Mathematik für Informatiker II (Analysis)

J. Schicho et alias

Id: skriptum.tex,v 1.3 2002/04/29 12:38:49 schicho Exp

Kapitel 1

Symmetrien

1.1 In der Ebene

1.2 Im Raum

1.2.1 Symmetrien 1. Art

1.2.2 Symmetrien 2. Art

1.3 Symmetrien von Graphen

Kapitel 2

Reelle Zahlen

Das Kontinuum, oder der Raum der reellen Zahlen, ist von fundamentaler Bedeutung für die Analysis. Die Vorstellung, die mit einer reellen Zahl verbunden ist, ist die einer meßbaren Größe, z. B. die Länge. Eine reelle Zahl kann vorgestellt worden als das Verhältnis zweier Größen, nämlich der zu messenden Größe und der Maßeinheit (z.B. “Meter”).

Zur mathematischen Erfaßung des Begriffes der reellen Zahlen ist es notwendig, Axiome anzugeben, die von den reellen Zahlen erfüllt werden. Jede weitere mathematische Aussage über das Kontinuum von diesen Axiomen herzuleiten.

Axiomensysteme für die reellen Zahlen waren bereits in der Antike bekannt. Die am schwierigsten zu erfassende Eigenschaft, nämlich die Vollständigkeit des Kontinuums, wurde im 19. Jahrhundert von Dedekind in bisher nicht dagewesener Klarheit begriffen.

2.1 Axiome von $\mathbb{R}$

In der ersten Gruppe von Axiomen werden die Eigenschaften festgehalten, die zum Rechnen mit reellen Zahlen benötigt werden.

2.1.1 Axiom (Körperaxiome)

$$\begin{array}{rrrr} a+(b+c)=(a+b)+c, & a+b=b+a, & a+0=a, & a+(-a)=0; \\ a\cdot(b\cdot c)=(a\cdot b)\cdot c, & a\cdot b=b\cdot a, & a\cdot 1=a, & a\neq 0 \implies a\cdot a^{-1}=1; \\ a\cdot(b+c)=a\cdot b+a\cdot c. & & & \end{array}$$

In der zweiten Gruppe von Axiomen werden die Eigenschaften festgehalten, die Vergleiche von reellen Zahlen betreffen.

2.1.2 Axiom (Ordnungsaxiome)

$$\begin{aligned}
&a \leq a, \quad a \leq b \wedge b \leq a \implies a = b, \quad a \leq b \wedge b \leq c \implies a \leq c;\\
&a \leq b \vee b \leq a;\\
&a \leq b \implies a + c \leq b + c,\\
&0 \leq a \wedge 0 \leq b \implies 0 \leq a \cdot b.
\end{aligned}$$

2.1.3 Definition Ein ***geordneter Körper*** ist eine Struktur, in der sowohl die Körperaxiome als auch die Ordnungsaxiome gelten.

2.1.4 Beispiel Die Menge $\mathbb{Q}$ der rationalen Zahlen ist ein geordneter Körper.

2.1.5 Bemerkung Die Menge $\mathbb{D}$ der Dezimalzahlen bildet nur einen geordneten Ring: Es gelten alle Regeln für Addition, Subtraktion, und Multiplikation, aber nicht jedes Element, das von 0 verschieden ist, hat ein Inverses bezüglich der Multiplikation.

Die Menge $\mathbb{Q}$ der rationalen Zahlen ist "unvollständig": es befinden sich "zwischen" den rationalen Zahlen Löcher, entsprechend den nichtrationalen Zahlen (z.B.$\sqrt{2}$). Diesen nichtrationalen Zahlen kann man sich beliebig knapp nähern, sie aber nicht erreichen ohne den Bereich $\mathbb{Q}$ zu verlassen. Um solche Löcher aufzufinden bzw. mathematisch dingfest zu machen, wird ein Grenzprozeß benötigt. Man kann entweder Cauchy-Folgen oder Intervallschachtelungen verwenden.

2.1.6 Definition Sei $(a_n)_n$ eine Folge von reellen Zahlen. Dann sagt man, daß diese Folge gegen einen ***Grenzwert*** $a \in \mathbb{R}$ ***konvergiert***, und schreibt $\lim_{n\to\infty} a_n = a$, falls es für alle $\epsilon > 0$ ein N gibt, sodaß $|a_{N+m} - a| < \epsilon$, für alle m.

Die Folge heißt eine ***Cauchy-Folge*** falls es für alle $\epsilon > 0$ ein N gibt, sodaß $|a_{N+m} - a_{N+n}| < \epsilon$, für alle m und n.

2.1.7 Beispiel Die ersten Glieder der durch $a_0 := 1$ und $a_{n+1} := \frac{a_n + \frac{2}{a_n}}{2}$ rekursiv definierten Folge sind:

$$\begin{aligned}
a_0 &= 1.00000000000000000000\ldots\\
a_1 &= 1.50000000000000000000\ldots\\
a_2 &= 1.46666666666666666666\ldots\\
a_3 &= 1.41421568627450980392\ldots\\
a_4 &= 1.41421356237468991062\ldots\\
a_5 &= 1.41421356237309504880\ldots\\
a_6 &= 1.41421356237309504880\ldots\\
a_7 &= 1.41421356237309504880\ldots\\
&\qquad\qquad\qquad\qquad\quad\ldots
\end{aligned}$$

Offensichtlich ändern sich vom 5-ten Glied an die ersten 20 Stellen hinter dem Komma nicht mehr. Daher gilt für $\epsilon = 10^{-20}$ und $N = 5$ die Ungleichung $|a_{N+m} - a_{N+n}| \leq \epsilon$ für alle m, n.

Eine ähnliche Beobachtung läßt sich nicht nur für $\epsilon = 10^{-20}$ machen, sondern auch für $\epsilon = 10^{-50}$, $\epsilon = 10^{-100}$, überhaupt für jedes noch so kleine positive ϵ. Die vorliegende Folge ist daher eine Cauchy-Folge.

Jede konvergente Folge ist automatisch eine Cauchy-Folge. Umgekehrt ist es aber vorstellbar, daß der Grenzwert einer Cauchy-Folge fehlt, bzw. sich außerhalb der betrachteten Struktur befindet, diese also Löcher hat. Die reellen Zahlen sollten keine Löcher haber, also vollständig sein.

2.1.8 Axiom (Vollständigkeit) *Jede Cauchy-Folge von reellen Zahlen hat einen Grenzwert in $\mathbb{R}$.*

2.1.9 Beispiel Die durch $a_0 := 1$ und $a_{n+1} := \frac{a_n + \frac{2}{a_n}}{2}$ rekursiv definierte Folge ist eine Cauchy-Folge von rationalen Zahlen. Der Grenzwert $\sqrt{2}$ befindet sich aber außerhalb der rationalen Zahlen. Daher ist $\mathbb{Q}$ nicht vollständig.

2.1.10 Definition Eine ***Intervallschachtelung*** ist eine Folge von Intervallen

$$I_n = [a_n, b_n] = \{\, x \mid a_n \leq x \leq b_n \,\},$$

sodaß gilt

$$\begin{aligned}
I_n &\neq \varnothing;\\
I_{n+1} &\subseteq I_n;\\
\lim_{n\to\infty} \text{länge}(I_n) = \lim_{n\to\infty} b_n - a_n &= 0.
\end{aligned}$$

2.1.11 Axiom (Intervallschachtelung) *Jede Invervallschachtelung hat einen innersten Punkt.*

2.1.12 Theorem *Das Axiom für Intervallschachtelungen ist eine äquivalente Formulierung des Vollständigkeitsaxioms.*

Beweis Es sind zwei Implikationen zu zeigen: Cauchy-Vollständigkeit (nämlich daß jede Cauchy-Folge einen Grenzwert hat) impliziert Intervallschachtelungs-Vollständigkeit (nämlich daß jede Intervallschachtelung einen innersten Punkt hat); und umgekehrt, Intervallschachtelungs-Vollständigkeit impliziert Cauchy-Vollständigkeit.

Die Idee für die erste Implikation ist die folgende. Ausgehend von einer beliebigen Intervallschachtelung $(I_n)_n$ konstruieren wir die Folge der Intervall-Mittelpunkte. Wir zeigen, daß diese Folge eine Cauchy-Folge ist. Mit der Annahme der Cauchy-Vollständigkeit können wir sagen, daß ein Grenzwert existiert. Wir zeigen daß dieser Grenzwert innerster Punkt von $(I_n)_n$ ist. Damit ist die IS-Vollständigkeit gezeigt.

Die Idee für die zweite Implikation ist die folgende. Ausgehend von einer beliebigen Cauchy-Folge $(a_n)_n$ konstruieren wir eine Folge von Intervallen, die die Cauchy-Folge einfangen (d.h. ab einem gewissen Index liegen alle Folgenglieder im Intervall). Wir zeigen, daß diese Folge eine Intervallschachtelung ist. Mit der Annahme der IS-Vollständigkeit können wir sagen, daß ein innerster Punkt existiert. Wir zeigen, daß dieser innerste Punkt Grenzwert von $(a_n)_n$ ist. Damit ist die Cauchy-Vollständigkeit gezeigt.

2.1.13 Bemerkung Typische Beweise in der Mathematik haben kreative Elemente (die "Beweisidee") und technische Elemente. Die technischen Aspekte stellen manchmal große Hürden für das Verstehen oder Darstellen eines Beweises dar, auch dann wenn die Beweisidee im Grunde klar ist. Sie gehorchen aber einfachen Beweisprinzipien und können durch Übung wie eine formale Sprache erlernt werden.

Der obige Beweis ist technisch außerordentlich anspruchsvoll. Daher eignet er sich sehr gut als Beispiel für die erwähnten Beweisprinzipien. Eine detaillierte Ausarbeitung des Beweises ist auf der Vorlesungs-Webseite http://www.risc.uni-linz.ac.at/courses/ss2002/math1/ zu finden.

Das letzte Axiom besagt, daß jede reelle Zahl beliebig genau durch rationale Zahlen approximiert werden kann.

2.1.14 Bemerkung Typische Beweise in der Mathematik haben kreative Elemente (die "Beweisidee") und technische Elemente. Die technischen Aspekte stellen manchmal große Hürden für das Verstehen oder Darstellen eines Beweises dar, auch dann wenn die Beweisidee im Grunde klar ist. Sie gehorchen aber einfachen Beweisprinzipien und können durch Übung wie eine formale Sprache erlernt werden.

Der obige Beweis ist technisch außerordentlich anspruchsvoll. Daher eignet er sich sehr gut als Beispiel für die erwähnten Beweisprinzipien. Eine detaillierte Ausarbeitung des Beweises ist auf der Vorlesungs-Webseite ttp://www.risc.uni-linz.ac.at/courses/ss2002/math1/ zu finden.

Das letzte Axiom besagt, daß jede reelle Zahl beliebig genau durch rationale Zahlen approximiert werden kann.
»»»> 1.3

2.1.15 Axiom (Archimedes) *Für jede reelle Zahl* α *und jede positive reelle Zahl* ϵ *gibt es eine rationale* ϵ*-Approximation* a, *d.h.* $|a - \alpha| \leq \epsilon$.

Auch zu diesem Axiom erwähnen wir zwei äquivalente Formulierungen:

2.1.16 Axiom (Archimedes′) *Für jede reelle Zahl* α *und jede positive reelle Zahl* ϵ *gibt es eine Dezimalzahl* a *mit* $|a - \alpha| \leq \epsilon$.

2.1.17 Axiom (Archimedes″) *Für jede reelle Zahl* α *gibt es eine natürliche Zahl* n *mit* $\alpha \leq n$.

2.2 Approximation mit rationalen Zahlen

Die folgende „gute“ dezimale Approximation

$$\operatorname{approx}(\alpha, n) := \frac{\operatorname{round}(10^n \cdot \alpha)}{10^n}$$

erfüllt

$$|\alpha - \operatorname{approx}(\alpha, n)| \leq \frac{1}{2 \cdot 10^n}.$$

2.2.1 Definition Eine Dezimal-Approximation mit Nenner 10^n und Fehler $\epsilon \leq \frac{1}{2 \cdot 10^n}$ heißt ***korrekte Approximation der Ordnung*** n,

Durch Wahl eines geeigneten Nenners kann man oft wesentlich besser approximieren:

$$\left|\pi - \frac{3141593}{1000000}\right| \approx 0.35 \cdot 10^{-6},$$

$$\left|\pi - \frac{355}{113}\right| \approx 0.27 \cdot 10^{-6}.$$

Wie gut kann man approximieren?

2.2.2 Satz *Zu jeder reellen Zahl* α *und natürlichen Zahl* n *gibt es eine rationale Approximation mit Nenner* $q \leq n$ *und Fehler* $\leq \frac{1}{q^n}$.

Wie kann man eine gute Approximation effizient berechnen?
$p_{-1} := 1$; $q_{-1} := 0$;
$\alpha_0 := \alpha$; $a_0 := \lfloor \alpha_0 \rfloor$; $p_0 := a_0$; $q_0 := 1$;
$i := 1$;

repeat
$\alpha_i := \frac{1}{\alpha_{i-1}-a_{i-1}}$;
$a_i := \lfloor \alpha_i \rfloor$;
$p_i := p_{i-1}a_i + p_{i-2}$;
$q_i := q_{i-1}a_i + q_{i-2}$;
$i := i + 1$;
until „Abbruchbedingung“

2.2.3 Beispiel Führt man den Algorithmus mit der Eingabe $\alpha = \pi = 3.1415926535$ durch, ergeben sich in den ersten vier Schritten die folgenden Werte:

i	α_i	a_i	p_i	q_i	p_i/q_i
-1	-	-	1	0	-
0	3.1415926535	3	3	1	3.0000000000
1	7.0625133059	7	22	7	3.1428571428
2	15.9965944066	15	333	106	3.1415094339
3	1.0034172310	1	355	113	3.1415929203
4	292.6345910143	292	103993	33102	3.1415926530

2.2.4 Satz $\left|\alpha - \frac{p_i}{q_i}\right| < \frac{1}{q_i q_{i+1}} < \frac{1}{q_i^2}$.

Beweis Für $i \geq 1$ gilt $\alpha_i \geq 1$, $a_i \geq 1$, $q_i > q_{i-1}$, $q_i > 0$. Der Wert des Ausdrucks $p_i q_{i-1} - p_{i-1} q_i$ ändert bei jeder Erhöhung von i um eins sein Vorzeichen, sein Betrag bleibt aber konstant; denn es gilt

$$p_{i+1}q_i - p_i q_{i+1} = (p_i a_{i+1} + p_{i-1})q_i - p_i(q_i a_{i+1} + q_{i-1}) = -(p_i q_{i-1} - p_{i-1} q_i).$$

Für $i = 0$ ist der Wert gleich -1, daher ist der Wert stets gleich $(-1)^{i+1}$. Daher ändert der Wert des Ausdrucks

$$\frac{p_i}{q_i} - \frac{p_{i-1}}{q_{i-1}} = \frac{p_i q_{i-1} - p_{i-1} q_i}{p_i q_i} = \frac{(-1)^{i+1}}{p_i q_i}$$

bei jeder Erhöhung das Vorzeichen, der Betrag wird jedoch bei jedem Schritt kleiner und konvergiert gegen Null. Es folgt, daß die Folge von Intervallen

$$\left(\left[\frac{p_0}{q_0}, \frac{p_1}{q_1}\right], \left[\frac{p_2}{q_2}, \frac{p_1}{q_1}\right], \left[\frac{p_2}{q_2}, \frac{p_3}{q_3}\right], \left[\frac{p_4}{q_4}, \frac{p_3}{q_3}\right], \ldots\right)$$

eine Intervallschachtelung ist. Wenn β der innerste Punkt ist, dann gilt für alle i, daß β zwischen $\frac{p_i}{q_i}$ und $\frac{p_{i+1}}{q_{i+1}}$ liegt, daher gilt

$$\left|\beta - \frac{p_i}{q_i}\right| \leq \left|\frac{p_i}{q_i} - \frac{p_{i+1}}{q_{i+1}}\right| = \frac{1}{q_i q_{i+1}} < \frac{1}{q_i^2}.$$

Wir behaupten, daß $\alpha = \beta$ gilt, daß also α der innerste Punkt ist. Dazu zeigen wir für jedes $i \geq 1$, daß α zwischen $\frac{p_i}{q_i}$ und $\frac{p_{i-1}}{q_{i-1}}$ liegt.

Der Wert des Ausdrucks $\frac{p_{i-1}\alpha_i+p_{i-2}}{q_{i-1}\alpha_i+q_{i-2}}$ ändert sich nicht, wenn i um eins erhöht wird, denn es gilt

$$\frac{p_i\alpha_{i+1}+p_{i-1}}{q_i\alpha_{i+1}+q_{i-1}} = \frac{(p_{i-1}a_i+p_{i-2})\frac{1}{\alpha_i-a_i}+p_{i-1}}{(q_{i-1}a_i+q_{i-2})\frac{1}{\alpha_i-a_i}+q_{i-1}} = \frac{p_{i-1}\alpha_i+p_{i-2}}{q_{i-1}\alpha_i+q_{i-2}}.$$

Für $i=0$ ist dieser Wert gleich α, also ist der für jedes i gleich α.

Allgemein kann gezeigt werden (Übung), daß $\frac{a+b}{c+d}$ zwischen $\frac{a}{c}$ und $\frac{b}{d}$ liegt. Daher liegt α zwischen $\frac{p_i\alpha_{i+1}}{q_i\alpha_{i+1}} = \frac{p_i}{q_i}$ und $\frac{p_{i-1}}{q_{i-1}}$. Damit ist alles gezeigt. □

Diese Art der Approximation heißt ***Kettenbruch***, weil gilt

$$\frac{p_i}{q_i} = a_o + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots \; a_{i-1} + \cfrac{1}{a_i}}}}}.$$

2.3 Darstellung reeller Zahlen im Computer

2.3.1 Variante 1: Unendliche Liste von Dezimalziffern

2.3.1 Bemerkung Unendliche Listen lassen sich mit Hilfe von funktionalen Programmiersprachen, definieren und manipulieren.

Die elementaren Operationen der funktionalen Programmiersprachen erfüllen die folgenden Gleichungen:

$$\mathit{head}[n:L] = n, \mathit{tail}[n:L] = L.$$

Die unten definierte Funktion *from* liefert dann die Liste aller natürlichen Zahlen $\geq n$. Sie ist durch eine unendliche Rekursion definiert.

$$\mathit{from}\, n := [n : \mathit{from}(n+1)]$$

Selbstverständlich ist es unmöglich, alle Elemente einer unendlich langen Liste anzuzeigen. Daher definieren wir eine Funktion *take*, welche die ersten n Elemente einer unendlich langen Liste L extrahiert.

$$\begin{aligned} \mathit{take}\, n\, L := \text{if } n = 0 \text{ then } & [] \\ \text{else } & [\mathit{head}\, L : \mathit{take}(n-1)(\mathit{tail}\, L)] \end{aligned}$$

Beispiel:

$$\mathit{take}\, 3\, \mathit{from}\, 4 \rightsquigarrow [4,5,6].$$

Damit kann man zwar die Dezimalzahlen, auch unendlich lange, im Computer darstellen, es gibt aber ein kleines technisches Problem: Die Arithmetik stürzt manchmal ab. Sei etwa $\alpha := [3 : \alpha]$, also die Dezimalentwicklung von $\frac{1}{3}$. Dann ist $3 \cdot \alpha$ nicht berechenbar, weil etwa zur Bestimmung von $(3 \cdot \alpha)$ die Kenntnis *aller* Dezimalen notwendig wäre, was aber unendlich viel Zeit benötigt. Man kann versuchen, das Problem zu ignorieren; das Problem tritt aber doch häufiger auf als man meinen könnte, zum Beispiel jedesmal wenn wir eine Zahl von sich selber abziehen. Oder man kann, wie folgt, eine kleine Modifiktion vornehmen.

2.3.2 Variante 2: Statt der Ziffern $0, \ldots, 9$, verwende $-5, \ldots, 5$

Es gilt dann beispielsweise

$$0.2\overline{2}2\overline{2} = 2 \cdot 10^{-1} + (-2) \cdot 10^{-2} + 2 \cdot 10^{-3} + (-2) \cdot 10^{-4}.$$

(Das Minuszeichen wird aus Gründen des Schriftbilds über der Ziffer geschrieben.) Da wir eine Ziffer mehr verwenden (nämlich 11 statt 10), kommt es dabei zu Mehrdeutigkeiten, etwa $0.35 = 0.4\overline{5}$ oder $0.444444\ldots = 1.\overline{555555}\ldots$. (Wir hatten jedoch auch schon bei Variante 1 Mehrdeutigkeiten, etwa $0.999999\ldots = 1.000000\ldots$.)

Dafür sind jetzt alle Grundrechnungsarten effektiv ausführbar.

Addition

Definiere zuerst die Ziffernoperationen

$$c_1(a,b) := \begin{cases} 0 & \text{falls } -5 \leq a+b \leq 4, \\ 1 & \text{falls } a+b \geq 5, \\ -1 & \text{falls } a+b \leq -6; \end{cases} \qquad c_2(a,b) := \begin{cases} 0 & \text{falls } -4 \leq a+b \leq 5, \\ 1 & \text{falls } a+b \geq 6, \\ -1 & \text{falls } a+b \leq -5; \end{cases}$$

$$s_1(a,b) := a+b-10 \cdot c_1(a,b), \qquad s_2(a,b) := a+b-10 \cdot c_2(a,b).$$

Offensichtlich liegt jeder Wert von s_1 zwischen -5 und 4, während jeder Wert von s_2 zwischen -4 und 5 liegt.

$$\begin{array}{rr} \text{Sei } \alpha \in \mathbb{R} \text{ gegeben durch} & a_{-n} \ldots a_{-1} a_0 . a_1 a_2 a_3 \ldots, \\ \beta \in \mathbb{R} \text{ gegeben durch} & b_{-n} \ldots b_{-1} b_0 . b_1 b_2 b_3 \ldots, \\ \text{dann ist } \alpha + \beta \text{ gegeben durch} & c_{-n-1} c_{-n} \ldots c_{-1} c_o . c_1 c_2 c_3 \ldots, \end{array}$$

wobei $c_i := s_1\big(s_2(a_i,b_i),\, c_2(a_{i+1},b_{i+1})\big) + c_1\big(s_2(a_{i+1},b_{i+1}),\, c_2(a_{i+2},b_{i+2})\big)$. Es ist nicht von vornherein klar, daß die c_i tatsächlich gültige Ziffern sind. Es könnte ja theoretisch sein, daß $s_1\big(s_2(a_i,b_i),\, c_2(a_{i+1},b_{i+1})\big) = 5$ und $c_1\big(s_2(a_{i+1},b_{i+1}),\, c_2(a_{i+2},b_{i+2})\big) = 1$, dann wäre die Summe 6, das ist keine Ziffer. Das ist aber nicht möglich, weil jeder Wert von s_1 zwischen -5 und 4 ist. Oder aber $s_1\big(s_2(a_i,b_i),\, c_2(a_{i+1},b_{i+1})\big) = -5$ und

$c_1\big(s_2(a_{i+1}, b_{i+1}),\, c_2(a_{i+2}, b_{i+2})\big) = -1$ (dann wäre die Summe gleich -6, d.h. keine Ziffer). Aber $c_1\big(s_2(a_{i+1}, b_{i+1}),\, c_2(a_{i+2}, b_{i+2})\big) = -1$ ist nur möglich, wenn $s_2(a_{i+1}, b_{i+1}) = -5$ und $c_2(a_{i+2}, b_{i+2}) = -1$ ist, und das ist wiederum ausgeschlossen, weil jeder Wert von s_2 zwischen -4 und 5 liegt.

Mit dieser Addition pflanzt sich also der Übertrag maximal bis zur zweiten Stelle fort.

Subtraktion

Wie in Addition (Multiplikation mit -1 ist trivial).

Multiplikation, Division

Gehen auch ähnlich, aber komplizierter.
$\rightsquigarrow$ Arithmetik mit unendlichen Listen solcher Ziffern stürzt nicht ab!

Was kann man mit solchen Zahlen machen?

Eine reelle Zahl α kann genau dann als unendliche Liste wie in Variante 2 dargestellt werden, wenn es ein Programm gibt, daß für jede Eingabe $n \in \mathbb{N}$ eine rationale Approximation a mit Fehler $|\alpha - a| \leq 1/n$ berechnen kann. In der Tat bestimmen die ersten m Ziffern die Zahl bis auf eine Genauigkeit von $0.5555\ldots \cdot 10^{-m} = \frac{5}{9 \cdot 10^m}$. Um eine rationale Approximation mit Fehler kleiner gleich $1/n$ zu bestimmen, reicht es, ein m zu bestimmen sodaß $\frac{5}{9 \cdot 10^m} \leq \frac{1}{n}$ gilt, und die ersten m Ziffern auszurechnen. Umgekehrt können aus der Kenntnis von α bis auf eine Genauigkeit von $\frac{1}{2 \cdot 9 \cdot 10^m}$ mindestens m Ziffern berechnet werden, die zu einer unendlichen Darstellung von α gemäß Variante 2 vervollständigt werden können. Dann läßt sich aber auch ein Programm `digit` schreiben, daß bei Eingabe n die n-te Ziffer einer solchen Darstellung berechnet. Die Zahl α läßt sich dann darstellen durch die unendliche Liste `alpha`, definiert durch

```
digits_from(n) = [ digit(n) :  digits_from(n+1) ]
alpha = digits_from(0)
```

Wir nennen eine reelle Zahl ***berechenbar***, wenn sie eine Darstellung wie oben besitzt, oder – anders ausgedrückt – wenn sie bis auf eine beliebige Genauigkeit berechnet werden kann. Die arithmetischen Operationen laßen sich für Variante 2 implementieren, das heißt die Summe bzw. Differenz usw. von zwei berechenbaren Zahlen ist wieder berechenbar und läßt sich durch ein Programm automatisch berechnen. Andererseits kann nicht automatisch entschieden werden, ob eine berechenbare Zahl gleich Null ist oder nicht: Dazu müßten alle Ziffern der unendlichen Liste bekannt sein, und ein Programm, das bei gegebener unendlichen Liste entscheidet, ob alle Elemente gleich Null sind, kann nicht terminieren.

Die folgende Zusammenstellung gibt eine Übersicht, welche Operationen mit berechenbaren Zahlen automatisch durchgeführt werden können und welche nicht.

ja: Arithmetic: $+, -, \cdot, /$ (falls Nenner $\neq 0$).

nein: Test auf Null. Die reelle Zahl 0 besitzt zwar eine eindeutige Darstellung, alle Ziffern müssen 0 sein, aber der Vergleich unendlicher Listen terminiert nicht, wenn die Listen gleich sind.

nein: Test auf Gleichheit zweier Zahlen.

ja: Maximum/Minimum. $\max(\alpha, \beta)$ und $\min(\alpha, \beta)$ sind berechenbar. Zur Berechnung einer Approximation von $\max(\alpha, \beta)$ bis auf eine Genauigkeit von ϵ reicht es, die Zahlen α und β bis auf eine Genauigkeit von ϵ zu kennen.

Man beachte, daß man zwar das Maximum in obigem Sinn berechnet werden kann, daß sich aber im allgemeinen nicht entscheiden läßt, welche von den beiden Zahlen nun die größere ist oder ob nicht doch die beiden gleich sind!

ja: Sortieren. Die Sortierung von (α, β) ist $(\min(\alpha, \beta), \max(\alpha, \beta))$. Also läßt sich das Sortier-Problem auf die Berechnung von Maximum und Minimum zurückführen.

ja: Konvertierung von Variante 1 nach Variante 2.

nein: Konvertierung von Variante 2 nach Variante 1.

nein: Berechnung einer korrekten Dezimalapproximation der Ordnung n. Um zum Beispiel eine korrekte Dezimalapproximation von α auf eine Stelle zu berechnen, müßte entschieden werden, ob α größer/gleich oder kleiner/gleich 0.05 ist; und das ist im allgemeinen unentscheidbar.

ja: Berechnung einer korrekten Dezimalapproximation der Ordnung n oder $n+1$. Wenn die $n+1$-te Ziffer zwischen -4 und 4 ist, kann eine Approximation der Ordnung n aus den ersten n Ziffern bestimmt werden. Wenn die $n+1$-te Ziffer gleich ± 5 ist, und die $n+2$-te Ziffer ungleich 0 ist, dann bestimmen diese beiden Ziffern, ob auf- oder abgerundet werden muß, und man kann wieder eine Approximation der Ordnung n bestimmen. Wenn die $n+2$-te Ziffer gleich 0 ist, dann bestimmen die ersten $n+1$ Ziffern eine Approximation der Ordnung $n+1$.

ja: exp, log (falls Argument > 0).

ja: Innerster Punkt einer Intervallschachtelung (gegeben als unendliche Liste von Paaren reeller Zahlen).

ja: exp, log (falls Argument > 0).

ja: Innerster Punkt einer Intervallschachtelung (gegeben als unendliche Liste von Paaren reeller Zahlen).

Grenzwerte von Cauchy-Folgen laßen sich im Allgemeinen nicht bestimmen. Solange nur endlich viele Folgenglieder bekannt sind, läßt sich nämlich über die Lage des Grenzwerts keine Aussage machen. Daher kann nicht einmal die erste Ziffer des Grenzwerts bestimmt werden, bevor nicht unendlich viele Folgenglieder bekannt sind, und das ist – nie.

Wenn jedoch ein Cauchy-Modulus für die Folge gegeben ist, das heißt eine Funktion $N : \mathbb{R}^+ \to \mathbb{N}$ mit

$$|a_{N(\epsilon)+m} - a_{N(\epsilon)+n}| \leq \epsilon$$

für alle $\epsilon \in \mathbb{R}^+$, $m, n \in \mathbb{N}$, dann läßt sich der Grenzwert berechnen. In diesem Fall ist nämlich $a_{N(\epsilon)}$ eine ϵ-Approximation, der Grenzwert läßt sich also beliebig genau berechnen und ist daher berechenbar.

Kapitel 3

Lösen von Gleichungen mittels Grenzwerten

3.0.2 Beispiel Wir betrachten die Gleichung

$$x - \cos x = 0.$$

Um sie zu lösen, brauchen wir nur den Cosinus zu iterieren:

$$\begin{aligned} a_0 &:= 0, \\ a_{n+1} &:= \cos a_n. \end{aligned}$$

Wir erhalten dann die Lösung $x = 0.73908513321516\ldots$.

3.0.3 Satz *Wenn die Folge $a_{n+1} := f(a_n)$ konvergiert und f stetig ist, dann ist der Grenzwert ein Fixpunkt.*

Dieser Satz liefert keine Aussage über die Kovergenzrate (Cauchy-Modulus). Die Konvergenz muß vorausgesetzt werden!

3.0.4 Definition Eine Funktion heißt ***kontrahierend*** mit ***Kontraktionszahl*** α, $0 < \alpha < 1$, wenn für alle x, y gilt

$$|f(x) - f(y)| \leq \alpha |x - y|.$$

3.0.5 Theorem (Banach) *Es sei $f\colon [a,b] \to [a,b]$ kontrahierend mit Kontraktionszahl α. Dann besitzt f genau einen Fixpunkt x. Sei $(x_n)_n$ die Folge mit $x_0 \in [a,b]$ beliebig und $x_{n+1} = f(x_n)$. Dann konvergiert $(x_n)_n$ gegen x. Die Konvergenzrate ist gegeben durch*

$$|x_n - x| \leq \alpha^n (b - a).$$

Beweis Wir betrachten die Folge von Intervallen definiert durch

$$I_0 := [a,b], I_{n+1} := f(I_n).$$

Wir zeigen, daß $(I_n)_n$ eine Intervallschachtelung ist.

Zunächst ist zu zeigen, daß $I_{n+1} \subset I_n$ ist. Wir zeigen diese Behauptung durch Induktion. Für $n = 0$ ist offensichtlich $I_1 = f([a,b]) \subset [a,b] = I_0$, damit ist der Induktionsanfang gemacht. Sei $n \geq 1$. Für zwei beliebige Intervalle gilt $J_1 \subset J_2 \Rightarrow f(J_1) \subset f(J_2)$. Daher ist $I_{n+2} = f(I_{n+1}) \subset f(I_n) = I_{n+1}$ nach Induktionsvoraussetzung.

Wir zeigen weiter, daß $L(I_n) \leq \alpha^n(b-a)$ ist. Für $n = 0$ ist die Behauptung offensichtlich richtig. Für $n \geq 1$ sei $I_n = [a_n, b_n]$. Dann sind a_n und b_n die Bilder von zwei Werten $c, d \in I_{n-1}$. Wegen der Kontraktionseigenschaft gilt $L(I_n) = b_n - a_n \leq \alpha|c-d| \leq \alpha L(I_{n-1})$, und diese Zahl ist aber kleiner oder gleich $\alpha^n(b-a)$ nach Induktionsvoraussetzung.

Da $\mathbb{R}$ vollständig ist, hat $(I_n)_n$ einen innersten Punkt y. Offensichtlich liegt das Folgenglied x_n im Intervall I_n, und daher gilt $|x_n - y| \leq L(I_n) = \alpha^n(b-a)$. Also ist konvergiert die Folge $(x_n)_n$ gegen y, und die Konvergenzrate ist wie behauptet.

Nun zeigen wir noch, daß y tatsächlich ein Fixpunkt ist. In der Tat muß $f(y)$ in jedem Intervall I_n liegen, also ist $f(y)$ ebenfalls innerster Punkt. Wegen der Eindeutigkeit des innersten Punktes ist $f(y) = y$.

Es bleibt noch zu zeigen, daß es nicht mehr als einen Fixpunkt gibt. Dieser Teil sei dem Leser als Übung überlassen.

3.0.6 Bemerkung Wenn f differenzierbar ist und $|f'(x)| \leq \alpha$, dann gilt $|f(x) - f(y)| \leq \alpha\,|x-y|$. Eine Schranke für die Ableitung ist daher eine Kontraktionszahl.

3.0.7 Beispiel Sei $f\colon [0,1] \to [0,1]$, $x \mapsto \cos x$. Dann gilt $|f'(x)| = |\sin x| \leq \sin(1) \leq 0.85$; daher konvergiert die obige Folge.

Nachteil der gegebenen Version: nur 1 Variable. Die passende Verallgemeinerung führt zum Konzept der metrischen Räume.

3.0.8 Definition Es sei X eine Menge. Eine ***Metrik*** auf X ist eine Funktion $d\colon X \times X \to \mathbb{R}$, sodaß gilt:

$$\begin{aligned} d(x,y) \geq 0 \text{ und } d(x,y) = 0 \iff x = y; \\ d(x,y) = d(y,x); \\ d(x,y) \leq d(x,z) + d(z,y). \end{aligned}$$

(X,d) heißt dann ein ***metrischer Raum***.

3.0.9 Beispiel

1. $\mathbb{R}^2$ mit $d(x,y) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$.

2. $\mathbb{R}^2$ mit $d(x,y) = \max(|x_1 - x_2|, |y_1 - y_2|)$.

3. $\mathbb{R}^2$ mit $d(x,y) = |x_1 - x_2| + |y_1 - y_2|$ ("Taxifahrer-Metrik").

3.0.10 Definition

1. Eine Folge $(x_n)_n$ heißt ***Cauchy-Folge***, wenn es für jedes $\epsilon > 0$ ein $N = N(\epsilon)$ gibt mit

$$d(x_{N+m}, x_{N+n}) \leq \epsilon, \text{ für alle } n, m \in \mathbb{N}.$$

2. Eine Folge $(x_n)_n$ ***konvergiert*** gegen x, wenn es für jedes $\epsilon > 0$ ein N gibt mit

$$d(x_{N+m}, x) \leq \epsilon, \text{ für alle } n \in \mathbb{N}.$$

Frage: Ist der Grenzwert überhaupt eindeutig bestimmt? (Übung)

3.0.11 Satz *Jede konvergente Folge ist Cauchy-Folge.*

Beweis Es sei $(x_n)_n$ eine Folge mit $\lim_{n\to\infty} x_n = x$, und mit Kontraktionsrate $M\colon \mathbb{R}^+ \to \mathbb{N}$, $\epsilon \mapsto M(\epsilon)$ (d.h. $d(x_{M(\epsilon)+n}, x) \leq \epsilon$, für alle $\epsilon > 0$, $n \in \mathbb{N}$. Dann ist $N\colon \mathbb{R}^+ \to \mathbb{N}$, $\epsilon \mapsto N(\epsilon) := M(\frac{\epsilon}{2})$ ein passender Cauchy-Modulus:

$$\begin{aligned} d(x_{N(\epsilon)+m}, x_{N(\epsilon)+n}) &= d(x_{M(\frac{\epsilon}{2})+m}, x_{M(\frac{\epsilon}{2})+n}) \\ \leq d(x_{M(\frac{\epsilon}{2})+m}, x) + d(x_{M(\frac{\epsilon}{2})+n} x) &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

3.0.12 Definition Wenn jede Cauchy-Folge konvergiert, dann heißt X bezüglich der Metrik d ***vollständig***.

3.0.13 Beispiel

1. $\mathbb{R}$ ist vollständig bezüglich $d(x,y) = |x - y|$.

2. $\mathbb{Q}$, mit der selben Metrik, ist nicht vollständig.

3. $[a,b] \subseteq \mathbb{R}$ ist vollständig.

4. Alle Beispiele in 3.0.9 sind vollständig.

3.0.14 Definition Sei (X,d) ein metrischer Raum und $A \subseteq X$. Dann heißt A ***abgeschlossen***, wenn für alle konvergenten Folgen $(x_n)_n$, $x_n \in A$, auch der Grenzwert $x = \lim_{n\to\infty} x_n$ in A liegt.

3.0.15 Beispiel

1. $[a, b] \subseteq \mathbb{R}$ ist abgeschlossen.

2. Die Vereinigung zweier abgeschlossenen Mengen ist abgeschloßen. Insbesondere ist die Vereinigung zweier Intervalle, also z.B. die Menge $\{x \mid 1 \leq x \leq 2 \vee 3 \leq x \leq 4\}$ abgeschlossen.

3. Die Menge $\mathbb{R}^+$ ist nicht abgeschlossen, weil die Folge $(\frac{1}{n})_n$ in $\mathbb{R}^+$ liegt, der Grenzwert 0 aber nicht.

3.0.16 Satz *Wenn (X, d) vollständig und $A \subseteq X$ abgeschlossen ist, dann ist auch (A, d) vollständig.*

3.0.17 Definition Es seien (X, d) metrischer Raum. Eine Funktion $f\colon X \to X$ heißt ***Lipschitz-stetig*** mit Konstante K, wenn für alle $x, y \in X$ gilt

$$d(f(x_1), f(x_2)) \leq K d(x_1, x_2).$$

3.0.18 Theorem (Banach'scher Fixpunktsatz) *Es sei (X, d) ein vollständiger metrischer Raum und $f\colon (X, d) \to (X, d)$ Lipschitz-stetig mit Konstante $\alpha < 1$. Dann hat f genau einen Fixpunkt x. Die Folge $x_{n+1} := f(x_n)$, mit $x_0 \in X$ beliebig, konvergiert gegen x und die Konvergenzrate ist gegeben durch*

$$d(x, x_n) \leq \frac{\alpha^n}{1-\alpha} d(x_0, x_1).$$

Beweis Aus der Voraussetzung erhalten wir mit Induktion

$$d(x_n, x_{n+1}) \leq \alpha^n d(x_0, x_1).$$

Die Dreiecksungleichung liefert dann, zusammen mit der Formel für die geometrische Reihe,

$$d(x_n, x_{n+m}) \leq \frac{\alpha^n}{1-\alpha} d(x_0, x_1).$$

Damit ergibt sich der Cauchy-Modulus

$$N(\epsilon) := \lceil \frac{\log \frac{\epsilon(1-\alpha)}{d(x_0, x_1)}}{\log \alpha} \rceil.$$

Ganz allgemein, ist die Cauchy-Modulus auch ein Konvergenzrate.

3.0.19 Anwendung Große lineare Gleichungssystem lassen sich folgendermaßen iterativ lösen. Gegeben sei dazu eine $n \times n$-Matrix $\mathbf{A} = \begin{pmatrix} a_{1\,1} & \dots & a_{1\,n} \\ \vdots & & \vdots \\ a_{n\,1} & \dots & a_{n\,n} \end{pmatrix}$ und ein Spaltenvektor $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$. Gesucht ist dann ein Vektor $\mathbf{x} \in \mathbb{R}^n$, sodaß $\mathbf{Ax} = \mathbf{b}$, d.h. eine Lösung des linearen Gleichungssystems

$$\begin{aligned} &a_{1,1}x_1 + \dots + a_{1\,n}x_n = b_1, \\ &\dots\dots\dots\dots\dots\dots\dots\dots \\ &a_{n\,1}x_1 + \dots + a_{n\,n}x_n = b_n. \end{aligned}$$

Dabei sei $\mathbf{A}$ nicht-singulär (sodaß eine eindeutige Lösung existiert). Nun wählen wir eine Matrix $\mathbf{B}$, sodaß die Abbildung $\mathbb{R} \to \mathbb{R}$,

$$\mathbf{x} \mapsto \mathbf{x} - \mathbf{B}(\mathbf{Ax})$$

Lipschitz-stetig mit Konstante $\alpha < 1$ ist. Solche Matrizen $\mathbf{B}$ sind leicht zu finden (z.B. $\mathbf{B} = \beta \cdot \mathbf{A}^{\mathrm{T}}$, mit $\beta > 0$, aber "'klein genug"'). Wir betrachten dann die Folge

$$\mathbf{x}_0 = \mathbf{b}, \quad \mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{B}(\mathbf{b} - \mathbf{Ax}_n).$$

Diese konvergiert dann gegen die Lösung, durch Anwendung des Banach'schen Fixpunktsatzes auf

$$f\colon \mathbb{R}^n \to \mathbb{R}^n, \quad \mathbf{x} \mapsto \mathbf{x} + \mathbf{B}(\mathbf{b} - \mathbf{Ax}).$$

Bei der Umformulierung eines Problems $f(x) = 0$ in ein Fixpunktproblem für eine kontrahierende Funktion ergeben sich die folgenden Probleme:

- dies ist nicht immer möglich;
- zusätzlich soll α möglichst klein sein.

3.0.3 Newton-Verfahren

Jede Nullstelle einer Funktion f ist Fixpunkt von

$$g(x) := x + a(x)f(x).$$

Wir wollen die Funktion a so wählen, daß $|g'(x)|$ möglichst klein ist ($|g'(x)|$ ist Lipschitz-Konstante!):

$$|g'(x)| = |1 + a(x)f'(x) + a'(x)f(x)|\,.$$

Wenn wir jetzt den ersten Teil $= 0$ setzen, also $1 + a(x)f'(x) = 0$, dann ergibt sich

$$a(x) := -\frac{1}{f'(x)},$$

und somit

$$|a'(x)f(x)| = \left|\frac{f''(x)f(x)}{f'(x)^2}\right|.$$

Daraus ergibt sich das ***Newton-Verfahren***:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

3.0.20 Satz *Wenn f zweimal differenzierbar ist, f' keine Nullstelle hat und*

$$\left|\frac{f''(x)f(x)}{f'(x)^2}\right| \leq \alpha < 1$$

für alle x gilt, dann konvergiert das Newton-Verfahren gegen die Lösung.

Das Newton-Verfahren läßt sich auf $\mathbb{R}^n$ (und sogar unendlich-dimensionale Räume) verallgemeinern. Dazu braucht man aber höher-dimensionale Ableitungen und lineare Algebra.

3.1 Stetige Funktionen

Es sei $f : \mathbb{R} \to \mathbb{R}$ eine Funktion. Wir wollen für gegebenes $x \in \mathbb{R}$ (gegeben durch eine unendliche Dezimalapproximation mit $-5, \ldots, 5$) den Wert $f(x)$ berechnen. Wie macht man das?

1. "'$f(x)$ zu berechnen"' ist äquivalent dazu, für gegebenes $\epsilon > 0$, eine Intervall zu finden, das Länge ϵ hat und $f(x)$ enthält.
2. Wir können x beliebig genau approximieren.
3. Wenn f z.B. Lipschitz-stetig wäre, mit Lipschitz-Konstante K, dann reicht es, x bis auf Genauigkeit $\delta = \frac{\epsilon}{K}$ zu kennen, denn aus $|x - x_\delta| \leq \frac{\epsilon}{K}$ folgt dann $|f(x) - f(x_\delta)| \leq \epsilon$.
4. Allgemein: wenn es eine Funktion $\mathbb{R}^+ \to \mathbb{R}^+$, $\epsilon \mapsto \delta(\epsilon)$ gibt mit

   $$|x - x_\delta| \leq \delta(\epsilon) \implies |f(x) - f(x_\delta)| \leq \epsilon,$$

   dann reicht es, x bis auf Genauigkeit $\delta(\epsilon)$ zu kennen.

3.1.1 Definition Eine Funktion heißt ***gleichmäßig stetig*** auf einem Intervall I, wenn es eine Funktion $\epsilon \mapsto \delta(\epsilon)$ gibt, sodaß

$$|x - x_\delta| \leq \delta(\epsilon) \implies |f(x) - f(x_\delta)| \leq \epsilon,$$

für alle $\epsilon > 0$, $x, y \in I$ gilt.

Die Funktion δ heißt dann ***Stetigkeitsmodul***.

3.1.2 Bemerkung Wenn f in einem Intervall I, das x enthält, gleichmäßig stetig ist, und er der Stetigkeitsmodul bekannt ist, und wenn wir f für die approximierenden Stellen beliebig genau berechnen können, dann läßt sich $f(x)$ berechnen.

Gibt es Funktionen, die stetig, aber nicht Lipschitz-stetig sind?

Sei $f\colon [0,1] \to \mathbb{R}$ die Wurzelfunktion $x \mapsto \sqrt{x}$. Es folgt: $\big(f(x) - f(y)\big)^2 = (\sqrt{x} - \sqrt{y})^2 \leq \left|\sqrt{x} - \sqrt{y}\right| \left|\sqrt{x} + \sqrt{y}\right| \leq |x - y|$, also

$$|x - y| \leq \epsilon^2 \implies |f(x) - f(y)| \leq \epsilon.$$

ϵ^2 ist also Stetigkeitsmodul.

Wäre f Lipschitz-stetig mit Konstante K, dann würde $\left|\sqrt{x} - \sqrt{y}\right| \leq K\,|x - y|$ gelten, also insbesondere (mit $x := 0$, $y = \frac{1}{2K^2}$) $\left|\frac{1}{2K} - 0\right| \leq K\left|\frac{1}{4K^2} - 0\right|$, d.h. $4 \leq 2$, Widerspruch!

Funktionen mit Stetigkeitsmodul $C \cdot \epsilon^r$ heißen ***Hölder-stetig***.

3.1.3 Satz *Jede algebraische Funktion ist Hölder-stetig (in jedem Intervall der Definitionsmenge).*

Es gibt auch Funktionen, die gleichmäßig stetig, aber nicht Hölder-stetig sind.

3.1.4 Definition Eine Funktion f heißt ***stetig an der Stelle*** x, wenn es einen "‘einseitigen"’ Stetigkeitsmodul gibt, d.h. ein $\delta(\epsilon) > 0$, sodaß

$$|x - y| \leq \delta(\epsilon) \implies |f(x) - f(y)| \leq \epsilon,$$

für alle $\epsilon > 0$ und y.

3.1.5 Definition Eine Funktion f heit ***stetig***, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

1. f ist gleichmäßig stetig in jedem endlichen Intervall im Definitionsbereich.
2. f ist stetig an jeder Stelle.
3. der Graph von f läßt sich zeichnen, "‘ohne mit dem Bleistift abzusetzen"’.
4. Falls $lim_{x\to\infty} x_n = x$ dann ist auch $lim_{x\to\infty} f(x_n) = f(x)$.

Die zweite und vierte Variante sind in offensichtlicher Weise auf beliebige metrische Räume erweiterbar.

Kapitel 4

Differential- und Integralrechnung

4.1 Das Integral

Beim Integral handelt es sich um einen grundlegenden Begriff zur mathematischen Beschreibung von Messungen. Der Begriff ist aber auch geschichtlich von Interesse, weil durch das Integral viele Probleme aus der klassisch/antiken Mathematik – die "Quadraturprobleme" – plötzlich mit einer einheitlichen Methode zugänglich gemacht wurden.

4.1.1 Definition Es sei $[a, b]$ ein Intervall. Eine *Zerlegung*

$$[a, b] = [a = x_0, x_1] \cup [x_1, x_2] \cup \cdots \cup [x_{n-1}, x_n = b]$$

in n Teilintervalle ist gegeben durch die Menge der Randpunkte $\{a = x_0, x_1, \ldots, [x_{n-1}, x_n = b\}$.

4.1.2 Beispiel

1. Die *equidistante Zerlegung* von $[a, b]$ in n Teilintervalle ist $\{a, a + \frac{b-a}{n}, a + 2\frac{b-a}{n}, \ldots, a + (n-1)\frac{b-a}{n}, b\}$.

2. Eine nicht equidistante Zerlegung von $[0, 1]$ ist etwa $\{0, \frac{1}{2}, \frac{3}{4}, \frac{7}{8}, 1\}$.

4.1.3 Definition Die *Schrittweite* der Zerlegung ist definiert als $\max_i(x_i - x_{i-1})$, die maximale Länge der Teilintervalle.

Eine Zerlegung Z_2 ist *feiner* als eine Zerlegung Z_1, wenn $Z_2 \subset Z_1$ ist, d.h. wenn jeder Randpunkt von Z_2 auch ein Randpunkt von Z_1 ist.

4.1.4 Beispiel

1. Die equidistante Zerlegung in 8 Teilintervalle ist feiner als die equidistante Zerlegung in 4 Teilintervalle.

2. Die equidistante Zerlegung in 5 Teilintervalle ist nicht feiner als die equidistante Zerlegung in 4 Teilintervalle.

4.1.5 Definition Eine Menge $X \subset \mathbb{R}$ heißt *nach oben beschränkt/nach unten beschränkt/beschränkt*, wenn es eine Zahl $M \in \mathbb{R}$ gibt, sodaß $x \leq M/x \geq M/|x| \leq M$ for all $x \in X$.

Eine Funktion $f : D \to \mathbb{R}$ heißt *nach oben beschränkt/nach unten beschränkt/beschränkt*, wenn der Wertebereich $f(D) \subset \mathbb{R}$ nach oben beschränkt/nach unten beschränkt/beschränkt ist.

4.1.6 Theorem *Jede nach oben beschränkte nichtleere Menge X besitzt eine kleinste obere Schranke (genannt "Supremum" $\sup(X)$).*

Jede nach unten beschränkte nichtleere Menge X besitzt eine größte untere Schranke (genannt "Infimum" $\inf(X)$).

Beweis Es sei $a \in X$, und b eine obere Schranke von X. Wir konstruieren eine Intervallschachtelung durch $I_0 := [a, b]$, und I_{n+1} ist entweder das linke Teilintervall $[a_n, \frac{a_n+b_n}{2}]$ oder das rechte Teilintervall $[\frac{a_n+b_n}{2}, b_n]$ von I_n, je nachdem ob der Mittelpunkt $\frac{a_n+b_n}{2}$ eine obere Schranke ist oder nicht. Nach Konstruktion gilt für jedes der Intervalle I_n, daß es mindestens ein Element aus X enthält, und daß der rechte Randpunkt eine obere Schranke ist.

Es sei s der innerste Punkt der Intervallschachtelung. Dann ist s der Grenzwert einer Folge von oberen Schranken, nämlich der rechten Randpunkte b_n. Für jeden Punkt $x \in X$ gilt $b_n \in [x, \infty)$; da $[x, \infty)$ abgeschloßen ist, ist auch $s \in [x, \infty)$, d.h. $s \geq x$. Es folgt, daß s eine obere Schranke ist.

Um zu zeigen, daß s die kleinste obere Schranke ist, nehmen wir indirekt an, daß $s' < s$ ebenfalls eine obere Schranke ist. Es sei I_n ein Intervall mit Länge kleiner als $s - s'$. Dann liegt s' links von I_n. Andererseits enthält I_n einen Punkt aus X, also ist s' keine obere Schranke – Widerspruch.

4.1.7 Definition Es sei $[a, b]$ ein Intervall. Es sei $f : [a, b] \to \mathbb{R}$ beschränkt. Es sei $Z = \{a = x_0, x_1, \ldots, x_{n-1}, x_n = b\}$ eine Zerlegung von $[a, b]$. Wir definieren *Obersumme*

$$O(f, a, b, Z) := \sum_{i=1}^{n} (x_i - x_{i-1}) \sup(f([x_{i-1}, x_i]))$$

und *Untersumme*

$$U(f, a, b, Z) := \sum_{i=1}^{n} (x_i - x_{i-1}) \inf(f([x_{i-1}, x_i])).$$

Offensichtlich gilt stets $U(f,a,b,Z) \leq O(f,a,b,Z)$.

4.1.8 Satz *Es seien Z_1, Z_2 zwei Zerlegungen, wobei Z_2 feiner sei als Z_1. Dann ist*

$$U(f,a,b,Z_1) \leq U(f,a,b,Z_2) \leq O(f,a,b,Z_2) \leq O(f,a,b,Z_1).$$

Beweis (geometrisch, durch Skizze an der Tafel.)

4.1.9 Satz *Es seien Z_1, Z_2 zwei beliebige Zerlegungen. Dann ist $U(f,a,b,Z_1) \leq O(f,a,b,Z_2)$.*

Beweis Die Zerlegung $Z_3 := Z_1 \cup Z_2$ ist feiner als Z_1 und feiner als Z_2. Nach Proposition 4.1.8 gilt daher

$$U(f,a,b,Z_1) \leq U(f,a,b,Z_3) \leq O(f,a,b,Z_3) \leq O(f,a,b,Z_2).$$

Aus Proposition 4.1.9 folgt, daß die Menge aller Untersummen nach oben beschränkt ist: jede Obersumme ist obere Schranke. Umgekehrt ist jede Untersumme eine untere Schranke für die Menge der Obersummen.

4.1.10 Definition Es sei $[a,b] \to \mathbb{R}$ eine beschränkte Funktion. Dann heißt f *Riemann-integrierbar*, wenn $\sup_Z U(f,a,b,Z) = \inf_Z O(f,a,b,Z)$ ist. Das *Integral* von f zwischen a und b ist definiert als

$$\int_a^b f(x)dx := \sup_Z U(f,a,b,Z) = \inf_Z O(f,a,b,Z).$$

Geometrisch entspricht dem Integral der Flächeninhalt eingeschlossen vom Funktionsgraphen, der $x - Achse$, und den beiden senkrechten Geraden $x = a$ und $x = b$.

4.1.11 Beispiel Es sei $f : [1,2] \to \mathbb{R}$, $x \mapsto x$. Wir berechnen Obersumme und Untersumme für die Zerlegung $Z_n := \{1, q, q^2, \ldots, q_n = 2\}$, wobei $q := \sqrt[n]{2}$ ist:

$$O(f,1,2,Z_n) = \sum_{i=1}^{n} (q^i - q^{i-1})q^i = \sum_{i=0}^{n-1} (q^{i+1} - q^i)q^{i+1}$$

$$= \sum_{i=0}^{n-1} (q^{2i+2} - q^{2i+1}) = \sum_{i=0}^{n-1} q^{2i}(q^2 - q) = (q^2 - q)\frac{q^{2n} - 1}{q^2 - 1}$$

$$= \frac{q(q-1)(q^{2n}-1)}{(q+1)(q-1)} = \frac{q(4-1)}{q+1} = \frac{3q}{q+1} = \frac{3\sqrt[n]{2}}{\sqrt[n]{2}+1}.$$

$$U(f,1,2,Z_n) = \sum_{i=1}^{n}(q^i - q^{i-1})q^{i-1} = \sum_{i=0}^{n-1}(q^{i+1} - q^i)q^i$$

$$= \sum_{i=0}^{n-1}(q^{2i+1} - q^{2i}) = \sum_{i=0}^{n-1} q^{2i}(q-1) = (q-1)\frac{q^{2n}-1}{q^2-1}$$

$$= \frac{(q-1)(q^{2n}-1)}{(q+1)(q-1)} = \frac{4-1}{q+1} = \frac{3}{q+1} = \frac{3}{\sqrt[n]{2}+1}.$$

Das Supremum *aller* Untersummen und das Infimum *aller* Obersummen liegt im Intervall $I_n := [U(f,1,2,Z_n), O(f,1,2,Z_n)]$. Die Länge des Intervalls $[U(f,1,2,Z_n), O(f,1,2,Z_n)]$ geht gegen 0 für $n \to \infty$:

$$\lim_{n\to\infty}\left(\frac{3\sqrt[n]{2}}{\sqrt[n]{2}+1} - \frac{3}{\sqrt[n]{2}+1}\right) = \frac{3\cdot 1}{1+1} - \frac{3}{1+1} = 0.$$

Daher ist das Supremum aller Untersummen gleich dem Infimum aller Obersummen und gleich dem innersten Punkt der Intervalle, also

$$\int_1^2 x dx = \frac{3}{1+1} = \frac{3}{2}.$$

Im obigen Beispiel gelingt es, das Integral zu berechnen, indem man eine Folge von Zerlegungen konstruiert, bei denen der Grenzwert der Obersummen gleich dem Grenzwert der Untersummen ist. Im folgenden Satz wird diese Methode für beliebige Funktionen verallgemeinert.

4.1.12 Theorem *Es sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Es sei Z_n eine Folge von Zerlegungen, sodaß $\lim_{n\to\infty}(O(f,a,b,Z_n) - U(f,a,b,Z_n)) = 0$ ist. Dann ist f integrierbar, und das Integral läßt sich als Grenzwert einer Folge $(a_n)_n$ berechnen, wobei a_n eine reelle Zahl zwischen $O(f,a,b,Z_n)$ und $U(f,a,b,Z_n)$ ist.*

(Eine einfache Methode, eine solche Zahl zu berechnen, ist, das Supremum bzw. Infimum in der Formel für Obersumme bzw. Untersumme durch einen beliebigen Funktionswert im entsprechenden Intervall zu ersetzen.)

Beweis Analog wie im obigen Beispiel.

4.1.13 Bemerkung Es stellt sich die Frage, wie man eine Folge von Zerlegungen findet, die die Voraussetzungen vom obigen Satz erfüllt. Gute Kandidaten sind solche Folgen, für die die Folge der Schrittweiten gegen Null konvergieren. Man kann sogar zeigen, daß jede Folge von dieser Art die Voraussetzungen vom obigen Satz erfüllt, wenn die Funktion f integrierbar ist.

Es gibt beschränkte Funktionen, die nicht integrierbar sind, wo also das Supremum aller Untersummen echt kleiner ist als das Infimum der Obersummen. Ein Beispiel ist die Funktion $f : [0,1] \to \mathbb{R}$, $x \mapsto 1$ falls x rational, und $x \mapsto 0$ falls x irrational. Hier ist jede Obersumme gleich 1, und jede Untersumme gleich 0.

4.1.14 Theorem (Fundamentalsatz der Integralrechnung) *Jede stetige Funktion ist integrierbar.*

Beweis Es sei $f : [a,b] \to \mathbb{R}$ stetig. Es sei $\delta : \mathbb{R}^+ \to \mathbb{R}^+$ ein Stetigkeitsmodul, d.h. es gilt

$$|x-y| \leq \delta(\epsilon) \Rightarrow |f(x) - f(y)| \leq \epsilon$$

für alle $\epsilon > 0$, $x, y \in [a,b]$. Es sei $\epsilon > 0$ beliebig. Es sei Z eine Zerlegung mit Schrittweite $h \leq \delta(\frac{\epsilon}{b-a})$. Dann unterscheiden sich in jedem Teilintervall von Z das Infimum und das Supremum höchstens um $\frac{\epsilon}{b-a}$. Daraus folgt, daß $O(f,a,b,Z) - U(f,a,b,Z) \leq \epsilon$ ist. Daher unterscheiden sich das Supremum aller Untersummen und das Infimum aller Obersummen höchstens um ϵ. Da aber ϵ beliebig gewählt werden kann, stimmen das Supremum aller Untersummen und das Infimum aller Obersummen überein.

4.1.15 Bemerkung Der Beweis ist konstruktiv: für eine gegebene Funktion mit gegebenem Stetigkeitsmodul läßt sich das Integral mit Hilfe der Beweis-Idee bis auf eine vorgegebene beliebige Genauigkeit $\epsilon > 0$ berechnen. Dazu reicht es, eine Zerlegung mit Schrittweite $h \leq \delta(\frac{\epsilon}{b-a})$ zu konstruieren und einen Wert zurückzugeben, der zwischen der Untersumme und der Obersumme liegt, zum Beispiel indem man das Infimum bzw. Supremum durch einen beliebigen Funktionswert des Teilintervalls ersetzt.

Umgekehrt gibt es sehr wohl integrierbare Funktionen, die nicht stetig sind. Eine weitere Klasse von Funktionen, die integrierbar sind, ist die Klasse der monotonen Funktionen. Eine Funktion heißt *monoton steigend/fallend*, wenn $x \leq y \Rightarrow f(x) \leq f(y)$ bzw. $x \leq y \Rightarrow f(x) \geq f(y)$ für alle x, y im Definitionsbereich gilt.

4.2 Eigenschaften des Integrals

Einige Eigenschaften des Integrals werden ohne Beweis angegeben. Die Beweise sind ausnahmslos "mathematische Routine", es sind keine wesentlichen Ideen nötig. Die Eigenschaften sind aber nützlich zur Manipulation mit Integralen (Technik der Integralrechnung).

4.2.1 Satz (Linearität) *Es seien $f, g : [a,b] \to \mathbb{R}$ integrierbar, $\lambda \in \mathbb{R}$. Dann sind auch $f + g : [a,b] \to \mathbb{R}$ und $\lambda \cdot f : [a,b] \to \mathbb{R}$ integrierbar, und es gilt*

$$\int_a^b (f+g)(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx,$$

$$\int_a^b (\lambda \cdot f)(x)dx = \lambda \int_a^b f(x)dx.$$

Unter Verwendung der Linearität lassen sich *Polynomfunktionen* integrieren, d.h. Funktionen der Art $x \mapsto a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$, für Koeffizienten $a_0, \ldots, a_n \in \mathbb{R}$. Dazu reicht es, die Integrale der Potenzfunktionen zu kennen:

$$\int_a^b x^n dx = \frac{b^{n+1}}{n+1} - \frac{a^{n+1}}{n+1}.$$

Da jede Polynomfunktion eine Linearkombination von Potenzfunktionen ist, ergibt sich das Integral der Polynomfunktion als die entsprechende Linearkombination der Integrale der Potenzfunktionen.

4.2.2 Beispiel

$$\int_1^2 (5x + 3x^3)dx = 5 \int_1^2 xdx + 3 \int_1^2 x^3 dx$$

$$= 5(\frac{2^2}{2} - \frac{1^2}{2}) + 3(\frac{2^4}{4} - \frac{1^4}{4}) = 18.75.$$

4.2.3 Satz (Additivität auf Intervallen) *Es seien $a, b, c \in \mathbb{R}$ mit $a < b < c$. Es sei $f : [a, c] \to \mathbb{R}$. Dann ist f genau dann integrierbar, wenn die Einschränkungen $f : [a, b] \to \mathbb{R}$ und $f : [b, c] \to \mathbb{R}$ integrierbar sind, und es gilt*

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

4.2.4 Satz (Monotonie) *Es seien $f, g : [a, b] \to \mathbb{R}$ integrierbar. Es sei $f(x) \leq g(x)$ für alle $x \in [a, b]$. Dann gilt*

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

4.2.5 Satz (Abschätzung) *Es sei $f : [a, b] \to \mathbb{R}$ integrierbar. Es sei M eine Schranke für f, d.h. $|f(x)| \leq M$ für alle $x \in [a, b]$. Dann gilt*

$$\left| \int_a^b f(x)dx \right| \leq (b - a)M.$$

4.3 Differentialrechnung

Die "Anderungsrate" einer Funktion beschreibt intuitiv, wie sich ein Funktionswert ändert, wenn sich die Stelle ändert. Es stellt sich heraus, daß viele Funktionen ein ähnliches Änderungsverhalten haben wie lineare Funktionen: Die Änderung des Funktionswertes ist, zumindest für kleine Änderungen, annähernd proportional zur Änderung der Stelle. Geometrisch entspricht diese Beobachtung der Anschauung, daß viele Funktionsgraphen eine Tangente haben. Im Berührpunkt schmiegt sich der Graph an die Tangente an, und der die Änderung des Funktionswertes entspricht dem Verhalten der linearen Funktion, die durch die Tangente gegeben ist. Differenzieren bedeutet also "Annähern einer Funktion durch eine lineare Funktion".

4.3.1 Definition Es sei $D \in \mathbb{R}$. Es sei $f : D \to \mathbb{R}$ stetig. Es sei $x \in D$ ein nicht isolierter Punkt (d.h. es gibt eine Folge in $D - \{x\}$, die gegen x konvergiert.) Dann heißt f an der Stelle x differenzierbar, wenn es eine stetige Funktion $g : D \to \mathbb{R}$ gibt, sodaß

$$f(y) - f(x) = (y - x)g(y)$$

für alle $x, y \in \mathbb{R}$ gilt.

Der Wert $g(x)$ wird als Ableitung, $f'(x)$, bezeichnet.

4.3.2 Beispiel Es sei $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$. Dann ist für beliebige x, y

$$y^2 - x^2 = (y - x)(y + x),$$

und wir können $g(y) := y + x$ wählen. Die Ableitung ist $f'(x) = x + x = 2x$.

4.3.3 Satz *Es sei $f : D \to \mathbb{R}$ differenzierbar bei $x \in D$. Die Funktion g ist dann eindeutig bestimmt. Wenn $(x_n)_n$ eine Folge ist, die gegen x konvergiert, so ist*

$$f'(x) = \lim_{n\to\infty} \frac{f(x_n) - f(x)}{x_n - x}.$$

4.3.4 Beispiel Es sei $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto e^x$. Die Ableitung bei $x = 0$ wird berechnet durch die Nullfolge $(\frac{1}{n})_n$:

$$\lim_{n\to\infty} \frac{e^{1/n} - e^0}{1/n} = \lim_{n\to\infty} \frac{e^{1/n} - 1}{1/n}.$$

Dieser Grenzwert muß numerisch berechnet werden, und es stellt sich heraus, daß er gleich 1 ist. In der Tat wird die Zahl e gerade so bestimmt, daß der Grenzwert 1 ist.

Für andere Stellen kann die Ableitung wie folgt hergeleitet werden:

$$\lim_{n\to\infty} \frac{e^{x+1/n} - e^x}{1/n} \lim_{n\to\infty} \frac{e^x(e^{1/n} - 1)}{1/n} = e^x \lim_{n\to\infty} \frac{e^{1/n} - 1}{1/n} = e^x.$$