 1 formulierten klassischen Approximationssatzes ist, wollen wir hier noch die Approximation periodischer Funktionen durchtrigonometrische Polynome behandeln. Dazu betrachten wir für $K = \mathbb{R}$ oder $K = \mathbb{C}$ den Raum

$$C_{2\pi\text{-periodisch}}(\mathbb{R}, K) := \{f : \mathbb{R} \to K \mid f(t + 2\pi) = f(t) \text{ für alle } t \in \mathbb{R}\}.$$

Wir haben schon bewiesen, dass die Exponentialabbildung eine 2π-periodische Abbildung

$$\begin{aligned} \exp : \mathbb{R} &\to S^1 \\ t &\mapsto e^{it} \end{aligned}$$

induziert. Daraus folgt, dass durch die Zuordnung $f(t) := f^*(e^{it})$ eine Bijektion zwischen den Elementen $f \in C_{2\pi\text{-periodisch}}(\mathbb{R}, K)$ und den Elementen $f^* \in C(S^1, K)$ beschrieben wird, welche auch die Abstände erhält. Somit ist eine Teilmenge $R \subset C_{2\pi\text{-periodisch}}(\mathbb{R}, K)$ genau dann dicht, wenn die dazugehörende Teilmenge $R^* \subset C(S^1, K)$ dicht ist.

Definition 2. Ein (reelles) *trigonometrisches Polynom* ist eine Funktion $P : \mathbb{R} \to \mathbb{R}$ der Form

$$P(x) = A_0 + \sum_{k=1}^{n} \left(A_k \cos(kx) + B_k \sin(kx)\right).$$

Es ist klar aus der Definition, dass alle trigonometrischen Polynome 2π-periodisch sind. Wir erinnern uns auch, dass $e^{ikx} = \cos(kx) + i\sin(kx)$, so dass

$$A_k \cos(kx) + B_k \sin(kx) = \frac{1}{2}\left((A_k - iB_k)e^{ikx} + (A_k + iB_k)e^{-ikx}\right).$$

Also lässt sich ein reelles trigonometrisches Polynom auch als

$$P(x) = \sum_{k=-n}^{n} C_n e^{inx} \quad \text{mit } C_n \in \mathbb{C}, C_{-n} = \overline{C}_n$$

schreiben.

Definition 3. Ein komplexwertiges trigonometrisches Polynom ist eine 2π-periodische Funktion $P : \mathbb{R} \to \mathbb{C}$ der Form

$$P(x) = \sum_{k=-n}^{N} C_k e^{ikx} \quad \text{mit } C_k \in \mathbb{C}.$$

Als Folgerung aus dem Approximationssatz für reellwertige Funktionen erhalten wir nun sofort

Theorem 5 (Approximationssatz für periodische Funktionen).

1. *Ist $f \in C_{2\pi\text{-periodisch}}(\mathbb{R}, \mathbb{R})$, so gibt es eine Folge reeller trigonometrischer Polynome $P_n(x)$ mit $P_n \rightrightarrows f$.*

2. *Ist $f \in C_{2\pi\text{-periodisch}}(\mathbb{R}, \mathbb{C})$, so gibt es eine Folge komplexer trigonometrischer Polynome $P_n(x)$ mit $P_n \rightrightarrows f$.*

Beweis. zu 1.)
Wir bezeichnen mit $R \subset C_{2\pi\text{-periodisch}}(\mathbb{R}, \mathbb{R})$ die Teilmenge der trigonometrischen Polynome, und mit R^* die dazugehörende Teilmenge in $C(S^1, \mathbb{R})$. Dann enthält R (und somit auch R^*) nach Definition alle Konstanten, und aus den

Additionstheoremen folgt, dass R (und somit auch R^*) abgeschlossen unter Produktbildung ist, denn es gilt z.B.

$$\cos(kx) \cdot \cos(\ell x) = \frac{1}{2}\left(\cos((k+\ell)x) + \cos((k-\ell)x)\right) \quad \text{etc.}$$

Außerdem trennt R^* Punkte in S^1, denn für die Funktionen $P(x) = \sin(x)$ und $Q(x) = \cos(x)$ in R gilt $P^*(z) = \mathrm{Im}(z)$ und $Q^*(z) = \mathrm{Re}(z)$, und mindestens eine dieser Funktionen unterscheidet zwei verschiedene Punkte $z_1 \neq z_2 \in S^1$. Nach dem Approximationssatz für reellwertige Funktionen ist R^* also dicht in $C(S^1, \mathbb{R})$, und somit ist auch R dicht in $C_{2\pi\text{-periodisch}}(\mathbb{R}, \mathbb{R})$.
zu 2.)
Bezeichnet man mit $R \subset C(S^1, \mathbb{C})$ die Teilmenge der komplexwertigen trigonometrischen Polynome, so ist R ein Funktionenring. Da $\overline{e^{it}} = e^{-it}$, ist R abgeschlossen unter Konjugation. Außerdem enthält R alle Konstanten und alle reellwertigen trigonometrischen Polynome. Da diese Punkte trennen, folgt die Behauptung aus dem Approximationssatz für komplexwertige Funktionen. $\square$