Der Weierstraßsche Approximationssatz

Teilnehmer:

Anna Denkert	Heinrich-Hertz-Oberschule, Berlin
Linus Hagemann	Herder-Oberschule, Berlin
Oliver Hager	Herder-Oberschule, Berlin
Nicolas Klodt	Heinrich-Hertz-Oberschule, Berlin
Celine Liem	Herder-Oberschule, Berlin
Fabian Rivetta	Heinrich-Hertz-Oberschule, Berlin

Gruppenleiter:

Konrad Gröger	Humboldt-Universität zu Berlin

Die Lösungen vieler mathematischer Probleme (Differentialgleichungen, Variationsprobleme, u. a.) sind *Funktionen*. Je nach Fragestellung hat man es mit Funktionen unterschiedlicher Güte zu tun. Es hat sich herausgestellt, dass es möglich und für viele Fragestellungen auch hilfreich ist, „weniger gute“ Funktionen durch „bessere“ Funktionen anzunähern, zu approximieren. Der herausragende Satz in dieser Richtung ist der 1885 von Karl Weierstraß publizierte Satz über die Approximation stetiger Funktionen durch Polynome, der heute *Weierstraßscher Approximationssatz* genannt wird. Dieser Satz hat große Beachtung gefunden und schon bald nach seiner Publikation sind zahlreiche unterschiedliche Beweise für ihn angegeben worden.

Die Gruppe hat sich zuerst mit dem Weierstraßschen Approximationssatz vertraut gemacht und sich dann drei verschiedene Beweise des Satzes erarbeitet. Dabei ging es nicht nur darum, die Details der Beweise nachzuvollziehen, sondern auch darum, Beweisideen zu „verstehen“, d. h., zu verstehen, wie man auf die Idee kommen konnte, den Beweis gerade in der jeweils angegebenen Art zu versuchen. Beispielsweise arbeitet einer der Beweise mit wahrscheinlichkeitstheoretischen Argumenten, obwohl der zu beweisende Sachverhalt mit Zufälligkeit oder Wahrscheinlichkeit nicht das Geringste zu tun hat.

Formulierung des Satzes und Vorbemerkungen

Wir bezeichnen mit X immer ein beschränktes abgeschlossenes Teilintervall der reellen Achse und mit $C(X)$ die Menge der auf dem Intervall X definierten reellwertigen stetigen Funktionen.

Satz 1. *Weierstraßscher Approximationssatz.*

Zu jeder Funktion $f \in C(X)$ und jedem $\varepsilon > 0$ existiert ein Polynom p mit der Eigenschaft

$$\sup_{x \in X} |p(x) - f(x)| < \varepsilon.$$

Definition 1. Eine Funktion $f \in C(X)$ nennt man *gleichmäßig stetig*, wenn

$$\underset{\varepsilon > 0}{\forall} \varepsilon \ \underset{\delta > 0}{\exists} \delta : \ |t - x| < \delta \Rightarrow |f(t) - f(x)| < \epsilon \quad \text{für } x, t \in X. \tag{1}$$

Die Funktionen aus $C(X)$ besitzen die folgenden wichtigen Eigenschaften:
1. Jede Funktion $f \in C(X)$ ist auf X gleichmäßig stetig.
2. Jede Funktion $f \in C(X)$ ist auf X beschränkt.

Diese Eigenschaften kann man indirekt mit dem Satz von Bolzano-Weierstraß beweisen.

Definition 2. Für $f \in C(X)$ nennt man

$$\|f\| := \sup_{x \in X} |f(x)|$$

Supremumsnorm von f. Das Supremum ist auf Grund der Beschränktheit von f endlich, also $\|f\|$ eine reelle Zahl.

Für beliebige Funktionen f und g aus $C(X)$ deutet man $\|f - g\|$ als *Abstand* von f und g.

Definition 3. Eine Folge (f_n) von Funktionen aus $C(X)$ heißt auf X *gleichmäßig konvergent* gegen $f^* \in C(X)$, falls $\lim_{n \to \infty} \|f_n - f^*\| = 0$ ist.

Im Folgenden betrachten wir nur das Intervall $X = [0, 1]$. Nach den Beweisen folgt eine Begründung, warum dies genügt.

Beweis nach S. N. Bernstein

Für den Beweis wird das sogenannte *Bernsteinpolynom* B_n zu f verwendet:

$$B_n(x) := \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k} \text{ für } n \in \mathbb{N},\ x \in X.$$

Hierbei ist $f \in C(X)$ eine gegebene Funktion. Wir setzen

$$g_{n,k}(x) := \binom{n}{k} x^k (1-x)^{n-k} \text{ für } x \in X.$$

Nach dem Binomischen Satz ist

$$\sum_{k=0}^{n} g_{n,k}(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = (x + (1-x))^n = 1^n = 1. \tag{2}$$

Daher ist $B_n(x)$ ein gewichtetes Mittel der Werte $f(\frac{k}{n})$, $k = 0, \ldots, n$. Wir beweisen den Weierstraßschen Approximationssatz, indem wir Folgendes zeigen:

$$\lim_{n\to\infty} \|B_n - f\| = 0. \tag{3}$$

Dazu betrachten wir die Differenz der Funktionswerte von B_n und f:

$$|B_n(x) - f(x)| = \left|\sum_{k=0}^{n} \left(f\left(\frac{k}{n}\right) - f(x)\right) g_{n,k}(x)\right| \leq \sum_{k=0}^{n} \left|\left(f\left(\frac{k}{n}\right) - f(x)\right)\right| g_{n,k}(x).$$

(Die letzte Ungleichung folgt aus der Dreiecksungleichung.)
Um die Behauptung (3) zu beweisen, zeigen wir, dass die Summe für $n \to \infty$ gleichmäßig bezüglich $x \in X$ gegen 0 konvergiert. Wir wählen $\varepsilon > 0$ beliebig und dazu nach (1) ein passendes $\delta > 0$. Wir betrachten für k zwei Fälle:

1. Fall $|\frac{k}{n} - x| < \delta$.

Für diese k kann die gleichmäßige Stetigkeit von f ausgenutzt werden. Es ist

$$\sum_{\{k; |\frac{k}{n} - x| < \delta\}} \left|f\left(\frac{k}{n}\right) - f(x)\right| g_{n,k}(x) \leq \sum_{\{k; |\frac{k}{n} - x| < \delta\}} \varepsilon \cdot g_{n,k}(x) \leq \varepsilon \sum_{\{k; |\frac{k}{n} - x| < \delta\}} g_{n,k}(x) \leq \varepsilon.$$

2. Fall $|\frac{k}{n} - x| \geq \delta$.

In diesem Fall gilt auf Grund der Dreiecksungleichung

$$\sum_{\{k;|\frac{k}{n}-x|\geq\delta\}} \left|f\left(\frac{k}{n}\right) - f(x)\right| g_{n,k}(x) \leq 2\|f\| \sum_{\{k;|\frac{k}{n}-x|\geq\delta\}} g_{n,k}(x)$$

und es ist

$$\sum_{\{k;|\frac{k}{n}-x|\geq\delta\}} g_{n,k}(x) \leq \sum_{\{k;|\frac{k}{n}-x|\geq\delta\}} \left(\frac{\frac{k}{n}-x}{\delta}\right)^2 g_{n,k}(x) \leq \frac{1}{\delta^2}\sum_{k=0}^{n}\left(\frac{k}{n}-x\right)^2 g_{n,k}(x),$$

da $\left(\frac{k}{n} - x\right)^2 \geq \delta^2$, also $\left(\frac{\frac{k}{n}-x}{\delta}\right)^2 \geq 1$ gilt.

Wir beweisen anschließend die folgende Behauptung:

$$\sum_{k=0}^{n}\left(\frac{k}{n} - x\right)^2 g_{n,k}(x) = \frac{x(1-x)}{n} \leq \frac{1}{n} \quad \text{für } x \in X. \tag{4}$$

Damit ist dann gezeigt, dass die Summe aus Fall 2 für hinreichend großes n unabhängig von x kleiner als ε wird. Nach Definition der gleichmäßigen Konvergenz ist somit dann auch die gewünschte Aussage (3) bewiesen.

Beweis der Behauptung (4): Für $k > 0$ ist

$$\frac{k}{n}\cdot\binom{n}{k} = \frac{k}{n}\cdot\frac{n!}{k!(n-k)!} = \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} = \binom{n-1}{k-1}$$

und damit auch

$$\frac{k(k-1)}{n(n-1)}\cdot\binom{n}{k} = \frac{k-1}{n-1}\cdot\binom{n-1}{k-1} = \binom{n-2}{k-2} \text{ für } k > 1.$$

Folglich gilt

$$\begin{aligned}\sum_{k=0}^{n}\frac{k}{n}\binom{n}{k}x^k(1-k)^{n-k} &= x\cdot\sum_{k=1}^{n}\binom{n-1}{k-1}x^{k-1}(1-x)^{(n-1)-(k-1)}\\ &= x\cdot\sum_{k=0}^{n-1}\binom{n-1}{k}x^k(1-x)^{n-1-k} = x\cdot\sum_{k=0}^{n-1}g_{n-1,k}(x) = x\end{aligned} \tag{5}$$

und analog

$$\sum_{k=0}^{n}\frac{k(k-1)}{n(n-1)}g_{n,k}(x) = x^2. \tag{6}$$

Die Behauptung (4) ergibt sich nun durch elementare Rechnungen aus (2), (5) und (6).

Wie oben bemerkt, ist damit die Beziehung $\lim\limits_{n\to\infty} \|B_n - f\| = 0$ bewiesen, also der erste Beweis des Weierstraßschen Approximationssatzes abgeschlossen.

Bernstein hat die Gewichte $g_{n,k}(x)$ durch wahrscheinlichkeitstheoretische Überlegungen gefunden. Diese werden hier kurz beschrieben.

Für $x \in [0,1]$ sei E_x ein Experiment mit genau zwei möglichen Ergebnissen, genannt „Erfolg“ und „Misserfolg“; dabei sei x die Wahrscheinlichkeit für das Eintreten des Erfolgs. Dann ist $1-x$ die Wahrscheinlichkeit des Misserfolgs. (Modell für E_x ist ein Glücksrad mit Anteil x für Erfolg und Anteil $1-x$ für Misserfolg.) Die Zahl $g_{n,k}(x)$ ist die Wahrscheinlichkeit dafür, dass bei n-maliger Durchführung des Experiments E_x genau k-mal Erfolg eintritt. Die Wahrscheinlichkeitstheorie lässt erwarten, dass bei vorgegebenem $\delta > 0$ die Summe der Wahrscheinlichkeiten $g_{n,k}(x)$ für diejenigen k, für die $\frac{k}{n}$ in der δ-Umgebung von x liegt, mit wachsendem n gegen 1 geht. Damit erhalten bei wachsendem n diejenigen Summanden der Bernsteinpolynome, für die $\frac{k}{n}$ nahe x ist (und damit $f(\frac{k}{n})$ nahe $f(x)$), ein immer größeres Gewicht.

Beweis durch Mittelung

Bei der Beweismethode, die nun folgt, wird eine Funktion $f \in C(X)$ durch Funktionen angenähert, die sich durch gewichtete Mittelung *aller* Werte von f ergeben. Das schafft man durch eine geeignete Folge von Gewichtsfunktionen. Edmund Landau hat Gewichtsfunktionen entdeckt, die bei dieser Approximation sehr hilfreich sind. Er nutzte aus, dass sich wichtige Eigenschaften der Gewichtsfunktionen auf die gemittelten Funktionen übertragen. Insbesondere gilt dies auch für die Eigenschaft, ein Polynom zu sein. Er wählte:

$$g_n(x) = c_n(1-x^2)^n \text{ für } x \in [-1,1].$$

Die Konstante c_n wird hierbei so bestimmt, dass gilt:

$$\int_{-1}^{1} g_n(x)\,\mathrm{d}x = 1. \tag{7}$$

Gemittelte Funktionen zum gegebenen $f \in C(X)$ definiert man nun wie folgt:

$$p_n(x) := \int_0^1 g_n(x-t)f(t)\,\mathrm{d}t \quad \text{für } x \in X.$$

Wie bereits erwähnt, überträgt sich die Eigenschaft, ein Polynom zu sein, von g_n auf p_n. Wir wollen das beweisen. Offensichtlich gilt

$$g_n(x-t) = c_n(1-(x-t)^2)^n = \sum_{k=0}^{2n} q_{n,k}(x)t^k$$

für geeignet gewählte Polynome $q_{n,k}$. Daher ist

$$p_n(x) = \int_0^1 \Big(\sum_{k=0}^{2n} q_{n,k}(x)t^k\Big) f(t)\,\mathrm{d}t = \sum_{k=0}^{2n} q_{n,k}(x)\int_0^1 t^k f(t)\,\mathrm{d}t.$$

Da die Integrale nun unabhängig von x sind, sind die Produkte Polynome nach x. Somit ist auch die Summe p_n ein Polynom (genauer: ein auf X eingeschränktes Polynom).

Wir wollen nun zeigen, dass die Funktionenfolge (p_n) mit wachsendem n gleichmäßig gegen f konvergiert, d. h., dass die Beziehung

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

gilt. Dazu unterwerfen wir f zunächst einer Zusatzbedingung, von der wir uns aber am Ende wieder frei machen werden. Die Bedingung lautet

$$f(0) = f(1) = 0.$$

Wir definieren $f(t) := 0$ für $t \notin X$. Auf Grund der Zusatzbedingung ist auch die so fortgesetzte Funktion f stetig, also auf jedem beschränkten abgeschlossenen Intervall gleichmäßig stetig. Man kann nun p_n wie folgt schreiben:

$$p_n(x) = \int_{x-1}^{x+1} g_n(x-t)f(t)\,\mathrm{d}t.$$

Durch Substitution von $s := x - t$ ergibt sich:

$$p_n(x) = \int_{-1}^{1} g_n(s)f(x-s)\,\mathrm{d}s.$$

Nun können wir den Abstand zwischen p_n und f betrachten. Für $x \in X$ ist

$$\begin{aligned} \big|p_n(x) - f(x)\big| &= \Big|\int_{-1}^{1} g_n(s)\big(f(x-s) - f(x)\big)\,\mathrm{d}s\Big| \\ &\leq \int_{-1}^{1} g_n(s)\big|f(x-s) - f(x)\big|\,\mathrm{d}s. \end{aligned} \tag{8}$$

Wir wählen nun $\varepsilon > 0$ beliebig. Dazu gibt es gemäß (1) ein passendes δ mit $0 < \delta < 1$. Wir teilen das letzte Integral nun in drei Teilintegrale und zeigen, dass man alle unabhängig von x beliebig klein machen kann. Auf Grund der Wahl von δ und der Gleichung (7) ist

$$\int_{-\delta}^{\delta} g_n(s)|f(x-s) - f(x)|\,\mathrm{d}s < \int_{-\delta}^{\delta} g_n(s)\varepsilon\,\mathrm{d}s < \varepsilon.$$

Wir betrachten jetzt das Integral von δ bis 1 und analog auch das Integral von -1 bis $-\delta$:

$$\int_{\delta}^{1} g_n(s)|f(x-s) - f(x)|\,\mathrm{d}s \leq 2||f|| \int_{\delta}^{1} g_n(s)\mathrm{d}s.$$

Wir zeigen nun, dass Folgendes gilt:

$$\lim_{n\to\infty} \int_{\delta}^{1} g_n(s)\ \mathrm{d}s = 0.$$

Zunächst wählt man η mit $0 < \eta < \delta$, um c_n nach oben abzuschätzen. Wegen der Bedingung (7) gilt

$$\frac{1}{c_n} = \int_{-1}^{1} (1-x^2)^n\,\mathrm{d}x \geq \int_{-\eta}^{\eta} (1-x^2)^n\,\mathrm{d}x \geq 2\eta(1-\eta^2)^n.$$

Mit Hilfe dieser Ungleichung erhält man

$$\begin{aligned} \int_{\delta}^{1} c_n(1-x^2)^n\,\mathrm{d}x &\leq \frac{1}{2\eta(1-\eta^2)^n} \int_{\delta}^{1} (1-x^2)^n\,\mathrm{d}x \\ &\leq \frac{1}{2\eta(1-\eta^2)^n}(1-\delta^2)^n \\ &= \frac{1}{2\eta}\left(\frac{1-\delta^2}{1-\eta^2}\right)^n < \varepsilon \ \text{ für hinreichend großes } n. \end{aligned}$$

Damit kann man das gesamte Integral aus (8) für hinreichend großes n unabhängig von x kleiner als 3ε machen. Das besagt, dass die Folge p_n gleichmäßig gegen f konvergiert, und beweist so den Weierstraßschen Approximationssatz unter der Zusatzbedingung $f(0) = f(1) = 0$.

Es sei nun $f \in C(X)$ beliebig gegeben. Wir setzen

$$q(x) := f(0) + x(f(1) - f(0)) \text{ für } x \in X.$$

Dann ist $f_1 := f - q \in C(X)$ und $f_1(0) = f_1(1) = 0$. Auf Grund des vorangegangenen Beweises gibt es zu gegebenem $\varepsilon > 0$ ein Polynom p auf X derart, dass $\|f_1 - p\| < \varepsilon$ ist. Dann ist $\|f - (p + q)\| = \|f_1 - p\| < \varepsilon$. Da auch $p + q$ ein Polynom auf X ist, beendet diese Feststellung den zweiten Beweis des Weierstraßschen Approximationssatzes.

Ein elementarer Beweis

In dem nun folgenden dritten Beweis des Weierstraßschen Approximationssatzes wird eine Funktion $f \in C(X)$ zuerst durch eine stetige, stückweise affine Funktion approximiert. Anschließend wird gezeigt, dass sich die Approximation stetiger, stückweise affiner Funktionen durch Polynome auf die Approximation einer einzigen solchen Funktion zurückführen lässt.

Es sei erneut $X = [0, 1]$ und f eine gegebene Funktion aus $C(X)$. Wir wählen ein beliebiges $\varepsilon > 0$ und dazu $\delta > 0$ so, dass die Beziehung

$$|f(t) - f(x)| < \varepsilon, \text{ falls } |t - x| < \delta,$$

gilt. Weiter wählen wir ein $N \in \mathbb{N}$ mit $\frac{1}{N} < \delta$ und teilen X in N gleichlange Teilintervalle der Länge $\frac{1}{N}$. Wir konstruieren eine Funktion a derart, dass a in den Stützpunkten $x_k := \frac{k}{N}$, $k = 0, \ldots, n$, mit f übereinstimmt und dass a zwischen je zwei benachbarten Stützpunkten affin ist. Dann gilt $a \in C(X)$. Da zwei benachbarte Stützpunkte x_k einen kleineren Abstand als δ besitzen, unterscheiden sich die Funktionswerte $f(x)$ auf einem Teilintervall um höchstens ε. Für die konstruierte Funktion a folgt daraus $||f - a|| \leq \varepsilon$.

Definition 4. Für $x \in \mathbb{R}$ wird $x^+ := \max\{0, x\}$ *Positivteil* von x genannt.

Behauptung. Mit geeigneten Konstanten $c_1, \dots, c_N$ kann a in der Form

$$a(x) = a(0) + \sum_{k=1}^{N} c_k (x - x_{k-1})^+ \quad \text{für } x \in X \tag{9}$$

dargestellt werden.

Da beide Seiten der Gleichung (9) in den Teilintervallen $[x_{k-1}, x_k]$ affin sind, genügt es, diese Gleichung für $x = x_k$, $k = 0, \dots, N$, zu erfüllen. Für $x = x_0 = 0$ ist (9) immer erfüllt. Auf Grund der Definition des Positivteils lautet die Gleichung (9) für $x = x_1$

$$a(x_1) = a(0) + c_1(x_1 - x_0).$$

Dies lässt sich durch Wahl von c_1 immer erfüllen, denn es ist $x_1 - x_0 \neq 0$. Analog lautet (9) für $x = x_2$

$$a(x_2) = a(0) + c_1(x_2 - x_0) + c_2(x_2 - x_1).$$

Die Zahl c_1 ist durch $a(x_1)$ bereits festgelegt, aber c_2 ist noch frei wählbar. So zeigt man induktiv für $k = 1, \dots, N$, dass man durch geeignetete Wahl von c_k die Gleichung (9) im Stützpunkt x_k erfüllen kann. Damit ist die Behauptung über die Darstellung von a bewiesen.

Zur Approximation von a durch ein Polynom genügt es wegen (9), jede der Funktionen $x \longmapsto (x - x_{k-1})^+$ auf X durch ein Polynom zu approximieren. Das ist sicher möglich, wenn man die Positivteilfunktion $x \longmapsto x^+$ auf dem Intervall $[-1, 1]$ durch ein Polynom gleichmäßig approximieren kann. Ist die Approximation für jeden Summanden genauer als $\frac{\varepsilon}{N}$, so bleibt der Approximationsfehler für die gesamte Summe kleiner als ε. Wegen der Identität

$$x^+ = \frac{1}{2}(|x| + x) \qquad \text{für } x \in \mathbb{R}$$

genügt es, statt der Positivteilfunktion die Betragsfunktion auf $[-1, 1]$ gleichmäßig beliebig gut durch ein Polynom zu approximieren.

Die Approximation der Betragsfunktion realisieren wir durch eine iterativ definierte Folge von Näherungspolynomen. Wir setzen

$$\begin{aligned} p_0(x) &:= 1 && \text{für } x \in [-1, 1], \\ p_n(x) &:= p_{n-1}(x) - \frac{1}{2}\left((p_{n-1}(x))^2 - x^2\right) && \text{für } x \in [-1, 1]. \end{aligned}$$

Wir schließen den dritten Beweis des Weierstraßschen Approximationssatzes dadurch ab, dass wir die gleichmäßige Konvergenz der Folge (p_n) gegen die Betragsfunktion auf dem Intervall $[-1,1]$ zeigen.
Mit einfachen Argumenten beweist man die Beziehung

$$|x| \le t \le 1 \Rightarrow |x| \le F_x(t) \le t \le 1 \text{ für } x \in [-1,1]$$

für die durch die Vorschrift

$$F_x(t) := t - \frac{1}{2}(t^2 - x^2) \text{ für } t \in [0,1]$$

definierte Funktion F_x. Wegen $p_n(x) = F_x(p_{n-1}(x))$ und $p_0(x) = 1$ gilt also

$$|x| \le p_n(x) \le p_{n-1}(x) \le 1 \text{ für } x \in [-1,1]. \tag{10}$$

Es folgt:

$$p^*(x) := \lim_{n\to\infty} p_n(x) \text{ existiert und } p^*(x) \ge |x| \text{ für } x \in [-1,1].$$

Für $n \to \infty$ folgt aus der Definition der p_n

$$p^*(x) = p^*(x) - \frac{1}{2}((p^*(x))^2 - x^2).$$

Folglich ist $p^*(x) = |x|$ für $x \in [-1,1]$. Es ist noch zu zeigen, dass die Folge (p_n) nicht nur punktweise, sondern auch gleichmäßig konvergiert. Es sei

$$d_n := \sup_{x\in[-1,1]} \big| p_n(x) - |x| \big|. \tag{11}$$

Wir haben zu beweisen, dass $\lim\limits_{n\to\infty} d_n = 0$ ist. Es gilt $0 \le d_n \le d_{n-1}$, weil $p_{n-1}(x) \ge p_n(x) \ge |x|$ ist. Daraus folgt, dass $d^* := \lim\limits_{n\to\infty} d_n$ existiert.

Es gelte nun $\lim_{n\to\infty} \delta_n = 0$ und $\delta_n > 0$ für $n \in \mathbb{N}$. Auf Grund von (11) existiert dann ein $x_n \in [-1,1]$, für das $\big|p_n(x_n) - |x_n|\big| > d_n - \delta_n$ ist. Ohne Beschränkung der Allgemeinheit darf man annehmen, dass die Folge (x_n) konvergiert (erforderlichenfalls geht man unter Berufung auf den Satz von Bolzano-Weierstraß zu einer Teilfolge über). Es gilt (vgl. (10))

$$p_m(x_n) - x_n \ge p_n(x_n) - x_n \ge d_n - \delta_n \qquad \text{für } n \ge m.$$

Setzt man $x^* := \lim\limits_{n\to\infty} x_n$, so ergibt sich daraus für $n \to \infty$

$$p_m(x^*) - |x^*| \ge d^* - 0.$$

Führt man nun noch den Grenzübergang $m \to \infty$ aus, so findet man $0 \geq d^*$. Da d^* nicht negativ sein kann, muss $d^* = 0$ gelten. Damit ist gezeigt, dass $\lim_{n\to\infty} \sup_{x\in[-1,1]} \big|p_n(x) - |x|\big| = 0$ ist, d. h., dass die Folge (p_n) auf $[-1, 1]$ gleichmäßig gegen die Betragsfunktion konvergiert. Diese Feststellung beendet den dritten Beweis des Weierstraßschen Approximationssatzes.

Schlussbemerkung

Bisher wurde der Weierstraßsche Approximationssatz lediglich für auf dem Intervall $[0, 1]$ definierte stetige Funktionen bewiesen. Es lässt sich jedoch leicht zeigen, dass der Satz für alle stetigen Funktionen auf abgeschlossenen und beschränkten Intervallen gültig ist:

Eine Funktion $f \in C([x_0, x_1])$ kann durch Verschiebung und Streckung bzw. Stauchung des Arguments in eine Funktion aus $C([0, 1])$ überführt werden. Diese kann nach den oben geführten Beweisen durch ein Polynom auf $[0, 1]$ beliebig genau angenähert werden. Dieses Polynom kann durch Rückverschiebung und Stauchung bzw. Streckung wieder auf eine Funktion auf dem Intervall $[x_0, x_1]$ abgebildet werden; dabei entsteht ein Polynom auf $[x_0, x_1]$, das die Funktion f beliebig genau approximiert.