Grundlagen der Schulmathematik

Universität Tübingen
Sommersemester 2014

Frank Loose

20. Juni 2014

Vorwort

Mit dieser Vorlesung möchte ich, ganz im Sinne von *Felix Kleins* „Schulmathematik vom höheren Standpunkt“ die Grundlagen der Schulmathematik im Bereich der Algebra und der Analysis etwas tiefer gehend beleuchten als das im Schulunterricht möglich sein wird. Ich verknüpfe damit die Hoffnung, dass die Hintergründe der Schulthemen den zukünftigen Lehrerinnen und Lehrern bei der Ausübung ihres Berufes eine reiche Hilfe sein wird. Wenn sie freilich den Studierenden durch das Studium an sich gefallen, so soll mir das durchaus auch schon zum Gelingen des Projektes reichen.

Geplant ist in naher Zukunft ein zweiter Teil, der die schulrelevanten Themen im Bereich der Geometrie und Stochastik beleuchten soll.

Tübingen, im Frühjahr 2014

Inhaltsverzeichnis

6 Reelle Funktionen einer reellen Veränderlichen

(6.1) Wir wollen nun mit dem Studium von reellen Funktionen $f\colon I \to \mathbb{R}$ beginnen, wobei $I \subseteq \mathbb{R}$ meist ein *abgeschlossenes Intervall* ist, d.i.

$$I = [a,b],\ (-\infty,b],\ [a,\infty),\ (-\infty,\infty) = \mathbb{R}$$

(mit $a,b \in \mathbb{R}$, $a \leq b$, wobei wir den entarteten Fall $a = b$ ausschließen wollen, weil er langweilig ist). Später, vor allem in der Differentialrechnung, wird $I \subseteq \mathbb{R}$ häufig ein *offenes Intervall* sein, d.i.

$$I = (a,b),\ (-\infty,b),\ (a,\infty),\ (-\infty,\infty) = \mathbb{R}.$$

Manchmal werden aus I endlich viele Punkte herausgenommen, vor allem dann, wenn es um rationale Funktionen geht, wo der Nenner an endlich vielen Stellen gleich Null werden kann.

Beginnen wir mit den einfachsten Funktionen, die sich direkt aus der Körperstruktur von $\mathbb{R}$ ergeben, die man also bei jedem Körper K hat. Da sind zunächst die *konstanten Funktionen*, d.i. $f\colon \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = c, \quad \forall x \in \mathbb{R},$$

und einem festen $c \in \mathbb{R}$, insbesondere die *Nullfunktion*, wenn $c = 0$ ist. Dann gibt es die *Identität* $f = \mathrm{id}\colon \mathbb{R} \to \mathbb{R}$,

$$\mathrm{id}(x) = x, \quad \forall x \in \mathbb{R},$$

die *linearen Funktionen* $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) = ax, \quad \forall x \in \mathbb{R},$$

bei einem festen $a \in \mathbb{R}$, welche die Nullfunktion und die Identität als Spezialfälle (nämlich für $a = 0$ bzw. $a = 1$) enthalten. Der nächste Schritt sind die *affin-linearen Funktionen*, d.i. $f\colon \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = ax + b, \quad \forall x \in \mathbb{R},$$

bei festen Wahlen $a,b \in \mathbb{R}$, die alle bisher vorgestellten Funktionen als Spezialfall enthalten. Schließlich gibt es die so genannten *Polynomfunktionen* oder auch *ganzrationalen Funktionen*, d.i. ein $f\colon \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = a_n x^n + \cdots + a_1 x + a_0, \quad \forall x \in \mathbb{R},$$

bei Vorgabe von $n \in \mathbb{N}$ und $a_0, \dots, a_n \in \mathbb{R}$. Wir haben in §4 bereits diskutiert, wie man diese Funktionen dadurch bekommt, dass man in die Polynome $p \in \mathbb{R}[X]$ jeweils alle reellen Zahlen einsetzt,

$$\Phi: \mathbb{R}[X] \to \mathrm{Abb}(\mathbb{R}, \mathbb{R}),\ p \mapsto f_p.$$

Ganz am Ende von §4 haben wir auch den Quotientenkörper von $\mathbb{R}[X]$, notiert mit $\mathbb{R}(X)$, vorgestellt, dessen Elemente durch *Brüche* (d.s. Äquivalenzklassen bzgl. einer naheliegenden Äquivalenzrelation auf $\mathbb{R}[X] \times (\mathbb{R}[X] \setminus \{0\})$, ähnlich wie bei den rationalen Zahlen) $\frac{p}{q}$ mit $p \in \mathbb{R}[X]$ und $q \in \mathbb{R}[X] \setminus \{0\}$ gegeben sind. Sie geben Anlass zu den so genannten *rationalen Funktionen* $f: I \to \mathbb{R}$, gegeben durch

$$f(x) = \frac{p(x)}{q(x)}, \quad \forall x \in I,$$

wobei $p, q: \mathbb{R} \to \mathbb{R}$ Polynomfunktionen sind und q nicht die Nullfunktion ist, $q \not\equiv 0$. (Diese Notation bedeutet, dass q nicht *identisch Null* ist, wohl aber Nullstellen haben kann.) Der Definitionsbereich $I \subseteq \mathbb{R}$ ist dann

$$I = \mathbb{R} \setminus \{x \in \mathbb{R} :\ q(x) = 0\},$$

d.h.: aus $\mathbb{R}$ müssen die maximal n Nullstellen von q, wenn $n = \deg(q)$ ist, herausgenommen werden. Hier wollen wir annehmen, dass p und q keine gemeinsamen linearen Faktoren haben oder sogar vollständig gekürzt sind. Man erinnere sich dabei an die Bemerkung in §4, dass auch $\mathbb{R}[X]$ faktoriell ist.)

Beachte, dass wir von dem *Grad einer Polynomfunktion* sprechen können, da es wegen der Unendlichkeit von $\mathbb{R}$ zu q genau ein $\tilde{q} \in \mathbb{R}[X]$ gibt mit $f_{\tilde{q}} = q$ und den Grad von q definieren wir dann als den Grad von $\tilde{q}$.

Diese rationalen Funktionen auf $\mathbb{R}$, die wir nun auch mit $\mathbb{R}(X)$ bezeichnen (weil die entsprechende Zuordnung von $\mathbb{R}(X)$ in die reellen Funktionen, die außerhalb einer endlichen Teilmenge von $\mathbb{R}$ definiert sind, injektiv ist) bilden einen gewissen Abschluss gegenüber dem, was man an Funktionen auf $\mathbb{R}$ bekommen kann, wenn man nur die *Grundrechenarten*, d.h. die Körperverknüpfungen von $\mathbb{R}$ verwendet. Wir wollen hier auch noch vermerken, dass man aus $\mathbb{R}(X)$ nicht nur keine neuen Funktionen bekommt, wenn man zwei Elemente f, g miteinander addiert, multipliziert, sie voneinander subtrahiert oder sie durcheinander dividiert ($g \not\equiv 0$), sondern auch, wenn man sie miteinander *verkettet*, also $g \circ f: I \to \mathbb{R}$ betrachtet ($I \subseteq \mathbb{R}$ geeignet) mit

$$g \circ f(x) = g(f(x)).$$

(6.2) Der erste Prozess, der in natürlicher Weise aus dieser Funktionenklasse herausführt und damit unseren Funktionenvorrat vergrößern wird, und den wir hier vorstellen wollen, ist der Übergang zu *Umkehrfunktionen*. Man erinnere sich daran, dass für den Fall, wo wir eine Funktion $f: I \to J$ haben – sagen wir, wo $I, J \subseteq \mathbb{R}$ Intervalle sind – die bijektiv ist, wir in natürlicher Weise die Umkehrfunktion g haben, also jene eindeutig bestimmte Funktion $g: J \to I$, die

$$g \circ f = \mathrm{id}_I \quad \text{und} \quad f \circ g = \mathrm{id}_J$$

erfüllt. Es ist dann also g durch die Bedingung gegeben, dass für alle $x \in I$ und $y \in J$ gilt:

$$y = f(x) \Longleftrightarrow g(y) = x.$$

Beispiel 6.1 Die Funktion $f: [0, \infty) \to [0, \infty)$, $f(x) = x^2$, ist zunächst injektiv, weil aus $x_1^2 = x_2^2$ und $x_1, x_2 \geq 0$ aus

$$0 = x_2^2 - x_1^2 = (x_2 + x_1)(x_2 - x_1)$$

folgt, dass $x_1 = x_2$ sein muss. Sie ist auch surjektiv, weil mit der babylonischen Folge (siehe (3.5), mit einem beliebigen $c \geq 0$ statt $c = 2$), dass es für jedes $c \geq 0$ ein $x \geq 0$ gibt mit $x^2 = c$. Es ist also f bijektiv und wir haben somit $g := f^{-1}$, welches wir die *Quadratwurzelfunktion* nennen und mit

$$\sqrt{\cdot} := g: [0, \infty) \to [0, \infty)$$

notieren. Es ist also defnitionsgemäß für $x, y \geq 0$:

$$x = \sqrt{y} \;\Leftrightarrow\; x^2 = y.$$

Obwohl natürlich f eine rationale Funktion ist, ist es $g = \sqrt{\cdot}$ nicht mehr, wie man sich überlegen kann (Übung).

Diese Konstruktion möchte man nun für möglichst viele Funktionen machen können, z.B. geeigneten Einschränkungen von rationalen Funktionen, wie

$$f = \mathrm{pot}_n: [0, \infty) \to \mathbb{R},\; f(x) = x^n,$$

für $n \in \mathbb{N}_+$.Während man die Injektivität versucht ähnlich wie bei $f = \mathrm{pot}_2$ in den Griff zu bekommen, benötigt man bei der Surjektivität ein grundsätzlicheres Argument, denn man kann kaum für jede Funktion und jedes c aus dem (potentiellen) Bildbereich eine neue babylonische Folge konstruieren, die gegen das Urbild geht. Überhaupt ist zunächst keineswegs klar, dass – sagen wir – bei einer bijektiven Funktion $f\colon [a, b] \to J \subseteq \mathbb{R}$ das Bild wieder ein (abgeschlossenes) Intervall ist (was i.a. auch falsch ist).

Ein wichtiges Instrument bei dieser Frage wird uns nun dann zur Verfügung gestellt, wenn f eine *stetige Funktion* ist. Man erinnere sich dazu an folgende Definition.

Definition 6.2 Sei $f\colon I \to \mathbb{R}$ eine Funktion, $I \subseteq \mathbb{R}$ ein Intervall und $a \in I$. Dann sagen wir, dass f *für x gegen a gegen $c \in \mathbb{R}$ konvergiert*, und schreiben

$$\lim_{x \to a} f(x) = c,$$

wenn es für alle $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in I$ mit $0 < |x - a| < \delta$ gilt:

$$|f(x) - c| < \varepsilon.$$

Man beachte, dass der Limes, wenn er denn existiert, auch eindeutig bestimmt ist, sobald das Intervall nicht-entartet ist. (Im entarteten Fall ist *jedes* $c \in \mathbb{R}$ Grenzwert.)

Eine äquivalente Formulierung wäre folgende (dessen Äquivalenz wir als Übung überlassen): Für alle Folgen (x_n) in $I \setminus \{a\}$ mit $\lim(x_n) = a$ gilt: $\lim f(x_n) = c$ (wobei der Grenzwertbegriff für Folgen ja schon in §3 eingeführt wurde).

Mit Hilfe dieses Grenzwertbegriffes für Funktionen, den wir später (siehe §**??**) auch noch für die Differenzierbarkeit heranziehen werden, kann man Stetigkeit nun so definieren:

Definition 6.3 Sei $I \subseteq \mathbb{R}$ ein Intervall und $f\colon I \to \mathbb{R}$ eine Funktion.

(a) Für $a \in I$ heißt f *stetig in a*, wenn gilt:

$$\lim_{x \to a} f(x) = f(a),$$

(b) f heißt *stetig (schlechthin)*, wenn f stetig ist in a, für alle $a \in I$.

Eine Funktion $f\colon I \to \mathbb{R}$ ist damit genau dann stetig in $a \in I$, wenn für alle Folgen (x_n) in I (wir brauchen hier a nicht herausnehmen) mit $(x_n) \to a$ gilt: $(f(x_n)) \to f(a)$.

Alle rationalen Funktionen auf $\mathbb{R}$ sind stetig. (Da, wo sie definiert sind! Die Frage, ob z.B. $f\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $x \mapsto 1/x$, stetig in $x = 0$ ist, wird gar nicht gestellt, da $x = 0$ nicht zum Definitionsbereich von f gehört.) Das kann man z.B. so sehen, dass man sich dies für die konstanten Funktionen und die Identität unmittelbar aus der Definition heraus klar macht, was einfach ist. Dann beweist man allgemeine Sätze von folgender Bauart: Seien $f, g\colon I \to \mathbb{R}$ stetig in $a \in I$ und $\lambda \in \mathbb{R}$. Dann gilt:

(i) Auch $f + g\colon I \to \mathbb{R}$ ist stetig in a;

(ii) auch $\lambda f\colon I \to \mathbb{R}$ ist stetig in a;

(iii) auch $f \cdot g\colon I \to \mathbb{R}$ ist stetig in a;

(iv) ist $g(a) \neq 0$, so ist $f/g\colon J \to \mathbb{R}$ (mit $J := I \setminus \{x \in I : \ g(x) = 0\}$) stetig in a.

Nützlich ist zudem: Ist $f\colon I \to \mathbb{R}$ stetig in $a \in I$, $g\colon J \to \mathbb{R}$ stetig in $b \in J$, $f(I) \subseteq J$ und $b = f(a)$, so ist auch $g \circ f\colon I \to \mathbb{R}$ stetig in a. Mit Hilfe dieser *Regeln* ist unmittelbar klar, dass alle rationalen Funktionen stetig sind.

Um nun an das Bild von bijektiven stetigen Funktionen heranzukommen, benutzt man folgenden wichtigen Satz über stetige Funktionen, der uns auch später noch sehr nützlich sein wird (z.B. in der Integrationstheorie in §**??**) und dessen Kraft wir auch in §4 (siehe (4.11)) schon angedeutet haben:

Zwischenwertsatz 6.4 *Sei $f\colon [a,b] \to \mathbb{R}$ stetig mit $f(a) < 0 < f(b)$. Dann gibt es ein $\xi \in (a,b)$ („zwischen a und b“), so dass gilt: $f(\xi) = 0$.*

Man beachte, dass dieser Satz starken Gebrauch von der Vollständigkeit der reellen Zahlen macht, ja (in einem zu präzisierenden Sinne) zur Vollständigkeit von $\mathbb{R}$ äquivalent ist. Über den rationalen Zahlen würde er nicht gelten, wie z.B. die rationale Funktion $f\colon [1,2] \cap \mathbb{Q} \to \mathbb{Q}$, $f(x) = x^2 - 2$ deutlich macht. (Die Definition der Stetigkeit könnte man freilich, so wie sie ist, übertragen.)

Zum Beweis macht man eine *Intervallschachtelung* $(I_n)_{n\in\mathbb{N}}$, d.h. $I_{n+1} \subseteq I_n$, $\forall n \in \mathbb{N}$, und $\lim(b_n - a_n) = 0$, wenn $I_n = [a_n, b_n]$ ist. Die Vollständigkeit von $\mathbb{R}$ impliziert dann, dass (I_n) einen *Kern* hat, d.h.: es gibt ein $\xi \in \mathbb{R}$, so dass $\xi \in I_n$ ist, $\forall n \in \mathbb{N}$. Es ist nämlich dann sowohl (a_n) als auch (b_n) eine Cauchy-Folge und sie haben den gleichen Grenzwert, nämlich ξ. Die Formulierung, dass jede Intervallschachtelung einen Kern hat, ist übrigens äquivalent zum Vollständigkeitsaxiom (Übung).

Beweis von 6.4. Man halbiert nun nämlich einfach das Intervall $I_0 := [a,b]$ und setzt dann

$$a_1 := \frac{1}{2}(a+b),\ b_1 := b, \quad \text{falls } f(\tfrac{1}{2}(a+b)) < 0$$

ist und

$$a_1 := a,\ b_1 := \frac{1}{2}(a+b), \quad \text{falls } f(\tfrac{1}{2}(a+b)) > 0$$

ist. Falls$f(\frac{1}{2}(a+b)) = 0$ ist, ist man bei $\xi := \frac{1}{2}(a+b)$ sowieso fertig. Und so fährt man fort: Es ist also dann stets $f(a_n) < 0$ und $f(b_n) > 0$, falls man nicht sowieso in der Folge (a_n) bzw. (b_n) eine Nullstelle von f findet. Da nun $\xi = \lim(a_n)$ ist, folgt aus der Stetigkeit von f (in ξ), dass

$$f(\xi) = f(\lim(a_n)) = \lim \underbrace{f(a_n)}_{<0} \leq 0$$

ist und wegen $\xi = \lim(b_n)$ aber ebenso

$$f(\xi) = \lim \underbrace{f(b_n)}_{>0} \geq 0.$$

Es ist also $f(\xi) = 0$. □

Abbildung 9: zum Beweis des Zwischenwertsatzes

(6.3) Als Konsequenz daraus sieht man unmittelbar, dass ein stetiges $f\colon [a,b] \to \mathbb{R}$ mit $f(a) < 0$ und $f(b) > 0$ nicht nur eine Nullstelle hat, sondern dass jeder *Zwischenwert* $c \in [f(a), f(b)]$ von f angenommen wird, in dem man nämlich den Zwischenwertsatz auf die Funktion $g = f - c$ anwendet.

Als nächstes sieht man, dass ein injektives und stetiges $f\colon [a,b] \to \mathbb{R}$ notwendig *streng monoton* sein muss. Es heißt f *streng monoton wachsend*, wenn aus $x_1 < x_2$ stets $f(x_1) < f(x_2)$ folgt, für alle $a \leq x_1 < x_2 \leq b$ (und entsprechend *streng monoton fallend*). Nehmen wir z.B. $f(a) < f(b)$ an. Wäre f nicht streng monoton wachsend, so würde mindestens einer der folgenden drei Fälle auftreten:

(i) $\exists x \in [a,b] : \; f(x) < f(a)$,

(ii) $\exists y \in [a,b] : \; f(y) > f(b)$,

(iii) $\exists a < x < y < b :$
$$f(a) < f(y) < f(x) < f(b).$$

Dann würden aber im Fall (i) nach dem Zwischenwertsatz für $c \in (f(x), f(a)$ sowohl ein $\xi_1 \in (a, x)$ als auch ein $\xi_2 \in (x, b)$ existieren mit

$$f(\xi_1) = f(\xi_2) = c$$

und ähnlich für ein $c \in (f(b), f(y))$ im Fall (ii). Auch im Fall (iii) gibt es nach dem Zwischenwertsatz für $c \in (f(y), f(x))$ ein $\xi_1 \in (a, x)$ und ein $\xi_2 \in (y, b)$ mit

$$f(\xi_1) = f(\xi_2) = c$$

(und sogar noch mindestens einen dritten Punkt $\xi_3 \in (x, y)$ mit $f(\xi_3) = c$). Aber das geht wegen der Injektivität nicht.

Abbildung 10: zur Monotonie einer injektiven, stetigen Funktion

Und jetzt haben wir, auf was wir aus waren: Ist $f\colon [a, b] \to \mathbb{R}$ injektiv, und sagen wir $f(a) < f(b)$ (und ähnlich argumentiert man im monoton fallenden Fall), dann ist

$$\mathrm{im}(f) = [f(a), f(b)],$$

d.h.: jeder Zwischenwert $c \in [f(a), f(b)]$ wird von f getroffen.

Abbildung 11: zur Surjektivität einer injektiven, stetigen Funktion

Setzt man $J := [f(a), f(b)|$, so ist also $f\colon I \to J$ bijektiv und die Umkehrfunktion f^{-1} tatsächlich wieder auf einem Intervall definiert.

Beispiel 6.5 Betrachte für jedes $n \in \mathbb{N}_+$ die *Potenzfunktion* (oder *Monomfunktion*) $f = \mathrm{pot}_n \colon [0, \infty) \to \mathbb{R}$,

$$f(x) = x^n.$$

Es ist f dann tatsächlich injektiv, weil aus $f(x) = f(y)$ folgt:

$$0 = y^n - x^n = (y - x)(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1})$$

und der 2. Faktor kann wegen $x, y \geq 0$ nur verschwinden, wenn $x = y = 0$ ist. Ansonsten muss der 1. Faktor verschwinden, also $x = y$ sein. In jedem Fall ist $x = y$.

Wegen des obigen Zwischenwertargumentes sieht man nun aber auch, dass $\mathrm{im}(f) = [0, \infty)$ ist. Ist nämlich $c \geq 0$ beliebig, so wähle man zunächst ein $b > 0$ mit $b^n \geq c$, was geht, da $\lim_{x \to \infty} x^n = \infty$ ist (mit einer naheliegenden Erweiterung der Grenzwertdefinition von Folgen (3.3)). Dann gibt es ein (eindeutiges) $\xi \in [0, b]$ mit $f(\xi) = c$.

Wir können damit also nun die *n-te Wurzelfunktion*

$$\sqrt[n]{\cdot} := \mathrm{pot}_n^{-1} \colon [0, \infty) \to [0, \infty)$$

definieren, also

$$\sqrt[n]{y} = x \iff y = x^n, \quad \forall x, y \in [0, \infty).$$

(6.4) Es ist nun eine naheliegende Frage, ob die so konstruierten Umkehrfunktionen stetiger Funktionen $f \colon I \to J$ nicht auch wieder stetig sind. Ein weiteres Hilfsmittel zur Beantwortung dieser Frage – und auch einiger, die noch folgen – wollen wir hier abschließend vorstellen. Wir nennen dabei $\xi \in \mathbb{R}$ einen *Häufungspunkt (H.P.) einer Folge* (x_n) in $\mathbb{R}$, wenn es für jedes $\varepsilon > 0$ unendlich viele Folgenglieder gibt, die ins Intervall $(\xi - \varepsilon, \xi + \varepsilon)$ fallen. (Aber nicht unbedingt *fast alle*, wenn ξ sogar Grenzwert ist.) Der folgende Satz drückt nun wiederum die Vollständigkeit der reellen Zahlen aus, ja ist äquivalent dazu, und wird mit einer ähnlichen Intervallhalbierungsmethode bewiesen wie der Zwischenwertsatz.

Satz 6.6 (von *Bolzano-Weierstraß*). *Sei (x_n) eine Folge reeller Zahlen im Intervall $[a, b] \subseteq \mathbb{R}$. Dann hat (x_n) einen Häufungspunkt in $[a, b]$.*

Beweis. Man halbiere das Intervall $I_0 = [a, b]$ wieder sukzessive und entscheide sich bei jedem Schritt für eines der beiden Teilintervalle, in dem unendlich viele Folgenglieder liegen. Der Kern $\xi \in [a, b]$ dieser Intervallschachtelung (I_n) ist dann ein H.P. von (x_n). $\square$

Damit kann man nun sehen, dass Umkehrfunktionen stetiger Funktionen $f\colon I \to J$ wieder stetig sein müssen. Nehmen wir dazu ein $b \in J$ beliebig und (y_n) eine Folge in J mit $(y_n) \to b$. Sei (x_n) die Bildfolge unter $g := f^{-1}$ und $a := g(b)$. Wir müssen zeigen, dass $(x_n) \to a$, weil das

$$\lim_{y \to b} g(y) = g(b)$$

bedeutet. Aber (x_n) hat nach (6.6) einen H.P., sagen wir $\xi \in [a, b]$. Es gibt dann eine *Teilfolge* $(x_{n_k})_{k\in\mathbb{N}}$ von (x_n), die gegen ξ konvergiert und wegen der Stetigkeit von f (in ξ) folgt dann

$$f(\xi) = f(\lim_{k\to\infty}(x_{n_k})) = \lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty}(y_{n_k}) = b$$

und damit, wegen der Injektivität von f: $\xi = a$. Das Argument zeigt aber nun, dass *jeder* H.P. von (x_n) die Zahl a ist, (x_n) also nur *einen* H.P. hat. Dann kann man aber, wiederum mit (6.6) sehen, dass a sogar Grenzwert von (x_n) sein muss. Gäbe es nämlich ein $\varepsilon > 0$, so dass in $(a - \varepsilon, a + \varepsilon)$ nicht fast alle Folgenglieder liegen, so müsste die (unendliche) Teilfolge derer, die in $I \setminus (a - \varepsilon, a + \varepsilon)$ liegen, wieder einen H.P. haben, der nicht a sein kann, sondern in $I \setminus (a - \varepsilon, a + \varepsilon)$ liegen muss. Dieser wäre auch ein H.P. von (x_n), den es aber wie gesehen nicht gibt.

Wir wollen abschließend eine weitere wichtige Konsequenz aus (6.6) erwähnen, die insbesondere bei der Erschließung der Differentialrechnung (siehe §**??**) wichtig sein wird.

Satz 6.7 (von *Weierstraß*). *Sei $f\colon [a, b] \to \mathbb{R}$ stetig. Dann nimmt f sein Supremum an.*

Eine Teilmenge $M \subseteq \mathbb{R}$ der reellen Zahlen heißt *nach oben beschränkt*, wenn es ein $c \in \mathbb{R}$ gibt, so dass $x \leq c$ ist, für alle $x \in M$. Jede solche Zahl $c \in \mathbb{R}$ nennen wir eine *obere Schranke von M*. Eine obere Schranke $c \in \mathbb{R}$ heißt *Supremum von M* (oder *kleinste obere Schranke*), wenn gilt: Ist $b \in \mathbb{R}$ eine weitere obere Schranke von M, so gilt: $c \leq b$. Eine solche kleinste obere Schranke ist, wenn sie existiert, eindeutig bestimmt. Wir schreiben dann $c =: \sup(M)$ dafür. Die Tatsache, dass jede nicht-leere, nach oben beschränkte Teilmenge $M \subseteq \mathbb{R}$ ein Supremum hat, folgt wiederum aus dem Vollständigkeitsaxiom. Sie ist sogar äquivalent zum Vollständigkeitsaxiom gemeinsam mit dem Archimedischen Axiom.

Übrigens setzt man für nicht nach oben beschränkte Teilmengen $M \subseteq \mathbb{R}$: $\sup(M) := \infty$ und für die leere Menge setzt man $\sup \emptyset := -\infty$, so dass man dann für *jede* Teilmenge $M \subseteq \mathbb{R}$ ihr Supremum $\sup(M) \in [-\infty, +\infty]$ definiert hat.

Setzt man also

$$c := \sup(\mathrm{im} f) \in (-\infty, \infty],$$

so ist also zunächst noch $c = \infty$ möglich. Man weiß also hier noch nicht, dass f *nach oben beschränkt* ist,, d.h.: $\mathrm{im}(f) \subseteq \mathbb{R}$ ist nach oben beschränkt. Satz (6.7) behauptet also insbesondere, dass f nach oben beschränkt ist, aber zudem, dass es ein $\xi \in [a, b]$ gibt mit $f(\xi) = \sup(f) := \sup(\mathrm{im} f)$.

Man nennt für eine Teilmenge $M \subseteq \mathbb{R}$ eine Zahl $c \in \mathbb{R}$ *Maximum von* M, wenn c obere Schranke für M ist und $c \in M$ ist. Natürlich ist c dann wieder eindeutig bestimmt und wir schreiben dann $c =: \max(M)$. Man beachte aber, dass ein Maximum einer Teilmenge M, auch wenn sie nach oben beschränkt ist, nicht immer zu existieren braucht, z.B. für $M = \{x \in \mathbb{R} : x < 0\}$ ist das der Fall. Wenn es existiert, gilt natürlich stets:

$$\max(M) = \sup(M).$$

Den Satz von Weierstraß (6.7) kann man dann auch einfach so ausdrücken, dass das Bild von f ein Maximum hat.

Man beachte schließlich, dass die Stelle $\xi \in [a, b]$, wo f sein Maximum annimmt, natürlich keineswegs eindeutig bestimmt sein muss. Man sollte diese Stellen vielleicht als *Maximumsstellen von* f bezeichnen, bezeichnet sie allerdings häufig auch (etwas verwirrenderweise) als ein Maximum von f.

Wenn eine Teilmenge $M \subseteq \mathbb{R}$ i.a. nun kein Maximum hat, so gibt es aber direkt aus der Definition des Supremums von M wenigstens eine Folge (y_n) in M mit $(y_n) \to \sup(M)$ (mit der naheliegenden Erweiterung von Definition (3.3) für den Fall $\lim(y_n) = \infty$), denn wenn schon das Supremum von M nicht in M zu liegen braucht, so gibt es doch beliebig nahe darunter Elemente in M.

Beweis. Setzen wir nun also

$$c := \sup(f) \in (-\infty, \infty],$$

so wissen wir wenigstens, dass es eine Folge (x_n) in $[a, b]$ gibt mit $(f(x_n)) \to c$. (Eine solche Folge nennen wir eine *Maximalfolge von* f.) Nach Übergang zu einer Teilfolge von (x_n), dürfen wir nach Satz (6.6) sogar annehmen, dass (x_n) gegen ein $\xi \in [a, b]$ konvergiert. (Natürlich gilt das nur für eine Teilfolge, aber diese Teilfolge notieren wir einfach im Folgenden wieder mit (x_n).) Aber dann ist

$$c = \lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty})(x_n) = f(\xi),$$

wegen der Stetigkeit von f (in ξ). Es folgt: $c < \infty$ und wird tatsächlich angenommen (nämlich in ξ). $\square$

Man beachte auch noch, dass wir nun unseren Funktionenvorrat nicht etwa nur um die n-ten Wurzelfunktionen ($n \in \mathbb{N}_+$) erweitert haben, sondern dass durch die algebraischen Operationen und Verkettungen auch Funktionen wie z.B. $f\colon \mathbb{R} \to \mathbb{R}$,

$$f(x) = \sqrt{1 + x^2},$$

zugänglich geworden sind, und dass diese allesamt stetig sind.

Literatur

[1] M. Aigner und G. Ziegler: *Das BUCH der Beweise.* Springer-Verlag, Berlin, Heidelberg

[2] S. Bosch: *Algebra.* Springer-Verlag, Berlin, Heidelberg, New York

[3] H.-D. Ebbinghaus et al.: *Zahlen.* Springer-Verlag, Berlin, Heidelberg, New York, Tokio

[4] G. Fischer: *Lehrbuch der Algebra.* Friedr. Vieweg & Sohn Verlag, Wiesbaden

[5] O. Forster: *Analysis I.* Friedr. Vieweg Verlag, Braunschweig, Wiesbaden

[6] G. Frey: *Elementare Zahlentheorie.* Friedr. Vieweg & Sohn Verlag, Braunschweig, Wiesbaden

[7] U. Friedrichsdorf und A. Prestel: *Mengenlehre für den Mathematiker.* Friedr. Vieweg & Sohn Verlag, Braunschweig, Wiesbaden

[8] P. Halmos: *Naive Mengenlehre.* Vandenhoeck & Ruprecht Verlag, Göttingen

[9] M. Spivak: *Calculus.* Cambridge University Press, Cambridge

[10] U. Storch und H. Wiebe: *Lehrbuch der Mathematik, Band I: Analysis einer Veränderlichen.* Springer Verlag, Berlin, Heidelberg

Abbildungsverzeichnis

Index