Analysis 2 für Lehramt

SS 2011

Harald Woracek

Inhaltsverzeichnis

1 Topologie metrischer Räume **1**
- 1.1 Offene Mengen . . . 1
- 1.2 Kompaktheit . . . 8
- 1.3 Konvergenz gegen ∞ . . . 13

2 Reelle und komplexe Funktionen **21**
- 2.1 Stetigkeit . . . 21
- 2.2 Unstetigkeitsstellen reeller Funktionen . . . 28
- 2.3 Gleichmäßige Konvergenz . . . 32
- 2.4 Der Satz von Arzela-Ascoli . . . 39
- 2.5 Die Exponentialfunktion . . . 41
- 2.6 'Elementare' Funktionen . . . 44
- 2.7 Der Fundamentalsatz der Algebra . . . 47

3 Differentialrechnung **51**
- 3.1 Definition der Ableitung . . . 51
- 3.2 Eigenschaften der Ableitung . . . 55
- 3.3 Mittelwertsätze . . . 59
- 3.4 Der Taylorsche Lehrsatz . . . 67
- 3.5 Ein Fixpunktsatz . . . 76

4 Integralrechnung **83**
- 4.1 Flächenberechnungen . . . 83
- 4.2 Definition des Riemannschen Integrals . . . 86
- 4.3 Eigenschaften des Integrals . . . 95
- 4.4 Hauptsatz der Differential-Integralrechnung . . . 100
- 4.5 Eine Näherungsformel . . . 107
- 4.6 Differentialgeometrie . . . 111

Literaturverzeichnis **117**

Index **118**

Kapitel 1

Topologie metrischer Räume

1.1 Offene Mengen

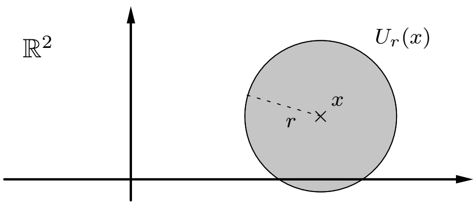
Sei $\langle X, d\rangle$ ein metrischer Raum, $x \in X$ und $r > 0$. Die Menge

$$U_r(x) := \big\{y \in X : d(y,x) < r\big\}$$

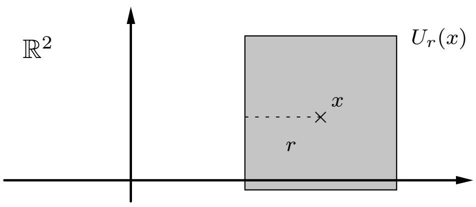
heißt *Kugel* mit Radius r und Mittelpunkt x.

1.1.1 Beispiel.

(*i*) Betrachte den Raum $\mathbb{R}^2$ mit der euklidischen Metrik. Dann ist die Kugel $U_r(x)$ tatsächlich gerade die Kreischeibe mit Radius r und Mittelpunkt x:

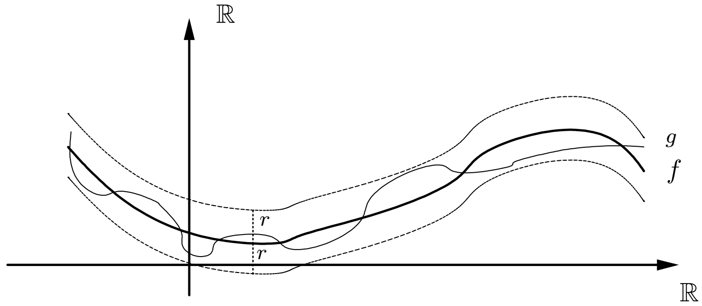


(*ii*) Betrachte den Raum $\mathbb{R}^2$ mit der Metrik $d_\infty(x,y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}$, $x = (x_1, x_2)$, $y = (y_1, y_2)$. Dann ist $U_r(x)$ das Quadrat mit Mittelpunkt x und Seitenlänge r:



(*iii*) Betrachte die Menge $\mathcal{B}(\mathbb{R})$ aller beschränkten Funktionen von $\mathbb{R}$ nach $\mathbb{R}$ versehen mit der Supremumsmetrik $d_\infty(f,g) := \sup_{x\in\mathbb{R}} |f(x) - g(x)|$. Die

Kugel $U_r(f)$ besteht dann aus allen Funktionen g die sich an keiner Stelle von f um mehr als um r unterscheiden, d.h. welche ganz in einem r-Schlauch um f verlaufen:



1.1.2 Definition. Sei $\langle X, d\rangle$ ein metrischer Raum. Eine Teilmenge $O \subseteq X$ heißt *offen*, wenn es zu jedem Element x von O eine ganze Kugel $U_r(x)$ gibt die in O enthalten ist. In Formeln ausgedrückt:

$$O \text{ offen } :\iff \quad \forall x \in O\, \exists r > 0 \,:\, U_r(x) \subseteq O\,.$$

Die Menge aller offenen Mengen von $\langle X, d\rangle$ bezeichnen wir mit $\mathcal{T}(X, d)$.

1.1.3 Beispiel. Sei $R > 0$ und $x_0 \in X$, dann ist die Kugel $U_R(x_0)$ offen. Um dies einzusehen, sei $x \in U_R(x_0)$ gegeben. Dann ist also $d(x, x_0) < R$. Setze $r := R - d(x, x_0)$ und betrachte die Kugel $U_r(x)$. Ist $y \in U_r(x)$, so gilt also $d(y, x) < r$, und daher

$$d(y, x_0) \leq d(y, x) + d(x, x_0) < (R - d(x, x_0)) + d(x, x_0) = R\,,$$

und damit $y \in U_R(x_0)$.

Wir fassen die grundlegenden Eigenschaften offener Mengen in der folgenden Proposition zusammen.

1.1.4 Proposition. *Sei $\langle X, d\rangle$ ein metrischer Raum. Dann gilt*

(O1) $\emptyset \in \mathcal{T}(X, d)$, $X \in \mathcal{T}(X, d)$;

(O2) *Ist $n \in \mathbb{N}$ und $O_1, \ldots, O_n \in \mathcal{T}(X, d)$, so folgt $\bigcap_{i=1}^n O_i \in \mathcal{T}(X, d)$;*

(O3) *Ist I eine beliebige Indexmenge und $O_i \in \mathcal{T}(X, d), i \in I$, so folgt $\bigcup_{i\in I} O_i \in \mathcal{T}(X, d)$.*

Eine weitere wichtige Eigenschaft ist

(T2) *Sind $x, y \in X$, $x \neq y$, so existieren offene Mengen O_x, O_y mit*

$$x \in O_x,\ y \in O_y \quad \textit{und} \quad O_x \cap O_y = \emptyset\,.$$

Beweis. Die leere Menge gehört zu $\mathcal{T}(X, d)$, denn sie enthält keine Punkte und erfüllt daher die offene Mengen definierende Bedingung trivialerweise. Für jedes $x \in X$ gilt offenbar $U_1(x) \subseteq X$, also ist auch X offen.

Seien nun $O_1, \ldots, O_n$ offen und sei $x \in O_1 \cap \ldots \cap O_n$. Dann gibt es $r_1, \ldots, r_n > 0$ mit $U_{r_i}(x) \subseteq O_i$, $i = 1, \ldots, n$. Setzt man $r := \min\{r_1, \ldots, r_n\}$, dann ist $r > 0$

und es gilt $U_r(x) \subseteq U_{r_i}(x)$, $i = 1, \dots, n$. Damit folgt $U_r(x) \subseteq U_{r_1} \cap \dots \cap U_{r_n} \subseteq O_1 \cap \dots \cap O_n$.

Seien O_i, $i \in I$, offen, und sei $x \in \bigcup_{i \in I} O_i$. Dann existiert $i_0 \in I$ mit $x \in O_{i_0}$. Wähle $r > 0$ mit $U_r(x) \subseteq O_{i_0}$, dann gilt auch $U_r(x) \subseteq \bigcup_{i \in I} O_i$.

Schliesslich seien zwei verschiedene Punkte $x, y \in X$ gegeben. Setze $r := \frac{1}{3} d(x,y)$ und $O_x := U_r(x)$, $O_y := U_r(y)$. Dann ist $r > 0$, also O_x, O_y offen. Angenommen es wäre $z \in O_x \cap O_y$, dann würde folgen

$$d(x,y) \leq d(x,z) + d(z,y) < 2r = \frac{2}{3} d(x,y)\,,$$

ein Widerspruch da $d(x,y) > 0$.

❑

Dual zum Konzept der offenen Menge ist das der abgeschlossenen Menge.

1.1.5 Definition. Sei $\langle X, d\rangle$ ein metrischer Raum. Eine Menge $A \subseteq X$ heißt *abgeschlossen*, wenn ihr Komplement $X \setminus A$ offen ist.

Geht man in Proposition 1.1.4 zu den Komplementen über, so erhält man:

1.1.6 Korollar. *Sei $\langle X, d\rangle$ ein metrischer Raum. Dann gilt*

- **(A1)** *$\emptyset, X$ sind abgeschlossen;*
- **(A2)** *Sind $n \in \mathbb{N}$ und $A_1, \dots, A_n$ abgeschlossen, so folgt das auch $\bigcup_{i=1}^n A_i$ abgeschlossen ist;*
- **(A3)** *Ist I eine beliebige Indexmenge und sind alle Mengen $A_i, i \in I$, abgeschlossen, so folgt das auch $\bigcap_{i \in I} A_i$ abgeschlossen ist.*

Weiters gilt

- **(T1)** *Für jedes $x \in X$ ist die einpunktige Menge $\{x\}$ abgeschlossen.*

Beweis. Es gilt $\emptyset^c = X$, $X^c = \emptyset$, sowie

$$\Big(\bigcup_{i=1}^n A_i\Big)^c = \bigcap_{i=1}^n (A_i^c),\ \ \Big(\bigcap_{i \in I} A_i\Big)^c = \bigcup_{i \in I} (A_i)^c\,.$$

Weiters existiert zu $y \neq x$ stets $r > 0$ sodass $x \notin U_r(y)$, z.B. $r := \frac{1}{2} d(x,y)$. Also ist $X \setminus \{x\}$ offen.

❑

Eine weitere allgemeine Begriffsbildung ist die des Häufungspunktes einer Menge.

1.1.7 Definition. Sei $\langle X, d\rangle$ ein metrischer Raum, $E \subseteq X$ und $x \in X$. Dann heißt x ein *Häufungspunkt* von E, wenn jede Kugel $U_r(x)$ einen von x verschiedenen Punkt aus E enthält. In Formeln ausgedrückt:

$$x \text{ Häufungspunkt von } E \ :\Longleftrightarrow\ \ \forall r > 0:\ E \cap \big(U_r(x) \setminus \{x\}\big) \neq \emptyset$$

1.1.8 Proposition. *Sei $\langle X, d\rangle$ ein metrischer Raum und sei $E \subseteq X$.*

(i) *Sei $E \subseteq X$ und $x \in X$. Ist x ein Häufungspunkt von E, so enthält jede Kugel $U_r(x)$ eine Folge von paarweise verschiedenen Punkten der Menge E welche gegen x konvergiert.*

(ii) *Sei $(x_n)_{n\in\mathbb{N}}$ eine konvergente Folge von Punkten aus X, $x := \lim_{n\to\infty} x_n$. Dann ist die Menge $E := \{x_n : n \in \mathbb{N}\}$ entweder endlich, oder sie hat genau einen Häufungspunkt, nämlich den Grenzwert x. Insbesondere ist $x \in E$ oder x ein Häufungspunkt von E.*

Beweis. Um die erste Behauptung zu zeigen sei eine Kugel $U_r(x)$ gegeben. Wir konstruieren rekursiv Folgen $(x_n)_{n\in\mathbb{N}}$ und $(r_n)_{n\in\mathbb{N}}$ mit $x_n \neq x$, $r_n > 0$.

Setze $r_1 := \min\{r, 1\}$. Da x ein Häufungspunkt von E ist, ist $E \cap (U_{r_1}(x) \setminus \{x\}) \neq \emptyset$. Wähle $x_1 \in E \cap (U_{r_1}(x) \setminus \{x\})$. Angenommen es sind schon Punkte $x_1, \dots, x_n$ und Radien $r_1, \dots, r_n$ konstruiert. Setze $r_{n+1} := \min\{d(x_1, x), \dots, d(x_n, x), \frac{1}{n+1}\}$. Dann ist $r_{n+1} > 0$. Da x ein Häufungspunkt von E ist, ist $E \cap (U_{r_{n+1}}(x) \setminus \{x\}) \neq \emptyset$. Wähle $x_{n+1} \in E \cap (U_{r_1}(x) \setminus \{x\})$.

Die Punkte x_n sind alle verschieden, denn für $m < n$ gilt stets $d(x_n, x) < d(x_m, x)$. Weiters gilt $d(x_n, x) < r_n \leq \frac{1}{n}$, also ist $x_n \to x$.

Um die zweite Behauptung einzusehen, zeigen wir zuerst, dass die Menge E nur endlich viele Elemente hat genau dann wenn die Folge $(x_n)_{n\in\mathbb{N}}$ ab einem Index konstant gleich x ist: Ist die Folge ab einem Index N konstant, so ist klarerweise $E = \{x_1, \dots, x_N\}$. Sei umgekehrt E endlich. Dann ist

$$\inf\{d(x_n, x) : n \in \mathbb{N}, x_n \neq x\} = \min\{d(x_n, x) : x_n \in E, x_n \neq x\} > 0\,.$$

Wähle ϵ positiv und kleiner als dieses Minimum. Da $x_n \to x$, existiert ein $N \in \mathbb{N}$ mit $d(x_n, x) < \epsilon$ für alle $n \geq N$. Da ϵ kleiner als obiges Minimum ist, folgt $x_n = x$ für $n \geq N$.

Sei nun angenommen, dass die Menge E nicht endlich ist, d.h. die Folge $(x_n)_{n\in\mathbb{N}}$ nicht ab einem Index konstant. Sei eine Kugel $U_r(x)$ gegeben, dann existiert $N \in \mathbb{N}$ sodaß $x_n \in U_r(x)$ für alle $n \geq N$. Da nicht alle x_n, $n \geq N$, gleich x sein können, enthält also $U_r(x) \cap E$ einen von x verschiedenen Punkt. Wir sehen, dass x ein Häufungspunkt von E ist. Sei nun y irgendein Häufungspunkt der Menge E. Dann existiert nach dem bereits bewiesenen ersten Punkt eine Folge $(x_{n(k)})_{k\in\mathbb{N}}$ von Punkten aus E die gegen y konvergiert. Nun ist aber, da $x_n \to x$, auch jede Teilfolge konvergent gegen den selben Grenzwert x. Da Grenzwerte eindeutig sind, muß $y = x$ sein.

❑

1.1.9 Proposition. *Die folgenden drei Aussagen sind äquivalent:*

(i) *E ist abgeschlossen.*

(ii) *Jeder Häufungspunkt von E liegt in E.*

(iii) *Für jede konvergente Folge $(x_n)_{n\in\mathbb{N}}$ von Punkten aus E, liegt der Grenzwert in E.*

Beweis.
$(i) \Rightarrow (ii)$: Sei E abgeschlossen und $x \notin E$. Die Menge E^c ist offen, also existiert $r > 0$ mit $U_r(x) \subseteq E^c$, d.h. $U_r(x) \cap E = \emptyset$. Damit kann x kein Häufungspunkt

von E sein.
$(ii) \Rightarrow (iii)$: Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge von Punkten aus E mit $x_n \to x$. Dann ist nach Proposition 1.1.8, (ii), also $x \in E$ oder zumindest ein Häufungspunkt von E. In zweitem Fall haben wir aber nach der Voraussetzung (ii) auch schon $x \in E$.
$(iii) \Rightarrow (i)$: Sei $x \notin E$. Angenommen es gilt für jedes $n \in \mathbb{N}$ das $U_{\frac{1}{n}}(x) \cap E \neq \emptyset$. Dann wähle $x_n \in U_{\frac{1}{n}}(x) \cap E$ und betrachte die Folge $(x_n)_{n\in\mathbb{N}}$. Wegen $d(x_n, x) < \frac{1}{n}$ gilt $x_n \to x$, also ist nach der Voraussetzung (iii) schon $x \in E$. Ein Widerspruch. Wir sehen, dass es ein $n_0 \in \mathbb{N}$ geben muss mit $U_{\frac{1}{n_0}}(x) \cap E = \emptyset$, d.h. mit $U_{\frac{1}{n_0}}(x) \subseteq E^c$. Es folgt, dass E^c offen ist.

❑

1.1.10 Definition. Sei $\langle X, d\rangle$ ein metrischer Raum und sei $E \subseteq X$. Dann heißt E *dicht* in X, wenn jeder Punkt von X zu E gehört oder ein Häufungspunkt von E ist.

1.1.11 Proposition. *Sei $\langle X, d\rangle$ ein metrischer Raum und sei $E \subseteq X$. Die folgenden Aussagen sind äquivalent:*

- (i) *E ist dicht in X;*
- (ii) *Für jeden Punkt $x \in X$ gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ von Punkten aus E mit $x_n \to x$.*
- (iii) *Für jede nichtleere offene Menge O gilt $O \cap E \neq \emptyset$.*

Beweis.
$(i) \Rightarrow (ii)$: Ist $x \in E$, so wähle die konstante Folge $x_n := x$, $n \in \mathbb{N}$. Ist $x \notin E$, so muß x ein Häufungspunkt von E sein, und nach Proposition 1.1.8, (i), gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$, $x_n \in E$, mit $x_n \to x$.
$(ii) \Rightarrow (iii)$: Sei O offen und nichtleer. Wähle $x \in O$ und $r > 0$ sodass $U_r(x) \subseteq O$. Wähle eine Folge $(x_n)_{n\in\mathbb{N}}$, $x_n \in E$, mit $x_n \to x$. Dann gilt für alle hinreichend grossen Indizes n stets $x_n \in U_r(x)$. Insbesondere ist $E \cap O \neq \emptyset$.
$(iii) \Rightarrow (i)$: Sei $x \in X$ gegeben. Für jedes $r > 0$ ist die Kugel $U_r(x)$ offen und nichtleer. Also gilt stets $E \cap U_r(x) \neq \emptyset$. Gibt es ein r mit $x \in E \cap U_r(x)$, so ist insbesondere $x \in E$. Andernfalls ist für jedes $r > 0$ die Menge $E \cap (U_r(x) \setminus \{x\})$ nichtleer, und daher x ein Häufungspunkt von E.

❑

1.1.12 Beispiel. Betrachte $\mathbb{R}^n$ versehen mit der euklidischen Metrik. Dann ist die Menge $\mathbb{Q}^n$ aller Vektoren mit rationalen Koordinaten dicht in $\mathbb{R}^n$.

Um dies einzusehen, betrachte zuerst den Fall $n = 1$. Sei $x \in \mathbb{R} \setminus \mathbb{Q}$ gegeben. Zu $n \in \mathbb{N}$ gibt es eine rationale Zahl x_n mit $x - \frac{1}{n} < x_n < x + \frac{1}{n}$, d.h. mit $|x_n - x| < \frac{1}{n}$. Es ist also $x_n \to x$, und wir erhalten dass x ein Häufungspunkt von $\mathbb{Q}$ ist.

Sei nun $(y_1, \ldots, y_n) \in \mathbb{R}^n$ gegeben. Wähle Folgen rationaler Zahlen $(x_{i,k})_{k\in\mathbb{N}}$ mit $\lim_{k\to\infty} x_{i,k} = y_i$, $i = 1, \ldots, n$. Dann folgt

$$\lim_{n\to\infty} (x_{1,k}, \ldots, x_{n,k}) = (y_1, \ldots, y_n)\,.$$

❑

1.1.13 Definition. Sei $\langle X, d\rangle$ ein metrischer Raum. Eine Teilmenge $E \subseteq X$ heißt *zusammenhängend*, wenn sie die folgende Eigenschaft hat: Sind $A, B \subseteq X$ offen und disjunkt, und gilt $E \subseteq A \cup B$, so folgt dass schon $E \subseteq A$ oder $E \subseteq B$.

Wir wollen herausfinden, welche Teilmengen von $\mathbb{R}$ zusammenhängend sind. Dazu zur Wiederholung die Definition eines *Intervalls* ($a, b \in \mathbb{R}$, $a \leq b$):

$$(a,b) := \{x \in \mathbb{R} : a < x < b\}$$
$$[a,b) := \{x \in \mathbb{R} : a \leq x < b\}$$
$$(-\infty,b) := \{x \in \mathbb{R} : x < b\}$$
$$(-\infty,b] := \{x \in \mathbb{R} : x \leq b\}$$
$$[a,b] := \{x \in \mathbb{R} : a \leq x \leq b\}$$
$$(a,b] := \{x \in \mathbb{R} : a < x \leq b\}$$
$$(a,\infty) := \{x \in \mathbb{R} : a < x\}$$
$$[a,\infty) := \{x \in \mathbb{R} : a \leq x\}$$
$$(-\infty,\infty) := \mathbb{R}$$

1.1.14 Proposition. *Sei $E \subseteq \mathbb{R}$. Dann ist E zusammenhängend, genau dann wenn E ein Intervall ist.*

Beweis.

Schritt 1: Wir zeigen, dass E genau dann zusammenhängend ist, wenn gilt: Sind $x, y \in E$, $x < y$, so folgt $[x, y] \subseteq E$.

Sei angenommen es existieren $x, y \in E$, $x < y$, sodaß $[x, y] \nsubseteq E$. Wähle $z \in [x, y] \setminus E$ und setze $A := (-\infty, z), B := (z, +\infty)$. Dann sind A, B offen, disjunkt, und es gilt $E \subseteq A \cup B$. Wegen $x \in A$ und $y \in B$ ist aber weder $E \subseteq A$ noch $E \subseteq B$. Also ist E nicht zusammenhängend.

Sei umgekehrt E nicht zusammenhängend. Dann gibt es offene und disjunkte Menge A, B mit $E \subseteq A \cup B$ aber $E \nsubseteq A$ und $E \nsubseteq B$. Wähle $x \in A \cap E$ und $y \in B \cap E$ und sei oBdA angenommen, dass $x < y$. Da A und B offen sind, enthält A sowie auch B mit jedem Punkt z noch ein ganzes offenes Intervall um z. Betrachte den Punkt $z_0 := \sup(A \cap [x, y])$. Da B ein ganzes Intervall um y enthält, und $A \cap B = \emptyset$ ist, folgt $z_0 < y$. Wäre $z_0 \in A$, so würde A noch ein ganzes Intervall um z_0 enthalten, ein Widerspruch da z_0 eine obere Schranke von $A \cap [x, y]$ ist. Wäre $z_0 \in B$, so würde B noch ein ganzes Intervall um z_0 enthalten, und da $A \cap B = \emptyset$ könnten wir eine kleinere obere Schranke von $A \cap [x, y]$ als z_0 finden, ein Widerspruch. Wir sehen, dass $z_0 \notin A \cup B$ und daher auch $z_0 \notin E$.

Schritt 2: Wir zeigen, dass es genau die Intervalle sind welche der obigen Eigenschaft genügen.

Zunächst enthält jedes Intervall offensichtlich mit je zwei Punkten x, y auch ganz $[x, y]$. Umgekehrt betrachte zum Beispiel den Fall das E nach oben beschränkt ist, aber nicht nach unten beschränkt. Setze $b := \sup E$, dann ist b eine obere Schranke von E, also $E \subseteq (-\infty, b]$. Sei $z < b$. Dann existieren $x, y \in E$ mit $x < z$ und $z < y$. Es folgt $z \in [x, y] \subseteq E$. D.h. wir haben $(-\infty, b) \subseteq E$. Je nachdem ob $b \in E$ oder $b \notin E$, ist also $E = (-\infty, b]$ oder $E = (-\infty, b)$.

Alle anderen Fälle behandelt man genauso.

❑

1.1.15 Definition. Seien $\langle X_1, d_1\rangle$ und $\langle X_2, d_2\rangle$ metrische Räume, und $\phi : X_1 \to X_2$. Dann heißt ϕ *isometrisch*, wenn

$$d_2\big(\phi(x), \phi(y)\big) = d_1(x, y),\ x, y \in X_1 .$$

1.1.16 Proposition. *Seien $\langle X_1, d_1\rangle$ und $\langle X_2, d_2\rangle$ metrische Räume, und sei $\phi : X_1 \to X_2$ bijektiv und isometrisch. Dann gilt:*

(i) $\phi^{-1} : X_2 \to X_1$ *ist bijektiv und isometrisch;*

(ii) $\phi(U_r^1(x)) = U_r^2(\phi(x))$*;*

(iii) $O \subseteq X_1$ *ist offen in* X_1*, genau dann wenn* $\phi(O)$ *offen in* X_2 *ist;*

(iv) $A \subseteq X_1$ *ist abgeschlossen in* X_1*, genau dann wenn* $\phi(A)$ *abgeschlossen in* X_2 *ist;*

(v) $E \subseteq X_1$ *ist zusammenhängend in* X_1*, genau dann wenn* $\phi(E)$ *zusammenhängend in* X_2 *ist;*

(vi) $E \subseteq X_1$ *ist dicht in* X_1*, genau dann wenn* $\phi(E)$ *dicht in* X_2 *ist.*

(vii) Eine Folge $(x_n)_{n\in\mathbb{N}}$ *in* X_1 *ist genau dann konvergent (eine Cauchy-Folge), wenn die Folge* $(\phi(x_n))_{n\in\mathbb{N}}$ *in* X_2 *konvergent (eine Cauchy-Folge) ist. Im Fall der Konvergenz gilt*

$$\lim_{n\to\infty} \phi(x_n) = \phi\big(\lim_{n\to\infty} x_n\big)\,.$$

Beweis.
`ad`(i): Seien $x_2, y_2 \in X_2$ gegeben, dann gilt

$$d_1\big(\phi^{-1}(x_2), \phi^{-1}(y_2)\big) = d_1\big(\phi(\phi^{-1}(x_2)), \phi(\phi^{-1}(y_2))\big) = d_2(x_2, y_2)\,.$$

`ad`(ii): Sei $y \in U_r^1(x)$, d.h. $d_1(y,x) < r$. Dann gilt

$$d_2\big(\phi(y), \phi(x)\big) = d_1(y,x) < r\,,$$

also ist $\phi(y) \in U_r^2(\phi(x))$. Umgekehrt sei $z \in U_r^2(\phi(x))$. Setze $y := \phi^{-1}(z)$, dann ist

$$d_1(y,x) = d_2\big(\phi(y), \phi(x)\big) = d_2\big(z, \phi(x)\big) < r\,,$$

also $y \in U_r^1(x)$. Offenbar gilt $\phi(y) = z$.

`ad`(iii): Sei O offen und $x \in \phi(O)$. Dann ist $\phi^{-1}(x) \in O$ und daher existiert $r > 0$ mit $U_r^2(\phi(x)) \subseteq O$. Es folgt

$$U_r^2(x) = \phi\big(U_r^1(\phi^{-1}(x))\big) \subseteq \phi(O)\,.$$

Also ist $\phi(O)$ offen.

Da ϕ^{-1} ebenfalls bijektiv und isometrisch ist, erhält man aus der Voraussetzung „$\phi(O)$ offen" dass $O = \phi^{-1}(\phi(O))$ offen ist.

`ad`(iv): Es gilt $\phi(A^c) = \phi(A)^c$.

`ad`(v): Sind A, B offen und disjunkt mit $E \subseteq A \cup B$, so sind auch $\phi(A), \phi(B)$ offen und es gilt $\phi(E) \subseteq \phi(A) \cup \phi(B)$. Umgekehrt wende man das gleiche Argument mit ϕ^{-1} an.

`ad`(vi): Ist O offen und nichtleer, so ist auch $\phi(O)$ offen und nichtleer. Weiters gilt $\phi(E \cap O) = \phi(E) \cap \phi(O)$. Umgekehrt wende das gleiche Argument mit ϕ^{-1} an.

`ad`(vii): Es gilt stets dass $d_1(x_n, x_m) = d_2(\phi(x_n), \phi(x_m))$ bzw. $d_1(x_n, x) = d_2(\phi(x_n), \phi(x))$.

❑

1.2 Kompaktheit

1.2.1 Definition. Sei $\langle X, d\rangle$ ein metrischer Raum und sei $E \subseteq X$. Eine *offene Überdeckung* von E ist eine Familie $\mathcal{G} = \{G_i : i \in I\}$ offener Mengen mit

$$\bigcup_{i\in I} G_i \supseteq E.$$

Sind $\mathcal{G}'$ und $\mathcal{G}$ offene Überdeckungen von E und gilt $\mathcal{G}' \subseteq \mathcal{G}$, dann bezeichnet man $\mathcal{G}'$ als *Teilüberdeckung* von $\mathcal{G}$.

1.2.2 Definition. Sei $\langle X, d\rangle$ ein metrischer Raum und sei $K \subseteq X$. Dann heißt K *kompakt*, wenn jede offene Überdeckung von K eine endliche Teilüberdeckung besitzt, d.h. eine die nur aus endlich vielen Mengen besteht.

Wir sammeln einige Aussagen über kompakte Mengen.

1.2.3 Proposition. *Sei $\langle X, d\rangle$ ein metrischer Raum. Dann gilt:*

- (i) *Ist $K \subseteq X$ kompakt, und $F \subseteq X$ abgeschlossen, dann ist auch $K \cap F$ kompakt.*
- (ii) *Ist $K \subseteq X$ kompakt, dann ist K auch abgeschlossen und beschränkt.*
- (iii) *Sei $(K_n)_{n\in\mathbb{N}}$, $K_n \subseteq X$, eine Folge nichtleerer kompakter Mengen mit $K_n \supseteq K_{n+1}$. Dann ist $\bigcap_{n\in\mathbb{N}} K_n \neq \emptyset$.*
- (iv) *Sei $K \subseteq X$ kompakt, dann hat jede unendliche Teilmenge von K einen Häufungspunkt der in K liegt.*
- (v) *Sei $K \subseteq X$ kompakt, dann hat jede Folge von Elementen von K eine konvergente Teilfolge deren Grenzwert in K liegt.*
- (vi) *Ist X kompakt, so ist $\langle X, d\rangle$ auch vollständig.*

Beweis.
`ad`(i): Sei K kompakt, F abgeschlossen, und sei $\{G_i : i \in I\}$ eine offene Überdeckung von $K\cap F$. Betrachte die Familie $\{G_i : i \in I\}\cup\{F^c\}$. Diese ist dann eine offene Überdeckung von K, und enthält daher eine endliche Teilüberdeckung, welche aus $G_{i_1}, \dots, G_{i_n}$ und unter Umständen noch F^c besteht. Dann ist aber $G_{i_1}, \dots, G_{i_n}$ eine Überdeckung von $K \cap F$.

`ad`(ii): Sei $K \subseteq X$ nicht abgeschlossen. Dann gibt es also einen Punkt $x \in X$ der zwar Häufungspunkt von K ist, aber nicht zu K gehört. Für $y \in K$ wähle Kugeln U_y, V_y, mit $y \in U_y, x \in V_y$, und $U_y \cap V_y = \emptyset$. Dann ist $\mathcal{G} := \{U_y : y \in K\}$ eine offene Überdeckung von K. Für je endlich viele $U_{y_1}, \dots, U_{y_n}$ gilt

$$\Big[U_{y_1} \cup \ldots \cup U_{y_n}\Big] \cap \Big[V_{y_1} \cap \ldots \cap V_{y_n}\Big] = \emptyset\,.$$

Da x Häufungspunkt von K ist und $V_{y_1} \cap \ldots \cap V_{y_n}$ eine gewisse Kugel mit Mittelpunkt x enthält, ist $[V_{y_1} \cap \ldots \cap V_{y_n}] \cap K \neq \emptyset$, also folgt

$$U_{y_1} \cup \ldots \cup U_{y_n} \not\supseteq K\,.$$

Die Überdeckung $\mathcal{G}$ kann also keine endliche Teilüberdeckung besitzen, d.h. K ist nicht kompakt.

Sei K kompakt. Dann hat die offene Überdeckung $\{U_1(y) : y \in K\}$ eine endliche Teilüberdeckung, also $K \subseteq U_1(y_1) \cup \ldots \cup U_1(y_n)$. Dann ist aber $K \subseteq U_R(y_1)$ mit $R := 1 + \max_{i=2,\ldots,n} d(y_i, y_1)$.

`ad`(iii): Angenommen $\bigcap_{n\in\mathbb{N}} K_n = \emptyset$. Geht man zu den Komplementen über, erhält man $\bigcup_{n\in\mathbb{N}} K_n^c = X$. Insbesondere ist $\mathcal{G} := \{K_n^c : n \in \mathbb{N}\}$ eine Überdeckung von K_1. Wegen dem bereits bewiesenen Punkt (ii), sind alle Mengen K_n^c offen. Da K_1 kompakt ist, gibt es $K_{i_1}, \ldots, K_{i_n}$ mit $\bigcup_{k=1}^n K_{i_k}^c \supseteq K$. Nun gilt $K_n^c \subseteq K_{n+1}^c$, also gilt

$$\bigcup_{k=1}^{n} K_{i_k}^c = K_i^c \text{ mit } i := \max\{i_1, \ldots, i_n\}\,.$$

Wir erhalten $K_i^c \supseteq K_1$, also $K_1 \cap K_i = \emptyset$. Wegen $K_i \subseteq K_1$, impliziert dies $K_i = \emptyset$, ein Widerspruch.

`ad`(iv): Sei K kompakt. Wir nehmen indirekt an, dass es eine unendliche Teilmenge E von K gibt, die keinen Häufungspunkt in K hat. Sei $x \in K$, $x \notin E$. Dann existiert eine offene Kugel G_x mit Mittelpunkt x mit $G_x \cap E = \emptyset$. Für $y \in E$ existiert eine offene Kugel H_y mit Mittelpunkt y mit $H_y \cap E = \{y\}$. Die Familie

$$\mathcal{G} := \{H_y : y \in E\} \cup \{G_x : x \in K \setminus E\}$$

ist eine offene Überdeckung von K. Keine der Mengen H_y kann weggelassen werden, denn keine der anderen Mengen enthält den Punkt y. Da E unendlich ist, kann man keine endliche Überdeckung auswählen. Ein Widerspruch. Also hat E einen Häufungspunkt. Da $E \subseteq K$, ist dieser auch ein Häufungspunkt von K, und liegt da K abgeschlossen ist daher in K.

`ad`(v): Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge von Punkten aus K. Ist die Menge $E := \{x_n : n \in \mathbb{N}\}$ endlich, so muß eines ihrer Elemente, z.B. x_{n_0}, in der Folge $(x_n)_{n\in\mathbb{N}}$ unendlich oft vorkommen. Wir können also eine Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ auswählen mit $x_{n_k} = x_{n_0}$, $k \in \mathbb{N}$. Diese ist trivialerweise konvergent, und ihr Grenzwert ist gleich x_{n_0}, liegt also in K.

Betrachte nun den Fall, dass die Menge E unendlich viele Elemente hat. Da K kompakt ist, besitzt E einen Häufungspunkt der in K liegt. Nach Proposition 1.1.8, (i), können wir eine Folge von paarweise verschiedenen Elementen von E finden die gegen x konvergiert. Diese können wir zu einer Teilfolge von $(x_n)_{n\in\mathbb{N}}$ anordnen.

`ad`(vi): Sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in X. Nach (v) hat sie eine konvergente Teilfolge, und ist daher schon selbst konvergent.

❑

1.2.4 Proposition. *Seien $\langle X_1, d_1\rangle$ und $\langle X_2, d_2\rangle$ metrische Räume, und sei $\phi : X_1 \to X_2$ bijektiv und isometrisch. Weiters sei $K \subseteq X_1$. Dann ist K kompakt in X_1 genau dann wenn $\phi(K)$ kompakt in X_2 ist.*

Beweis. Sei vorausgesetzt, dass K kompakt ist. Sei $\mathcal{G}^2 := \{G_i^2 : i \in I\}$ eine offene Überdeckung von $\phi(K)$. Setze $G_i^1 := \phi^{-1}(G_i^2)$, dann ist $\mathcal{G}^1 := \{G_i^1 : i \in$

$I\}$ eine offene Überdeckung von $\phi^{-1}(\phi(K)) = K$. Daher existiert eine endliche Teilüberdeckung, $K \subseteq G^1_{i_1} \cup \ldots \cup G^1_{i_n}$. Es folgt, dass

$$\phi(K) \subseteq \phi(G^1_{i_1}) \cup \ldots \cup \phi(G^1_{i_n}) = G^2_{i_1} \cup \ldots \cup G^2_{i_n}$$

Setzt man umgekehrt voraus, dass $\phi(K)$ kompakt ist, so folgt dass K kompakt ist, indem man das obige Argument auf ϕ^{-1} anwendet.

❑

Wir betrachten die folgenden Eigenschaften einer Menge $K \subseteq X$:

(AB) E ist abgeschlossen und beschränkt.

(HP) Jede unendliche Teilmenge E von K hat einen Häufungspunkt der in K liegt.

(TF) Jede Folge von Elementen aus K hat eine konvergente Teilfolge deren Grenzwert in K liegt.

Wie wir in Proposition 1.2.3 gesehen haben, ist jede dieser Eigenschaften schwächer als Kompaktheit, d.h. stets gilt

$$K \text{ kompakt } \Rightarrow K \text{ erfüllt (AB), (HP), (TF)}$$

Im Allgemeinen gilt keine der umgekehrten Implikationen. Es ist eine wichtige Tatsache, die wir im folgenden beweisen werden, dass für den speziellen Fall des Raumes $\mathbb{R}^n$ die Umkehrungen doch gelten.

1.2.5 Satz. *Sei K eine Teilmenge des $\mathbb{R}^n$ wobei dieser mit der euklidischen Metrik versehen ist. Dann sind äquivalent:*

(*i*) *K ist kompakt;*

(*ii*) *Jede unendliche Teilmenge E von K hat einen Häufungspunkt der in K liegt;*

(*iii*) *E ist abgeschlossen und beschränkt;*

(*iv*) *Jede Folge von Elementen aus K hat eine Teilfolge die gegen ein Element von K konvergiert.*

Die Äquivalenz von (*i*) und (*ii*) bezeichnet man auch als den *Satz von Heine*[1]*-Borel*[2]. Bevor wir zum Beweis von Satz 1.2.5 kommen, wollen wir uns überlegen, dass Mengen die (HP) erfüllen ganz ähnliche Eigenschaften haben wie kompakte Mengen.

1.2.6 Lemma. *Es gilt:*

(*i*) *Hat $K \subseteq X$ die Eigenschaft* (HP)*, und ist $F \subseteq X$ abgeschlossen, dann hat auch $K \cap F$ die Eigenschaft* (HP)*.*

(*ii*) *Hat $K \subseteq X$ die Eigenschaft* (HP)*, dann ist K auch abgeschlossen und beschränkt.*

[1] Heinrich Eduard Heine. 16.3.1821 Berlin - 21.10.1881 Halle
[2] Emile Borel. 7.1.1871 Saint-Affrique (Aveyron) - 3.2.1956 Paris

(*iii*) *Sei* $(K_n)_{n\in\mathbb{N}}$, $K_n \subseteq X$, $K_n \supseteq K_{n+1}$, *eine absteigende Folge nichtleerer Mengen die alle die Eigenschaft* (HP) *haben. Dann ist* $\bigcap_{n\in\mathbb{N}} K_n \neq \emptyset$.

(*iv*) *Hat* $K \subseteq X$ *die Eigenschaft* (HP), *dann hat jede Folge von Elementen von* K *eine konvergente Teilfolge deren Grenzwert in* K *liegt.*

Beweis.
ad(i): Sei $E \subseteq K \cap F$ eine unendliche Teilmenge. Dann ist $E \subseteq K$, also hat E einen Häufungspunkt x der in K liegt. Nun ist aber auch $E \subseteq F$. Ein Häufungspunkt von E ist also erst recht ein Häufungspunkt von F. Da F abgeschlossen ist, enthält F alle seine Häufungspunkte, also ist $x \in F$. Insgesamt folgt $x \in K \cap F$.

ad(ii): Sei $(x_n)_{n\in\mathbb{N}}$ eine konvergente Folge von Punkten aus K, $x := \lim_{n\to\infty} x_n$. Betrachte die Menge $E := \{x_n : n \in \mathbb{N}\}$. Ist diese Menge endlich, so ist die Folge ab einem Index konstant, und daher konstant gleich ihrem Grenzwert. Es folgt $x \in K$. Ist diese Menge unendlich, so besitzt sie einen Häufungspunkt der in K liegt. Nun hat E nach Proposition 1.1.8 genau einen Häufungspunkt, nämlich x. Auch in diesem Fall erhalten wir also $x \in K$. Es folgt das K abgeschlossen ist.

Angenommen K ist unbeschränkt. Wähle $x_0 \in X$ und Punkte $x_n \in K$ mit $d(x_n, x_0) \geq n$. Dann ist die Menge $E := \{x_n : n \in \mathbb{N}\}$ eine unendliche Teilmenge von K. Ist $y \in X$ und $r > 0$, so kann die Kugel $U_r(y)$ nur endlich viele Elemente von E enthalten. Also ist y kein Häufungspunkt von E. Wir sehen, dass E keine Häufungspunkte haben kann, ein Widerspruch.

ad(iii): Ist eine der Mengen K_n, zum Beispiel K_{n_0}, endlich, so muß ab einem gewissen Index N gelten das $K_n = K_N$, $n \geq N$. Denn die Menge K_{n_0} kann nur endlich oft echt verkleinert werden bevor sie leer ist. Also folgt $\bigcap_{n\in\mathbb{N}} K_n = K_N \neq \emptyset$.

Sei nun angenommen, dass jede Menge K_n unendlich ist. Wir wählen eine Folge $E = \{x_1, x_2, x_3, \ldots\}$ von paarweise verschiedenen Punkten wie folgt: $x_1 \in K_1$, $x_2 \in K_2$ mit $x_2 \neq x_1$, $x_3 \in K_3$ mit $x_3 \neq x_1, x_2$, usw. Eine solche Wahl ist möglich, da ja jedes K_n unendlich viele Punkte enthält. Dann hat E als unendliche Teilmenge von K_1 einen Häufungspunkt x.

Wäre $x \notin \bigcap_{n\in\mathbb{N}} K_n$, d.h. $x \notin K_N$ für ein gewisses N, so existierte, da K_N abgeschlossen ist, eine Kugel $U_r(x)$ mit $U_r(x) \cap K_N = \emptyset$. Wegen $x_n \in K_N$ für $n \geq N$ kann $U_r(x)$ nur endlich viele Punkte von E enthalten, ein Widerspruch zu Proposition 1.1.8.

ad(iv): Wir haben im Beweis von Proposition 1.2.3, (v), nur die Eigenschaft (HP) benützt.

❑

Weiters benötigen wir noch die folgende Aussage:

1.2.7 Lemma. *Es existiert eine Folge* $(V_n)_{n\in\mathbb{N}}$ *von offenen Mengen des* $\mathbb{R}^n$ *mit der folgenden Eigenschaft: Für jeden Punkt* $x \in \mathbb{R}^n$ *und jede offene Menge* G *mit* $x \in G$, *gibt es ein* $n \in \mathbb{N}$ *mit* $x \in V_n \subseteq G$.

Beweis. Betrachte die Menge aller Kugeln $U_r(y)$ mit $r \in \mathbb{Q}$ und $y \in \mathbb{Q}^n$.

Sei $x \in \mathbb{R}^n$ und G offen mit $x \in G$ gegeben. Dann existiert eine gewisse Kugel $U_{r_0}(x)$ mit $U_{r_0}(x) \subseteq G$. Da $\mathbb{Q}^n$ dicht ist, existiert ein Punkt $y \in \mathbb{Q}^n \cap U_{\frac{r_0}{2}}(x)$, d.h. $d(y,x) < \frac{r_0}{2}$. Wähle $r \in \mathbb{Q}$ mit $d(y,x) < r < \frac{r_0}{2}$. Dann ist $x \in U_r(y)$. Weiters gilt für jedes $z \in U_r(y)$

$$d(z,x) \le d(z,y) + d(y,x) < r + \frac{r}{2} < \frac{r_0}{2} + \frac{r_0}{2} = r_0 \,.$$

Insgesamt haben wir $x \in U_r(y) \subseteq U_{r_0}(x) \subseteq G$.

Betrachte die Menge $\mathbb{Q} \times \mathbb{Q}^n$. Wir wissen, dass man eine Bijektion finden kann von $\mathbb{N}$ auf $\mathbb{Q}$. Diese hatten wir durch „abzählen längs Diagonalen" erhalten. Genauso erhält man, wiederum durch diagonales abzählen, eine Bijektion von $\mathbb{N}$ auf $\mathbb{Q} \times \mathbb{Q}^n$. Das heißt also, dass man die obigen Mengen $U_r(y)$ als eine Folge $(V_n)_{n\in\mathbb{N}}$ anordnen kann.

❑

Beweis. (von Satz 1.2.5) Die Implikationen $(i) \Rightarrow (ii) \wedge (iii) \wedge (iv)$ und $(ii) \Rightarrow (iii) \wedge (iv)$ sind schon bekannt. Wir zeigen $(iii) \Rightarrow (ii) \Rightarrow (i)$ und $(iv) \Rightarrow (ii)$.

Schritt 1, $(iii) \Rightarrow (ii)$: Sei K abgeschlossen und beschränkt, und sei E eine unendliche Teilmenge von K. Wir müssen einen Häufungspunkt von E konstruieren der in K liegt. Wir führen die Konstruktion durch für den Fall $n = 1$. Im allgemeinen geht man genauso vor, nur das man anstelle von Intervallen entsprechend Quader verwendet.

Da E beschränkt ist, gibt es eine Zahl $M > 0$ sodaß $E \subseteq I_1$, wobei $I_1 := [-M, M]$. Eines der beiden Intervalle $[-M, 0]$ und $[0, M]$ muß unendlich viele Punkte von E enthalten. Bezeichne dieses mit I_2. Wir halbieren das Intervall I_2, dann muß eine der beiden Hälften unendlich viele Punkte enthalten. Sei diese I_3. Verfährt man induktiv so weiter, erhält man eine Folge $(I_n)_{n\in\mathbb{N}}$ von Intervallen mit $I_n \supseteq I_{n+1}$, $n \in \mathbb{N}$, wobei I_n die Länge $\frac{2M}{2^{n-1}}$ hat, und jedes I_n unendlich viele Punkte von E enthält.

Setze $x := \sup_{n\in\mathbb{N}}(\min I_n)$, und sei $r > 0$. Dann existiert $N \in \mathbb{N}$ sodass $\min I_N > x - r$ und $\frac{2M}{2^{N-1}} < r$. Für ein solches N gilt dann $I_N \subseteq U_r(x)$. Da I_N unendlich viele Punkte von E enthält, muss es auch einen von x verschiedenen Punkt E enthalten. Also haben wir $E \cap (U_r(x) \setminus \{x\}) \neq \emptyset$.

Die Menge E hat also den Häufungspunkt x. Da $E \subseteq K$, ist x auch ein Häufungspunkt von K. Da K abgeschlossen ist, liegt x in K.

Schritt 2: Wir zeigen: Ist $\mathcal{G} = \{G_i : i \in I\}$ eine Familie von offenen Mengen des $\mathbb{R}^n$, dann exisiert eine abzählbare Teilfamilie $\mathcal{G}' = \{G_{j(1)}, G_{j(2)}, \dots\}$ von $\mathcal{G}$ mit

$$\bigcup_{i\in I} G_i = \bigcup_{k\in\mathbb{N}} G_{j(k)} \,.$$

Wähle eine Folge $(V_n)_{n\in\mathbb{N}}$ wie in Lemma 1.2.7. Setze $O := \bigcup_{i\in I} G_i$. Für jedes $x \in O$ existieren Indizes $i \in I$ mit $x \in G_i$. Wähle einen solchen: $i(x)$. Nach Lemma 1.2.7 existieren Indizes $l \in \mathbb{N}$ sodaß $x \in V_l \subseteq G_{i(x)}$. Wähle einen solchen: $l(x)$.

Die Menge $M := \{l(x) : x \in O\}$ ist eine Teilmenge von $\mathbb{N}$ und wir können sie also anschreiben als $M = \{l_1, l_2, l_3, \dots\}$. Für $k \in \mathbb{N}$ definiere $j(k)$ wie folgt: Es ist $l_k \in M$, also gibt es $x \in O$ mit $l_k = l(x)$. Wähle ein solches: $x(k)$. Setze $j(k) := i(x(k))$. Beachte, dass stets gilt $V_{l_k} \subseteq G_{j(k)}$.

Wir zeigen $\bigcup_{k\in\mathbb{N}} G_{j(k)} = O$. Die Inklusion „$\subseteq$" ist trivial. Sei $y \in O$, dann ist $l(y) = l_k$ für ein gewisses $k \in \mathbb{N}$. Es folgt $y \in V_{l_k} \subseteq G_{j(k)}$, und wir erhalten $y \in \bigcup_{k\in\mathbb{N}} G_{j(k)}$.

Schritt 3, $(ii) \Rightarrow (i)$: Wegen Schritt 2 genügt es zu zeigen, dass jede abzählbare offene Überdeckung von K eine endliche Teilüberdeckung besitzt. Sei $(G_n)_{n\in\mathbb{N}}$ eine solche. Setze

$$H_n := G_1 \cup G_2 \cup \ldots \cup G_n,\ n \in \mathbb{N}\,.$$

Dann gilt $H_n \subseteq H_{n+1}$ und alle H_n sind offen. Weiters gilt $K \subseteq \bigcup_{n\in\mathbb{N}} H_n$. Sei $F_n := H_n^c \cap K$. Dann hat wegen Lemma 1.2.6, (i), F_n die Eigenschaft (HP). Weiters gilt stets $F_n \supseteq F_{n+1}$. Wegen $K \subseteq \bigcup_{n\in\mathbb{N}} H_n$ folgt $\bigcap_{n\in\mathbb{N}} F_n = \emptyset$, und es muß wegen Lemma 1.2.6, (iii), schon eine Menge F_N leer sein. Dann gilt also

$$K \subseteq H_N = G_1 \cup \ldots \cup G_N\,,$$

und wir haben eine endliche Teilüberdeckung gefunden.

Schritt 4, $(iv) \Rightarrow (ii)$: Sei E eine unendliche Teilmenge von K. Wähle eine Folge $(x_n)_{n\in\mathbb{N}}$ von paarweise verschiedenen Punkten aus E. Sei $(x_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge die gegen einen Punkt $x \in K$ konvergiert. Da die Punkte x_n paarweise verschieden sind, ist die Menge $\{x_{n_k} : k \in \mathbb{N}\}$ unendlich. Nach Proposition 1.1.8, (ii), ist x ein Häufungspunkt dieser Menge, und damit auch von E.

❑

Wir haben im Beweis von Satz 1.2.5, $(iii) \Rightarrow (ii)$, eigentlich die folgende Aussage gezeigt:

1.2.8 Korollar (Bolzano[3]-Weierstraß[4])**.** *Sei E eine beschränkte unendliche Teilmenge von $\mathbb{R}^n$. Dann hat E einen Häufungspunkt.* ❑

Schliesslich wollen wir eine oft praktische Tatsache explizit herausstellen:

1.2.9 Korollar. *Sei $K \subseteq \mathbb{R}$ kompakt. Dann besitzt K ein Supremum und ein Infimum und diese beiden Zahlen gehören zu K.*

Beweis. Die Menge K ist beschränkt und hat daher Supremum und Infimum. Für jedes $n \in \mathbb{N}$ wähle $x_n, y_n \in K$ mit $\sup K - x_n \leq \frac{1}{n}$ und $y_n - \inf K \leq \frac{1}{n}$. Dann ist $x_n \to \sup K$ und $y_n \to \inf K$. Da K abgeschlossen ist, folgt $\sup K, \inf K \in K$.

❑

1.3 Konvergenz gegen ∞

Betrachte den Raum $\mathbb{R}$ der reellen Zahlen, wie üblich versehen mit der euklidischen Metrik. Die Folge $(x_n)_{n\in\mathbb{N}}$ die definiert ist als

$$x_n := n,\ n \in \mathbb{N}\,,$$

ist nicht beschränkt und daher sicherlich nicht konvergent. Und doch zeigt sie ein irgendwie strebsames Verhalten; sie wächst und wächst und .... Man könnte also intuitiv sagen: Diese Folge „konvergiert gegen $+\infty$“.

Wir wollen dieser Idee nun Sinn geben, und sie exakt formulieren. Dazu betrachte die Menge $\overline{\mathbb{R}}$ die aus $\mathbb{R}$ und zwei zusätzlichen Elementen $+\infty$ und $-\infty$ besteht:

[3] Bernard Bolzano. 5.10.1781 Prag - 18.12.1848 Prag

[4] Karl Theodor Wilhelm Weierstraß. 31.10.1815 Ostenfelde (Westphalen) - 19.12.1897 Berlin

$$\begin{array}{ccccc} -\infty & & \mathbb{R} & & +\infty \\ \bullet & \cdots & \overline{\hspace{6cm}} & \cdots & \bullet \end{array}$$

Auf dieser Menge wollen wir eine Ordnungsrelation und eine Metrik definieren. Dazu zeigen wir zuerst eine Hilfsaussage.

1.3.1 Lemma. *Die Abbildung*

$$\sigma : \begin{cases} (-1,1) & \to & \mathbb{R} \\ \quad x & \mapsto & \frac{x}{1-|x|} \end{cases}$$

ist bijektiv und streng monoton wachsend.

Beweis. Es gilt $\sigma(x) = 0$ genau dann wenn $x = 0$, und stets ist $\sigma(-x) = -\sigma(x)$. Es genügt also zu zeigen, dass $\sigma|_{(0,1)}$ eine streng monoton wachsende Bijektion von $(0,1)$ auf $\mathbb{R}^+$ ist.

Seien $x, y \in (0,1)$, dann gilt

$$\sigma(x) - \sigma(y) = \frac{x}{1-x} - \frac{y}{1-y} = \frac{x(1-y) - y(1-x)}{(1-x)(1-y)} =$$
$$= \underbrace{\frac{1}{(1-x)(1-y)}}_{>0}(x-y)$$

Wir sehen, dass $\sigma(x) > \sigma(y)$ für $x > y$. Also ist σ streng monoton wachsend und daher auch injektiv.

Sei nun $z \in \mathbb{R}^+$ gegeben. Setze $x := \frac{z}{1+z}$, dann ist $x \in (0,1)$ und es gilt

$$\sigma(x) = \frac{\frac{z}{1+z}}{1 - \frac{z}{1+z}} = z$$

also ist $\sigma|_{(0,1)} : (0,1) \to \mathbb{R}^+$ surjektiv.

❑

Die Abbildung σ setzen wir nun in natürlicher Weise zu einer Bijektion von $[-1,1]$ auf $\overline{\mathbb{R}}$ fort, nämlich sei $\overline{\sigma} : [-1,1] \to \overline{\mathbb{R}}$ definiert als

$$\overline{\sigma}(x) := \begin{cases} \sigma(x) & , \ x \in (-1,1) \\ +\infty & , \ x = 1 \\ -\infty & , \ x = -1 \end{cases}$$

Bezeichne mit $\iota : \overline{\mathbb{R}} \to [-1,1]$ die Inverse von $\overline{\sigma}$. Aus dem Beweis von Lemma 1.3.1 sieht man leicht, dass ι explizit gegeben ist als

$$\iota(x) = \begin{cases} \frac{x}{1+|x|} & , \ x \in \mathbb{R} \\ 1 & , \ x = +\infty \\ -1 & , \ x = -\infty \end{cases}$$

1.3.2 Definition. Für $x, y \in \overline{\mathbb{R}}$ sei

$$x \leq y \ :\Longleftrightarrow \ \iota(x) \leq \iota(y), \qquad \overline{d}(x,y) := |\iota(x) - \overline{\iota}(y)|$$

1.3.3 Proposition. *Die oben definierte Relation „$\leq$“ auf $\overline{\mathbb{R}}$ ist eine Ordnungsrelation. Die oben definierte Abbildung $\overline{d} : \overline{\mathbb{R}} \times \overline{\mathbb{R}} \to \mathbb{R}$ ist eine Metrik. Die Abbildung $\iota : \overline{\mathbb{R}} \to [-1,1]$ ist bijektiv, ordnungserhaltend, und isometrisch. Es gilt*

$$x \leq y \ :\Longleftrightarrow \ x = -\infty \vee y = +\infty \vee (x, y \in \mathbb{R} \wedge x \leq y) \tag{1.1}$$

wobei das „$\leq$“ auf der rechten Seite die Ordnungsrelation auf $\mathbb{R}$ bezeichnet.

Beweis. Die Tatsache dass „$\leq$“ und $\overline{d}$ eine Ordnungsrelation bzw. eine Metrik sind, lässt sich aufgrund unserer Definition unmittelbar daraus herleiten, dass ι bijektiv ist und die Ordnungsrelation sowie die Betragsmetrik auf $[-1,1]$ eine Ordnungsrelation bzw. eine Metrik sind:
`ad`„Ordnungsrelation“: Zunächst ist $x \leq x$ genau dann wenn $\iota(x) \leq \iota(x)$ und das ist stets richtig. Sind $x, y \in \overline{\mathbb{R}}$ und gilt $x \leq y$ und $y \leq x$, so heißt das gerade $\iota(x) \leq \iota(y)$ und $\iota(y) \leq \iota(x)$. Also folgt $\iota(x) = \iota(y)$ und daher $x = y$. Schliesslich folgt aus $x \leq y$ und $y \leq z$ dass $\iota(x) \leq \iota(y)$ und $\iota(y) \leq \iota(z)$, damit $\iota(x) \leq \iota(z)$, und damit $x \leq z$.

`ad`„Metrik“: Es gilt

$$0 = \overline{d}(x, y) \iff 0 = |\iota(x) - \iota(y)| \iff \iota(x) = \iota(y) \iff x = y$$

Weiters ist

$$\overline{d}(y, x) = |\iota(y) - \iota(x)| = |\iota(x) - \iota(y)| = \overline{d}(x, y)\,,$$

und

$$\overline{d}(x, y) = |\iota(x) - \iota(y)| \leq |\iota(x) - \iota(z)| + |\iota(z) - \iota(y)| = \overline{d}(x, z) + \overline{d}(z, y)\,.$$

Die Tatsache dass die Bijektion ι ordnungserhaltend und isometrisch ist, ist offensichtlich, denn wir haben „$\leq$“ und $\overline{d}$ ja gerade so definiert.

Es bleibt zu zeigen, dass (1.1) gilt. Zuerst die Implikation „$\Leftarrow$“: Ist $x = -\infty$, so ist $\iota(x) = -1$, und daher $\iota(x) \leq \iota(y)$ für beliebiges $y \in \overline{\mathbb{R}}$. Ist $y = +\infty$, so ist $\iota(y) = 1$ und daher $\iota(x) \leq \iota(y)$ für beliebiges $x \in \overline{\mathbb{R}}$. Seien nun $x, y \in \mathbb{R}$ und sei $x \leq y$ wobei hier die Ordnungsrelation auf $\mathbb{R}$ gemeint ist. Da σ monoton ist, hat auch σ^{-1} diese Eigenschaft, und wir erhalten $\iota(x) = \sigma^{-1}(x) \leq \sigma^{-1}(y) = \iota(y)$, also $x \leq y$ wobei hier nun die Ordnungsrelation auf $\overline{\mathbb{R}}$ steht. Die umgekehrte Implikation „$\Rightarrow$“ sieht man genauso ein.

❏

1.3.4 Korollar.

(*i*) *Der metrische Raum $\langle \overline{\mathbb{R}}, \overline{d} \rangle$ ist kompakt (und daher vollständig) und zusammenhängend.*

(*ii*) *Jede nichtleere Teilmenge von $\overline{\mathbb{R}}$ hat ein Supremum und ein Infimum.*

(*iii*) *Jede monotone Folge von Punkten aus $\overline{\mathbb{R}}$ konvergiert (und zwar gegen $\sup\{x_n : n \in \mathbb{N}\}$ bzw. gegen $\inf\{x_n : n \in \mathbb{N}\}$ je nachdem ob sie monoton wachsend oder monoton fallend ist).*

Beweis.
`ad`(*i*): Folgt unmittelbar aus Proposition 1.1.16.

`ad`(ii): Das Intervall $[-1,1]$ hat diese Eigenschaft, denn: Eine nichtleere Teilmenge von $[-1,1]$ ist insbesondere eine nichtleere beschränkte Teilmenge von $\mathbb{R}$ und hat daher ein Supremum und ein Infimum in $\mathbb{R}$. Da 1 eine obere und -1 eine untere Schranke ist, liegt dieses Supremum und dieses Infimum in $[-1,1]$.

`ad`(iii): Ist $(x_n)_{n\in\mathbb{N}}$ eine monotone Folge in $\overline{\mathbb{R}}$, so ist $(\iota(x_n))_{n\in\mathbb{N}}$ eine monotone und beschränkte Folge in $\mathbb{R}$. Diese ist gegen eine reelle Zahl x konvergent, und zwar gegen $\sup\{\iota(x_n) : n \in \mathbb{N}\}$ bzw. gegen $\inf\{\iota(x_n) : n \in \mathbb{N}\}$, je nachdem ob sie monoton wachsend oder monoton fallend ist. Da $[-1,1]$ abgeschlossen ist, ist $x \in [-1,1]$. Also konvergiert $(x_n)_{n\in\mathbb{N}}$, und zwar gegen $\overline{\sigma}(x)$.

❑

1.3.5 Proposition. *Seien $x, y \in \mathbb{R} \subseteq \overline{\mathbb{R}}$. Dann gilt*

$$\overline{d}(x,y) \leq |x-y|\,.$$

Beweis. Seien zunächst $x, y \geq 0$. Dann gilt

$$\begin{aligned}|\iota(x)-\iota(y)| &= \frac{x}{1+x} - \frac{y}{1+y} = \frac{x(1+y)-y(1+x)}{(1+x)(1+y)} = \\ &= \underbrace{\frac{1}{(1+x)(1+y)}}_{\leq 1}|x-y| \leq |x-y|\end{aligned}$$

Seien nun $x, y \leq 0$, dann haben wir

$$|\iota(x)-\iota(y)| = |-\iota(-x)+\iota(-y)| = |\iota(y)-\iota(x)| \leq |y-x| = |x-y|\,.$$

Sei schliesslich $x > 0 > y$, dann ist

$$\begin{aligned}|\iota(x)-\iota(y)| = \iota(x)-\iota(y) = |\iota(x)-0|+|0-\iota(y)| = |\iota(x)-\iota(0)|+|\iota(0)-\iota(y)| \leq \\ \leq |x-0|+|0-y| = x-y = |x-y|\,.\end{aligned}$$

Der Fall $x < 0 < y$ folgt wegen der Symmetrie der Metrik.

❑

1.3.6 Korollar. *Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen, die in $\mathbb{R}$ gegen die Zahl x konvergiert. Betrachte x_n als Elemente von $\overline{\mathbb{R}}$. Dann ist $(x_n)_{n\in\mathbb{N}}$ in $\overline{\mathbb{R}}$ ebenfalls gegen x konvergent.*

Beweis. Sei $\epsilon > 0$ gegeben, dann wähle N mit $|x_n - x| < \epsilon$, $n \geq N$. Es folgt dass auch $\overline{d}(x_n, x) < \epsilon$, $n \geq N$.

❑

1.3.7 Beispiel. Die Umkehrung dieses Korollars ist nicht richtig. Betrachte zum Beispiel die Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n := n$. Diese ist in $\mathbb{R}$ nicht konvergent. Es gilt jedoch

$$\overline{d}(x_n,+\infty) = |\iota(x_n)-\iota(+\infty)| = \left|\frac{n}{1+n}-1\right| = \frac{1}{1+n} \to 0\,,$$

also haben wir $x_n \to +\infty$ in $\overline{\mathbb{R}}$.

1.3.8 Bemerkung.

(*i*) Wir wollen anmerken, dass sich die Tatsache $\lim_{n\to\infty} x_n = +\infty$ auch wie folgt formulieren läßt:

$$\forall C \in \mathbb{R}\, \exists N \in \mathbb{N} : x_n > C \text{ für } n \geq N\,.$$

Analog ist $\lim_{n\to\infty} x_n = -\infty$ genau dann, wenn $\forall C \in \mathbb{R}\, \exists N \in \mathbb{N} : x_n < C$ für $n \geq N$.

(*ii*) Achtung: Es sind keine vernünftigen algebraischen Operationen „$+, \cdot$“ auf $\overline{\mathbb{R}}$ definiert! Man bezeichnet häufig, um die Notation zu vereinfachen, für Elemente $x, y \in \overline{\mathbb{R}}$

$$x + y := \begin{cases} x + y & , \ x, y \in \mathbb{R} \\ +\infty & , \ (x \in \mathbb{R}, y = +\infty) \vee (x = +\infty, y \in \mathbb{R}) \\ -\infty & , \ (x \in \mathbb{R}, y = -\infty) \vee (x = -\infty, y \in \mathbb{R}) \end{cases}$$

Die Summe $(+\infty) + (-\infty)$ bleibt undefiniert!

Schliesslich wollen wir noch einige Rechenregeln für Konvergenz gegen ∞ zusammenstellen.

1.3.9 Korollar. *Seien $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$, Folgen reeller Zahlen. Dann gilt:*

(*i*) *Ist $\lim_{n\to\infty} x_n = +\infty$ und $y_n \geq C$ für eine gewisse reelle Zahl C, so folgt $\lim_{n\to\infty}(x_n + y_n) = +\infty$.*

(*ii*) *Es ist $\lim_{n\to\infty} x_n = +\infty$ genau dann wenn $\lim_{n\to\infty}(-x_n) = -\infty$.*

(*iii*) *Ist $\lim_{n\to\infty} x_n = +\infty$ und $y_n \geq C$ für eine gewisse positive reelle Zahl C, so folgt $\lim_{n\to\infty}(x_n y_n) = +\infty$.*

(*iv*) *Ist $x_n > 0$, $n \in \mathbb{N}$, so gilt $\lim_{n\to\infty} x_n = +\infty$ genau dann wenn $\lim_{n\to\infty} \frac{1}{x_n} = 0$.*

Beweis. Diese Aussagen erhält man unmittelbar aus Bemerkung 1.3.8, (*i*).

❑

Nun können wir den in der Analysis oft gebräuchlichen Begriff des Limes superior bzw. Limes inferior einer Folge diskutieren.

1.3.10 Definition. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge, $x_n \in \overline{\mathbb{R}}$. Dann definieren wir

$$\limsup_{n\to\infty} x_n := \inf_{N\in\mathbb{N}} \sup_{n\geq N} x_n, \ \liminf_{n\to\infty} x_n := \sup_{N\in\mathbb{N}} \inf_{n\geq N} x_n\,.$$

Wir nennen die Zahl $\limsup_{n\to\infty} x_n$ den *Limes superior* der Folge $(x_n)_{n\in\mathbb{N}}$ und $\liminf_{n\to\infty} x_n$ ihren *Limes inferior*.

Man kann die Zahlen $\limsup_{n\to\infty} x_n$ und $\liminf_{n\to\infty} x_n$ in verschiedener Weise anders charakterisieren.

1.3.11 Proposition. *Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $\overline{\mathbb{R}}$.*

(i) *Es gilt* $\limsup_{n\to\infty} x_n = \lim_{N\to\infty}(\sup_{n\geq N} x_n)$, *und* $\liminf_{n\to\infty} x_n = \lim_{N\to\infty}(\inf_{n\geq N} x_n)$.

(ii) *Es ist* $\limsup_{n\to\infty} x_n$ *die (eindeutig bestimmte) kleinste Zahl* $M \in \overline{\mathbb{R}}$ *mit der Eigenschaft*

$$\forall \epsilon \in \mathbb{R}^+ \, \exists N \in \mathbb{N} : x_n \leq M + \epsilon \text{ für alle } n \geq N \,. \tag{1.2}$$

Analog ist $\liminf_{n\to\infty} x_n$ *die (eindeutig bestimmte) größte Zahl* $m \in \overline{\mathbb{R}}$ *mit der Eigenschaft*

$$\forall \epsilon \in \mathbb{R}^+ \, \exists N \in \mathbb{N} : x_n \geq m - \epsilon \text{ für alle } n \geq N \,.$$

(iii) *Sei* A *die (wegen Korollar 1.3.4, (i), nichtleere) Menge*

$$A := \left\{ y \in \overline{\mathbb{R}} : \exists (y_n)_{n\in\mathbb{N}} \text{ Teilfolge von } (x_n)_{n\in\mathbb{N}} \text{ mit } y_n \to y \right\}.$$

Dann gilt

$$\liminf_{n\to\infty} x_n = \min A, \; \limsup_{n\to\infty} x_n = \max A \,.$$

Beweis. Wir zeigen die Behauptungen für den Limes superior. Den Fall des Limes inferior behandelt man analog.
`ad`(i): Die Folge $z_N := \sup_{n\geq N} x_n$ ist monoton fallend, denn für $N \geq N'$ ist $\{x_n : n \geq N\} \supseteq \{x_n : n \geq N'\}$. Also ist sie konvergent, und zwar gegen $\inf_{N\in\mathbb{N}} z_N$.

`ad`(ii): Bemerke, dass für $N \in \mathbb{N}$ gilt:

$$x_n \leq K, \; n \geq N \iff z_N = \sup_{n\geq N} x_n \leq K \,.$$

Sei M eine Zahl mit der Eigenschaft (1.2). Sei $\epsilon > 0$ gegeben, dann existiert also ein $N \in \mathbb{N}$ mit $z_N \leq M + \epsilon$. Daher ist auch $\lim_{N\to\infty} z_N \leq M + \epsilon$. Da $\epsilon > 0$ beliebig war, folgt $\limsup_{n\to\infty} x_n = \lim_{N\to\infty} z_N \leq M$.

Setze $M := \limsup_{n\to\infty} x_n$, und sei $\epsilon > 0$ gegeben. Wegen $M = \lim_{N\to\infty} z_N$, gibt es ein $N \in \mathbb{N}$, sodaß $z_N \leq M + \epsilon$. Also hat die Zahl $\limsup_{n\to\infty} x_n$ die Eigenschaft (1.2).

`ad`(iii): Sei $(x_{n_k})_{k\in\mathbb{N}}$ eine konvergente Teilfolge von $(x_n)_{n\in\mathbb{N}}$. Sei $\epsilon \in \mathbb{R}^+$ und wähle $N \in \mathbb{N}$ sodass $x_n \leq \limsup_{n\to\infty} x_n + \epsilon$ für alle $n \geq N$. Dann gilt insbesondere auch $x_{n_k} \leq \limsup_{n\to\infty} x_n + \epsilon$, $k \geq N$, und damit folgt $\lim_{k\to\infty} x_{n_k} \leq \limsup_{n\to\infty} x_n + \epsilon$. Da ϵ beliebig war, erhalten wir $\lim_{k\to\infty} x_{n_k} \leq \limsup_{n\to\infty} x_n$. Wir schliessen dass $\sup A \leq \limsup_{n\to\infty} x_n$.

Wir zeigen nun, dass $\limsup_{n\to\infty} x_n \in A$. Betrachte die Folge $a_n := \iota(x_n)$ und setze $b_N := \sup_{n\geq N} a_n$. Dann ist $\lim_{N\to\infty} b_N = \iota(\limsup_{n\to\infty} x_n) =: b$. Wir konstruieren induktiv eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$. Wähle N_1 sodass $|b_{N_1} - b| < 1$. Wähle $n_1 \geq N_1$ sodass $a_{n_1} \in [b_{N_1} - 1, b_{N_1}]$. Seien $n_1, \ldots, n_{m-1}$ bereits definiert. Wähle $N_m > n_{m-1}$ sodass $|b_{N_m} - b| < \frac{1}{m}$, und wähle $n_m \geq N_m$ sodass $a_{n_m} \in [b_{N_m} - \frac{1}{m}, b_{N_m}]$. Damit haben wir eine Teilfolge $(a_{n_m})_{m\in\mathbb{N}}$ definiert und für diese gilt $|a_{n_m} - b| < \frac{2}{m}$, also $a_{n_m} \to b$. Dann ist $(x_{n_m})_{m\in\mathbb{N}}$ eine Teilfolge von $(x_n)_{n\in\mathbb{N}}$,und es gilt $x_{n_m} \to \limsup_{n\to\infty} x_n$.

❑

1.3.12 Korollar. *Eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\overline{\mathbb{R}}$ ist genau dann konvergent wenn ihr Limes superior und Limes inferior übereinstimmen. Ihr Grenzwert ist in diesem Fall gleich diesem gemeinsamen Wert.*

Beweis. Sei $x_n \to x$, dann gilt auch für jede Teilfolge $(y_n)_{n\in\mathbb{N}}$ dass $y_n \to x$. Also ist $A = \{x\}$, und daher $\limsup_{n\to\infty} x_n = x = \liminf_{n\to\infty} x_n$.

Sei umgekehrt angenommen dass $\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n =: x$ gilt, d.h. dass $A = \{x\}$ ist. Sei $(y_n)_{n\in\mathbb{N}}$ eine Teilfolge von $(x_n)_{n\in\mathbb{N}}$. Da $\overline{\mathbb{R}}$ kompakt ist, existiert eine Teilfolge $(y_{n_k})_{k\in\mathbb{N}}$ die konvergiert. Wegen unserer Voraussetzung ist ihr Grenzwert gleich x. Also enthält jede Teilfolge eine Teilfolge die gegen x konvergiert, und es folgt dass $\lim_{n\to\infty} x_n = x$.

❑

Kapitel 2

Reelle und komplexe Funktionen

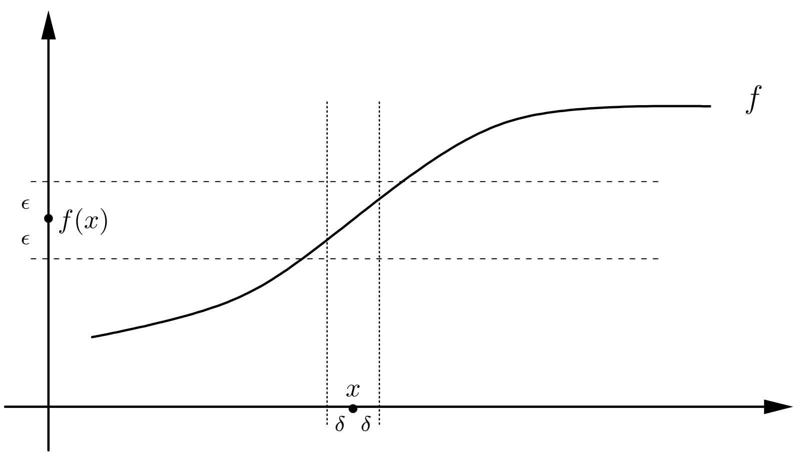
2.1 Stetigkeit

Sei f eine Funktion und sei x ein Punkt ihres Definitionsbereiches. Sagen wir dass diese Funktion stetig an der Stelle x ist, so verstehen wir darunter anschaulich, dass der Funktionswert $f(t)$ sich beliebig wenig von $f(x)$ unterscheidet, wenn nur t hinreichend nahe bei x ist. Wir sehen, dass man diesem Begriff Sinn geben kann, wenn man verlangt dass Definitionsbereich und Wertebereich der betrachteten Funktion metrische Räume sind.

2.1.1 Definition. Seien $\langle X, d_X\rangle$ und $\langle Y, d_Y\rangle$ metrische Räume, und sei $f : X \to Y$ eine Funktion. Weiters sei $x \in X$. Dann heißt f *stetig an der Stelle* x, wenn gilt

$$\forall \epsilon > 0\, \exists \delta > 0\, \forall t \in X : \quad d_X(t,x) < \delta \;\Rightarrow\; d_Y(f(t), f(x)) < \epsilon\,.$$

Ist f an jeder Stelle $x \in X$ stetig, so heißt f *stetig*.



2.1.2 Beispiel.

(i) Sei $f : X \to Y$ eine konstante Funktion, $f(x) := y_0$, $x \in X$. Dann ist f stetig. Denn: Sei $x \in X$ und $\epsilon > 0$ gegeben. Wähle $\delta := 1$, dann gilt für alle $t \in X$ mit $d_X(t,x) < \delta$ sicher

$$d_Y(f(t), f(x)) = d_Y(y_0, y_0) = 0 < \epsilon\,.$$

(ii) Sei $f : X \to X$ die identische Abbildung, $f(x) := x$, $x \in X$. Dann ist f stetig. Denn: Sei $x \in X$ und $\epsilon > 0$ gegeben. Wähle $\delta := \epsilon$, dann gilt für alle $t \in X$ mit $d_X(t,x) < \delta$

$$d_X(f(t), f(x)) = d_X(t,x) < \delta = \epsilon\,.$$

Allgemeiner gilt: Sei $f : X \to Y$ eine isometrische Abbildung. Dann ist f stetig. Denn: Sei $x \in X$ und $\epsilon > 0$ gegeben. Wähle $\delta := \epsilon$, dann gilt für alle $t \in X$ mit $d_X(t,x) < \delta$

$$d_Y(f(t), f(x)) = d_X(t,x) < \delta = \epsilon\,.$$

(iii) Sei $\langle X, d_X\rangle$ ein metrischer Raum, und sei $X \times X$ versehen mit der Metrik $d_+ : (X \times X)^2 \to \mathbb{R}$ die definiert ist als

$$d_+\big((x_1,y_1),(x_2,y_2)\big) := d_X(x_1,x_2) + d_X(y_1,y_2)\,.$$

Dann ist die Abbildung $d_X : X \times X \to \mathbb{R}$ stetig.

Um dies einzusehen, seien $(x_1,y_1),(x_2,y_2) \in X \times X$ gegeben. Dann gilt

$$d_X(x_1,y_1) \leq d_X(x_1,y_2) + d(y_2,y_1),\ d_X(x_1,y_2) \leq d_X(x_1,y_1) + d(y_1,y_2)\,,$$

also

$$\big|d_X(x_1,y_1) - d_X(x_1,y_2)\big| \leq d(y_1,y_2)\,.$$

Genauso sieht man $|d_X(x_1,y_2) - d_X(x_2,y_2)\big| \leq d(x_1,x_2)$, und es folgt

$$\big|d_X(x_1,y_1) - d_X(x_2,y_2)\big| \leq \big|d_X(x_1,y_1) - d_X(x_1,y_2)\big| +$$

$$+\big|d_X(x_1,y_2) - d_X(x_2,y_2)\big| \leq d(y_1,y_2) + d(x_1,x_2) = d_+\big((x_1,y_1),(x_2,y_2)\big)\,.$$

(iv) Sei $f : \mathbb{C} \to \mathbb{R}$ die Funktion $z \mapsto |z|$. Dann ist f stetig. Denn: Sei $z \in \mathbb{C}$ und $\epsilon > 0$ gegeben. Wähle $\delta := \epsilon$, dann gilt für alle $w \in \mathbb{C}$ mit $|w - z| < \delta$ wegen der Dreiecksungleichung nach unten

$$|f(w) - f(z)| = \big||w| - |z|\big| \leq |w - z| < \delta = \epsilon\,.$$

(v) Betrachte die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, die definiert ist als $f(x,y) := \max\{x,y\}$. Hierbei ist $\mathbb{R}^2$ (wie immer) versehen mit der euklidischen Metrik. Dann ist f stetig. Denn: Sei $(x,y) \in \mathbb{R}^2$, $\epsilon > 0$ gegeben, und wähle $\delta := \epsilon$. Ist $(a,b) \in \mathbb{R}^2$ mit $d((a,b),(x,y)) < \delta$, so folgt

$$|a - x|, |b - y| \leq d((a,b),(x,y)) < \delta\,,$$

und wir erhalten

$$a < x + \delta,\ b < y + \delta\,,$$

$$x < a + \delta,\ y < b + \delta\,.$$

Damit folgt

$$\max\{a, b\} < \max\{x + \delta, y + \delta\} = \max\{x, y\} + \delta\,,$$

$$\max\{x, y\} < \max\{a + \delta, b + \delta\} = \max\{a, b\} + \delta\,,$$

und wir erhalten

$$-\delta < \max\{a, b\} - \max\{x, y\} < \delta\,,$$

also $|\max\{a, b\} - \max\{x, y\}| < \delta = \epsilon$.

(vi) Sei $f : \mathbb{R} \to \mathbb{R}$ die Funktion $f(x) := [x]$. Dabei bezeichnet $[x]$ die größte ganze Zahl die $\leq x$ ist. Man bezeichnet diese Funktion auch als die *Gauß-klammer*. Diese Funktion ist stetig an jeder Stelle $x \in \mathbb{R} \setminus \mathbb{Z}$, und nicht stetig an jeder Stelle $x \in \mathbb{Z}$.

Denn: Ist $x \in \mathbb{R} \setminus \mathbb{Z}$, und ist $\epsilon > 0$ gegeben, so wähle $\delta > 0$ sodaß das Intervall $(x-\delta, x+\delta)$ keine ganze Zahl enthält. Dann ist f auf $(x-\delta, x+\delta)$ konstant, also gilt

$$|f(t) - f(x)| = 0 < \epsilon,\ |t - x| < \delta\,.$$

Ist dagegen $x \in \mathbb{Z}$, so enthält das Intervall $(x - \delta, x + \delta)$ für jedes $\delta > 0$ sowohl Zahlen t die größer als x sind, als auch Zahlen t die kleiner als x sind. Nun ist aber für $t_- < x$ und $t_+ > x$ sicher $f(t_-) < f(t_+)$ und daher, da ja beide Werte ganze Zahlen sind, $|f(t_-) - f(t_+)| \geq 1$. Wir können also für kein ϵ mit $0 < \epsilon < 1$ ein δ finden, dass der geforderten Bedingung genügt.

Nun wollen wir die Stetigkeit an einer Stelle anders charakterisieren, nämlich mit Hilfe von Grenzwerten. Dazu müssen wir zuerst den Begriff des Grenzwertes einer Funktion erklären.

2.1.3 Definition. Seien $\langle X, d_X\rangle$ und $\langle Y, d_Y\rangle$ metrische Räume, $D \subseteq X$, und sei $f : D \to Y$. Weiters sei $x_0 \in X$ ein Häufungspunkt von D. Dann sagen wir der *Grenzwert* $\lim_{x\to x_0} f(x)$ existiert und ist gleich y_0, wenn gilt

$$\forall\epsilon > 0\,\exists\delta > 0\,\forall x \in D : \quad 0 < d_X(x, x_0) < \delta\ \Rightarrow\ d_Y(f(x), y_0) < \epsilon$$

Dieser Limesbegriff steht in Zusammenhang mit dem bereits bekannten Begriff des Limes von Folgen.

2.1.4 Proposition. *Seien $\langle X, d_X\rangle$ und $\langle Y, d_Y\rangle$ metrische Räume, $D \subseteq X$, $f : D \to Y$, $x_0 \in X$ ein Häufungspunkt von D, und $y_0 \in Y$. Dann ist $\lim_{x\to x_0} f(x) = y_0$ genau dann, wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$ von Punkten aus $D \setminus \{x_0\}$ mit $\lim_{n\to\infty} x_n = x_0$ gilt dass $\lim_{n\to\infty} f(x_n) = y_0$ ist.*

Beweis. Sei vorausgesetzt dass $\lim_{x\to x_0} f(x) = y_0$, und sei $x_n \to x_0$. Ist $\epsilon > 0$ gegeben, so wähle $\delta > 0$ gemäß Definition 2.1.3, und $N \in \mathbb{N}$ mit $d_X(x_n, x_0) < \delta$, $n \geq N$. Dann gilt

$$d_Y(f(x_n), y_0) < \epsilon,\ n \geq N\,,$$

also ist $\lim_{n\to\infty} f(x_n) = y_0$.

Umgekehrt sei angenommen, dass die Aussage $\lim_{x\to x_0} f(x) = y_0$ falsch ist. Dann existiert also ein $\epsilon > 0$ sodass für jedes $\delta > 0$ ein Element $x(\delta)$ existiert mit $d_X(x(\delta), x_0) < \delta$ aber $d_Y(f(x(\delta)), y_0) \geq \epsilon$. Dann ist $x_n := x(\frac{1}{n})$ eine Folge von Punkten aus X mit $x_n \to x_0$, aber $f(x_n) \not\to y_0$.

❑

2.1.5 Korollar. *Seien $\langle X, d_X\rangle$ und $\langle Y, d_Y\rangle$ metrische Räume, $f : X \to Y$, und $x_0 \in X$. Ist x_0 ein Häufungspunkt von X, so sind äquivalent:*

(*i*) *f ist stetig an der Stelle x_0.*

(*ii*) *Es gilt $\lim_{x\to x_0} f(x) = f(x_0)$.*

(*iii*) *Für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} x_n = x_0$ gilt $\lim_{n\to\infty} f(x_n) = f(x_0)$.*

Ist x_0 kein Häufungspunkt von X, so ist f an der Stelle x_0 stetig.

Beweis. Die Äquivalenz von (*i*) und (*ii*) ist unmittelbar aus der Definition ersichtlich. Die Äquivalenz von (*ii*) und (*iii*) ist unmittelbar aus Proposition 2.1.4.

Ist x_0 kein Häufungspunkt von X, so existiert $\delta > 0$ mit $U_\delta(x_0) = \{x_0\}$. Diese Wahl von δ funktionert in der Definition der Stetigkeit daher für jedes $\epsilon > 0$.

❑

In anderen Worten ausgedrückt ist also f stetig, wenn man stets Limesbildung und Anwendung von f vertauschen darf, $f(\lim x_n) = \lim f(x_n)$.

2.1.6 Beispiel. Aus unseren Rechenregeln für Folgen erhalten wir nun unmittelbar: Die algebraischen Operationen auf $\mathbb{R}$

$$+ : \begin{cases} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x+y \end{cases} \qquad \cdot : \begin{cases} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x\cdot y \end{cases}$$

$$- : \begin{cases} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & -x \end{cases} \qquad .^{-1} : \begin{cases} \mathbb{R}\setminus\{0\} & \to & \mathbb{R}\setminus\{0\} \\ x & \mapsto & \frac{1}{x} \end{cases}$$

sind stetig. Aus genau dem gleichen Grund, sind die algebraischen Operationen auf $\mathbb{C}$ stetig.

Die folgende, mengentheoretisch orientierte, Charakterisierung der Stetigkeit einer Funktion spielt eine wichtige Rolle.

2.1.7 Satz. *Seien $\langle X, d_X\rangle$ und $\langle Y, d_Y\rangle$ metrische Räume, und sei $f : X \to Y$. Dann ist f stetig genau dann, wenn für jede in $\langle Y, d_Y\rangle$ offene Menge B das vollständige Urbild $f^-(B)$ offen in $\langle X, d_X\rangle$.*

Beweis. Sei vorausgesetzt dass f stetig ist, und sei $O \subseteq Y$ offen. Weiters sei $x \in f^-(O)$ gegeben. Wähle $\epsilon > 0$, sodass die Kugel $U^Y_\epsilon(f(x))$ ganz in O enthalten ist, und wähle $\delta > 0$ gemäß Definition 2.1.1. Ist $t \in U^X_\delta(x)$, so ist also $d_X(t,x) < \delta$. Daher folgt $d_Y(f(t), f(x)) < \epsilon$, d.h. $f(t) \in U^Y_\epsilon(f(x)) \subseteq O$. Wir sehen, dass $U^X_\delta(x) \subseteq f^-(O)$.

Umgekehrt sei vorausgesetzt dass das Urbild offener Mengen stets offen ist, und sei $x \in X$, $\epsilon > 0$, gegeben. Dann ist die Kugel $U_\epsilon^Y(f(x))$ eine offene Teilmenge von Y, und daher $f^-(U_\epsilon^Y(f(x)))$ offen in X. Nun gilt $x \in f^-(U_\epsilon^Y(f(x)))$, also existiert eine ganze Kugel $U_\delta^X(x) \subseteq f^-(U_\epsilon^Y(f(x)))$. Dieses δ ist genau jenes das wir für die Bedingung in Definition 2.1.1 brauchen, denn wir haben nun

$$d_X(t,x) < \delta \Rightarrow t \in U_\delta^X(x) \Rightarrow t \in f^-\big(U_\epsilon^Y(f(x))\big) \Rightarrow f(t) \in U_\epsilon^Y(f(x)) \Rightarrow$$

$$\Rightarrow d_Y(f(t), f(x)) < \epsilon$$

❑

Aus dieser Charakterisierung erhalten wir ganz einfach einige wesentliche Eigenschaften stetiger Funktionen.

2.1.8 Satz. *Seien $\langle X, d_X\rangle$ und $\langle Y, d_Y\rangle$ metrische Räume, und sei $f : X \to Y$ eine stetige Funktion.*

- (i) *Ist $K \subseteq X$ kompakt, so ist auch $f(K) \subseteq Y$ kompakt.*
- (ii) *Ist $E \subseteq X$ zusammenhängend, so ist auch $f(E) \subseteq Y$ zusammenhängend.*
- (iii) *Ist zusätzlich $\langle Z, d_Z\rangle$ ein metrischer Raum und $g : Y \to Z$ eine stetige Funktion, dann ist auch die Zusammensetzung $g \circ f : X \to Z$ stetig.*

Beweis.
`ad`(i): Sei $\mathcal{G} = \{G_i : i \in I\}$ eine Überdeckung von $f(K)$ durch in Y offene Mengen. Dann ist $\mathcal{G}^- := \{f^-(G_i) : i \in I\}$ eine Überdeckung von K durch in X offene Mengen. Da K kompakt ist, existieren $i_1, \dots, i_n \in I$ sodaß $\{f^-(G_{i_1}), \dots, f^-(G_{i_n})\}$ eine Überdeckung von K. Dann ist $\{G_{i_1}, \dots, G_{i_n}\}$ auch eine Überdeckung von $f(K)$.

`ad`(ii): Seien A, B in Y offene und disjunkte Mengen mit $f(E) \subseteq A \cup B$. Dann sind $f^-(A), f^-(B)$ offen in X und disjunkt, und es gilt $E \subseteq f^-(A) \cup f^-(B)$. Da E zusammenhängend ist, gilt entweder $E \subseteq f^-(A)$ oder $E \subseteq f^-(B)$. Es folgt das entweder $f(E) \subseteq A$ oder $f(E) \subseteq B$.

`ad`(iii): Sei A offen in Z. Dann ist $g^-(A)$ offen in Y und daher $f^-(g^-(A))$ offen in X. Nun gilt $(g \circ f)^-(A) = f^-(g^-(A))$.

❑

2.1.9 Korollar. *Sei $f : X \to \mathbb{R}$ stetig und sei $K \subseteq X$ kompakt. Dann ist f auf K beschränkt und nimmt ein Maximum und ein Minimum an, d.h. es gibt Punkte $x_+, x_- \in K$ mit*

$$f(x_+) = \max_{x \in K} f(x), \quad f(x_-) = \min_{x \in K} f(x)\,.$$

Insbesondere nimmt jede auf einem reellen Intervall $[a, b]$ definierte und stetige reellwertige Funktion ein Maximum und ein Minimum an.

Beweis. Es ist $f(K) \subseteq \mathbb{R}$ kompakt und besitzt und enthält daher Supremum und Infimum, d.h. es gibt Punkte x_+, x_- mit $f(x_+) = \sup_{x \in K} f(x)$, $f(x_-) = \inf_{x \in K} f(x)$.

❑

2.1.10 Korollar. *Sei $f : X \to \mathbb{R}$ stetig und sei $E \subseteq X$ zusammenhängend. Ist $c \in \mathbb{R}$ ein Wert mit $\inf_{x \in E} f(x) < c < \sup_{x \in E} f(x)$, so existiert ein Punkt $x \in E$ mit $f(x) = c$.*

Insbesondere gilt der Zwischenwertsatz*: Sei f stetig auf $[a, b]$ und $f(a) < f(b)$. Für jede Zahl c mit $f(a) < c < f(b)$ existiert ein Punkt $x \in (a, b)$ mit $f(x) = c$. Genauso für $f(a) > f(b)$.*

Beweis. Wähle $a, b \in f(E)$ mit $\inf_{x \in E} f(x) < a < c < b < \sup_{x \in E} f(x)$. Da $f(E) \subseteq \mathbb{R}$ zusammenhängend ist, enthält $f(E)$ das ganze Intervall $[a, b]$. Daher gibt es ein $x \in E$ mit $f(x) = c$.

❑

Als weiteres Korollar erhalten wir die Stetigkeit rationaler Funktionen.

2.1.11 Beispiel. Sei f eine rationale Funktion mit reellen oder komplexen Koeffizienten, d.h. sei f von der Gestalt

$$f(x) = \frac{p(x)}{q(x)}$$

mit gewissen Polynomen p und q in $\mathbb{R}[x]$ oder $\mathbb{C}[x]$. Dann ist

$$f : \{x \in \mathbb{R} \text{ bzw. } \mathbb{C} : \ q(x) \neq 0\} \to \mathbb{R} \text{ bzw. } \mathbb{C}$$

stetig.

Die Abbildung $x \mapsto x$ sowie die Abbildung $\cdot : (a, b) \mapsto a \cdot b$ ist stetig. Nun können wir die Abbildung $x \mapsto x^2$ auch anschreiben als Zusammensetzung $x^2 = \cdot(x, x)$. Sie ist also die Zusammensetzung von stetigen Funktionen, und daher selbst stetig. Verfährt man induktiv so weiter, folgt dass alle Potenzen $x \mapsto x^n$, $n \in \mathbb{N}$, stetig sind. Da konstante Funktionen sowie auch die Addition $(a, b) \mapsto a + b$ stetig sind, folgt, wieder mit dem gleichen Argument, dass jedes Polynom stetig ist. Nun bleibt noch übrig zu verwenden, dass das Invertieren $x \mapsto \frac{1}{x}$ (auf der Menge aller von Null verschiedenen Zahlen) stetig ist, um zu sehen dass alle rationalen Funktionen stetig sind. ❑

Anmerkung: Wem diese Tatsache trivial vorkommt, der versuche „zu Fuß“, d.h. durch explizite Angabe einer Zahl δ zu vorgegebenen ϵ und x, zu überprüfen das die Funktion

$$f(x) = \frac{x^3 + x + 1}{x^2 + 1}$$

auf ihrem Definitionsbereich $\mathbb{R}$ stetig ist.

Wir wollen noch einmal zur ursprünglichen Definition der Stetigkeit einer Funktion zurückkehren. Diese lautet ja, in logischen Formeln angeschrieben,

$$\forall x \in X \, \forall \epsilon > 0 \, \exists \delta > 0 \, \forall t \in X : d_X(t, x) < \delta \Rightarrow d_Y(f(t), f(x)) < \epsilon \,. \qquad (2.1)$$

Die Zahl δ, die die existieren muss, wird im allgemeinen natürlich sowohl von ϵ als auch von x abhängen.

2.1.12 Beispiel. Betrachte die Funktion $f(x) = x^{-1} : \mathbb{R}^+ \to \mathbb{R}^+$. Ist $x \in \mathbb{R}$ und $\epsilon > 0$ gegeben, so berechnet man:

$$\frac{1}{x} - \frac{1}{x+\delta} = \frac{\delta}{(x+\delta)x}\,.$$

Damit dieser Ausdruck $\leq \epsilon$ ist, darf δ höchstens $\frac{\epsilon x^2}{1-\epsilon x}$ sein. Man sieht, dass diese größtmögliche Wahl von δ immer kleiner wird je kleiner x wird, und tatsächlich für $x \to 0$ ebenfalls gegen 0 strebt. Man kann in diesem Beispiel also tatsächlich zu gegebenem ϵ kein δ finden das nicht mehr von x abhängt.

Sollte eine Funktion nun so beschaffen sein, dass dieses Phänomen nicht auftritt, dass also zu jedem gegebenem ϵ stets ein δ existiert welches für alle x funktioniert, so nennt man die Funktion gleichmäßig stetig:

2.1.13 Definition. Seien $\langle X, d_X\rangle$ und $\langle Y, d_Y\rangle$ metrische Räume, und sei $f : X \to Y$ eine Funktion. Dann heißt f *gleichmäßig stetig*, wenn gilt

$$\forall \epsilon > 0\, \exists \delta > 0\, \forall x \in X\, \forall t \in X : d_X(t,x) < \delta \Rightarrow d_Y(f(t), f(x)) < \epsilon\,.$$

Vergleicht man diese Definition mit der Formel (2.1), so sieht man, dass man hier den Allquantor $\forall x \in X$ und den Existenzquantor $\exists \delta > 0$ vertauscht hat. Dies wird also nicht den gleichen, sondern einen stärkeren Begriff liefern: Ist f gleichmäßig stetig, so ist f auch stetig.

Wie wir in Beispiel 2.1.12 gesehen haben, gilt die Umkehrung tatsächlich nicht. Interessant in diesem Zusammenhang ist nun der folgende Satz.

2.1.14 Satz. *Sei $f : X \to Y$ stetig. Ist X kompakt, so ist f sogar gleichmäßig stetig.*

Beweis. Sei $\epsilon > 0$ gegeben. Zu jedem $x \in X$ existiert $\delta(x) > 0$, sodaß für die Kugel $U^X_{\delta(x)}(x)$ gilt

$$t \in U^X_{\delta(x)}(x) \Rightarrow d_Y(f(t), f(x)) < \epsilon\,.$$

Setze $V(x) := U^X_{\frac{\delta(x)}{2}}(x)$. Diese offenen Kugeln überdecken X. Da X kompakt ist, gibt es eine endliche Teilüberdeckung

$$X = V(x_1) \cup V(x_2) \cup \ldots \cup V(x_n)\,.$$

Setze $\delta := \frac{1}{2}\min\{\delta(x_1), \ldots, \delta(x_n)\}$, dann ist $\delta > 0$.

Seien nun $t, x \in X$ mit $d_X(t,x) < \delta$ gegeben. Wähle m mit $t \in V(x_m)$. Dann folgt dass $x \in U_{\delta(x_m)}(x_m)$. Wir erhalten nach der Dreiecksungleichung

$$d_Y(f(t), f(x)) \leq d_Y(f(t), f(x_m)) + d_Y(f(x_m), f(x)) < 2\epsilon.$$

❑

Wir haben gesehen, dass stetige Funktionen auf kompakten Mengen einige sehr angenehme Eigenschaften haben. Das folgende Beispiel zeigt, dass die Kompaktheit des Definitionsbereiches wirklich wesentlich ist.

2.1.15 Beispiel. Sei $E \subseteq \mathbb{R}$ nicht kompakt. Dann gilt

(i) Es gibt eine auf E stetige Funktion die nicht beschränkt ist.

(ii) Es gibt eine auf E stetige und beschränkte Funktion, die kein Maximum hat.

(iii) Ist E beschränkt, so gibt es eine auf E stetige, aber nicht gleichmäßig stetige Funktion.

Wir betrachten zuerst den Fall, daß E einen Häufungspunkt x_0 hat, der nicht zu E gehört (dieser Fall tritt sicher immer dann ein, wenn E beschränkt ist): Die Funktion $f(x) := \frac{1}{x-x_0}$ ist stetig auf E, aber nicht beschränkt. Sie ist nicht gleichmäßig stetig. Die Funktion $F(x) := \frac{1}{1+(x-x_0)^2}$ ist stetig auf E und beschränkt ($0 < f(x) < 1$). Offenbar gilt $\lim_{x\to x_0} f(x_0) = 1$, also $\sup_{x\in E} f(x) = 1$.

Betrachte nun den Fall, daß E nicht beschränkt ist. (i) folgt mit $f(x) = x$, (ii) mit $f(x) = \frac{x^2}{1+x^2}$.

In (iii) kann man die Forderung das E beschränkt ist nicht ganz weglassen. Zum Beispiel betrachte $E = \mathbb{Z}$. Dann ist jede Funktion auf E gleichmäßig stetig, denn man kann stets irgendein $\delta < 1$ wählen, z.B. $\delta = \frac{1}{2}$. ❑

2.2 Unstetigkeitsstellen reeller Funktionen

Für reelle Funktionen kann man Unstetigkeitsstellen noch genauer klassifizieren. Dazu definieren wir den Begriff des links- bzw. rechtsseitigen Limes.

2.2.1 Definition. Sei $\langle Y, d_Y\rangle$ ein metrischer Raum, seien $a, b \in \mathbb{R}$, $a < b$, sei $f : [a, b) \to Y$ eine Funktion, und sei $y_0 \in Y$. Dann sagen wir der *linksseitige Grenzwert* von f nach b existiert und ist gleich y_0, wenn gilt

$$\forall\epsilon > 0\,\exists\delta > 0 :\ d_Y(f(x), y_0) < \epsilon \text{ wenn } x \in [a, b), |x - b| < \delta\,.$$

In diesem Fall schreibt man $\lim_{x\to b-} f(x) = y_0$.

In analoger Weise definiert man den *rechtsseitigen Grenzwert*, ist nämlich $f : (a, b] \to Y$, dann sagt man $\lim_{x\to a+} f(x) = y_0$ wenn

$$\forall\epsilon > 0\,\exists\delta > 0 :\ d_Y(f(x), y_0) < \epsilon \text{ wenn } x \in (a, b], |x - b| < \delta\,.$$

2.2.2 Bemerkung.

(i) Man kann, genauso wie in Proposition 2.1.4, die Tatsache dass $\lim_{x\to b-} f(x) = y_0$ dadurch charakterisieren, dass für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in [a, b)$, $x_n \to b$, gilt dass $f(x_n) \to y_0$. Analog für den rechtsseitigen Limes.

(ii) Sei $f : (a, b) \to Y$, und sei $x_0 \in (a, b)$, $y_0 \in Y$. Dann gilt $\lim_{x\to x_0} f(x) = y_0$ genau dann wenn $\lim_{x\to x_0-} f(x) = y_0$ und $\lim_{x\to x_0+} f(x) = y_0$.

2.2.3 Definition. Sei f definiert auf (a, b), sei $x \in (a, b)$ und f unstetig bei x. Man sagt f hat eine *Unstetigkeit 1.Art* bei x falls $\lim_{t\to x-} f(t)$ und $\lim_{t\to x+} f(t)$ existieren (aber nicht beide gleich $f(x)$ sind). Andernfalls spricht man von einer *Unstetigkeit 2.Art*.

Für Unstetigkeiten 1.Art gibt es zwei Möglichkeiten. Entweder $\lim_{t\to x-} f(t) \neq \lim_{t\to x-} f(t)$, in welchem Fall man von einer *Sprungstelle* spricht, oder $\lim_{t\to x-} f(t) = \lim_{t\to x+} f(t) \neq f(x)$, dann spricht man von einer *hebbaren Unstetigkeit*.

2.2.4 Beispiel.

(i) Betrachte die Funktion

$$f(x) = \begin{cases} 1 & , \ x \text{ rational} \\ 0 & , \ x \text{ irrational} \end{cases}$$

Diese Funktion hat an jeder Stelle x eine Unstetigkeit 2.Art.

(ii) Sei

$$f(x) = \begin{cases} x & , \ x \text{ rational} \\ 0 & , \ x \text{ irrational} \end{cases}$$

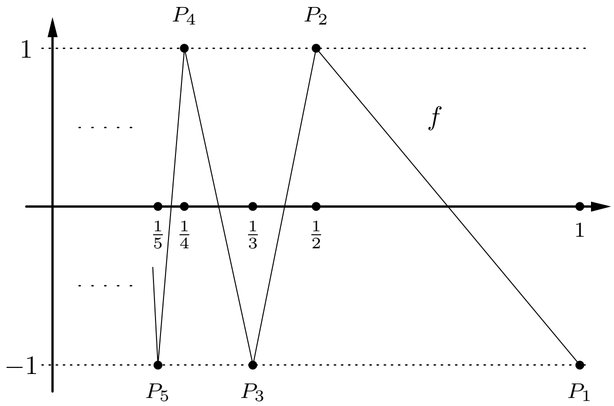
Dann ist f stetig bei 0, hat aber an jeder Stelle $x \neq 0$ eine Unstetigkeit 2.Art.

(iii) Sei

$$f(x) = \begin{cases} x+2 & , \ -3 < x < -2 \\ -x-2 & , \ -2 \leq x < 0 \\ x+2 & , \ 0 \leq x < 1 \end{cases}$$

Dann ist f stetig auf $(-3, 1) \setminus \{0\}$ und hat bei 0 eine Sprungstelle.

(iv) Betrachte die Funktion $f : [0, 1] \to \mathbb{R}$ die wie folgt definiert ist. Bezeichne mit P_n, $n \in \mathbb{N}$, den Punkt $P_n := (\frac{1}{n}, (-1)^n)$. Nun sei f der Polygonzug der durch die Punkte P_n geht, und sei $f(0) := 0$:



Dann ist f stetig auf $(0, 1]$ und hat eine Unstetigkeit 2.Art bei 0.

Wir sehen, dass alle möglichen Varianten von Unstetigkeiten auftreten können.

Bei sogenannten monotonen Funktionen ist die Situation wesentlich einfacher als im Allgemeinen. Dazu definieren wir: Seien $a, b \in \mathbb{R}$, $a < b$, und $f : (a, b) \to \mathbb{R}$. Man sagt f ist *monoton wachsend* falls gilt

$$x < y \Rightarrow f(x) \leq f(y) \,.$$

Gilt sogar $x < y \Rightarrow f(x) < f(y)$, so sagt man f sei *streng monoton wachsend*

Analog sagt man f sei *monoton fallend* falls $x < y \Rightarrow f(x) \geq f(y)$. Sollte $x < y$ sogar $f(x) > f(y)$ implizieren, so spricht man analog von einer *streng monoton fallenden* Funktion.

Wir wollen hier noch bemerken, dass eine streng monotone Funktion stets injektiv ist. Nun kommen wir zur Diskussion der Unstetigkeitsstellen monotoner Funktionen.

2.2.5 Proposition. *Sei f monoton wachsend auf (a,b). Dann existieren für jeden Punkt $x \in (a,b)$ sowohl $\lim_{t\to x-} f(t)$ als auch $\lim_{t\to x+} f(t)$. Es gilt*

$$\sup_{a<t<x} f(t) = \lim_{t\to x-} f(t) \leq f(x) \leq \lim_{t\to x+} f(t) = \inf_{x<t<b} f(t). \tag{2.2}$$

Ist $x < y$ so gilt

$$\lim_{t\to x+} f(t) \leq \lim_{t\to y-} f(t).$$

Eine analoge Aussage gilt für monoton fallende Funktionen.

Beweis. Die Menge aller Zahlen $f(t)$, $a < t < x$, ist durch $f(x)$ nach oben beschränkt, also existiert $A := \sup_{a<t<x} f(t) \in \mathbb{R}$. Klarerweise gilt $A \leq f(x)$.

Sei $\epsilon > 0$ gegeben. Dann existiert ein $\delta > 0$ so daß $A - \epsilon < f(x-\delta) \leq A$. Da f monoton ist gilt $f(x-\delta) \leq f(t)$ für alle t mit $x - \delta \leq t < x$. Also ist $|f(t) - A| < \epsilon$ für alle $t < x, |t-x| \leq \delta$, d.h. $\lim_{t\to x-} f(t) = A$. Die zweite Hälfte von (2.2) folgt analog.

Um den zweiten Teil der Behauptung einzusehen, genügt es zu bemerken daß gilt

$$\lim_{t\to x+} f(t) = \inf_{x<t<b} f(t) = \inf_{x<t<y} f(t),$$

$$\lim_{t\to y-} f(t) = \sup_{a<t<y} f(t) = \sup_{x<t<y} f(t).$$

❑

2.2.6 Korollar. *Sei $f : [a,b] \to [c,d]$ eine monoton wachsende Funktion mit $f(a) = c$, $f(b) = d$. Dann ist f stetig genau dann, wenn f surjektiv ist.*

Beweis. Ist f stetig, so nimmt f nach dem Zwischenwertsatz jeden Wert zwischen c und d an. Sei umgekehrt f surjektiv. Angenommen, f wäre an einer Stelle x nicht stetig, $\lim_{t\to x-} f(t) < \lim_{t\to x+} f(t)$. Dann gilt

$$f(t) \leq \lim_{t\to x-} f(t),\ t < x \text{ und } f(t) \geq \lim_{t\to x+} f(t),\ t > x\,.$$

Also kann f die Werte zwischen $\lim_{t\to x-} f(t)$ und $\lim_{t\to x+} f(t)$, bis auf unter Umständen einen (nämlich $f(x)$) nicht annehmen.

❑

2.2.7 Korollar. *Sei $f : (a,b) \to (c,d)$ eine streng monoton wachsende stetige Funktion mit $\lim_{x\to a+} f(x) = c$ und $\lim_{x\to b-} f(x) = d$. Dann ist f bijektiv, und die inverse Funktion $f^{-1} : (c,d) \to (a,b)$ ist ebenfalls streng monoton wachsend und stetig. Weiters gilt $\lim_{x\to c+} f^{-1}(x) = a$ und $\lim_{x\to d-} f^{-1}(x) = b$.*

Beweis. Streng monotone Funktionen sind stets injektiv. Stetigkeit ist daher äquivalent zur Bijektivität.

❑

2.2.8 Beispiel. Sei $n \in \mathbb{N}$, dann ist die Funktion $f(x) := x^n$ eine streng monoton wachsende Bijektion von $(0, \infty)$ auf $(0, \infty)$. Ihre Umkehrfunktion, das ist also $g(x) := \sqrt[n]{x}$, ist daher ebenfalls stetig. Weiters gilt $\lim_{x \to 0+} \sqrt[n]{x} = 0$ und $\lim_{x \to +\infty} \sqrt[n]{x} = +\infty$.

Wir wollen nun noch rationale Funktionen etwas genauer diskutieren. Sind $p, q \in \mathbb{C}[x]$ gegeben, $q \neq 0$, so ist durch $f(x) = \frac{p(x)}{q(x)}$ eine stetige Funktion $f : \{x \in \mathbb{C} : q(x) \neq 0\} \to \mathbb{C}$ definiert. Was passiert an den Nullstellen von q?

2.2.9 Proposition. *Seien $p, q \in \mathbb{C}[x]$, $q \neq 0$, und setze $f(x) = \frac{p(x)}{q(x)}$ für $q(x) \neq 0$. Sei $x_0 \in \mathbb{C}$ eine Nullstelle von q, dann gilt entweder*

- *der Grenzwert $\lim_{x \to x_0} f(x)$ existiert in $\mathbb{C}$*

oder

- *es ist $\lim_{x \to x_0} |f(x)| = +\infty$.*

Beweis. Schreibe $p(x) = a \prod (x - a_n)^{\alpha_n}$ und $q(x) = b \prod (x - b_n)^{\beta_n}$ wobei a_n und b_n die Nullstellen von p bzw. von q bezeichnen. Sei oBdA $x_0 = b_1$.

Fall 1: Angenommen $p(x_0) \neq 0$, dann gilt

$$f(x) = (x - x_0)^{-\beta_1} \cdot a \prod_{n \geq 2} (x - a_n)^{\alpha_n} \Big(b \prod_{n \geq 2} (x - b_n)^{\beta_n} \Big)^{-1}$$

und daher $\lim_{x \to x_0} |f(x)| = +\infty$.

Fall 2: Angenommen es ist $p(x_0) = 0$, oBdA $x_0 = a_1$, und es gilt $\alpha_1 < \beta_1$. Dann hat man

$$f(x) = a \prod (x - a_n)^{\alpha_n} \Big(b \prod (x - b_n)^{\beta_n} \Big)^{-1} =$$

$$= (x - x_0)^{\alpha_1 - \beta_1} \cdot a \prod_{n \geq 2} (x - a_n)^{\alpha_n} \Big(b \prod_{n \geq 2} (x - b_n)^{\beta_n} \Big)^{-1}$$

und es folgt wieder $\lim_{x \to x_0} |f(x)| = +\infty$.

Fall 3: Sei schliesslich $p(x_0) = 0$, oBdA $x_0 = a_1$, wobei $\alpha_1 \geq \beta_1$. Dann zeigt die gleiche Darstellung wie im letzten Fall, dass

$$\lim_{x \to x_0} f(x) = \begin{cases} 0 & , \ \alpha_1 > \beta_1 \\ a \prod_{n \geq 2} (x_0 - a_n)^{\alpha_n} \Big(b \prod_{n \geq 2} (x_0 - b_n)^{\beta_n} \Big)^{-1} & , \ \alpha_1 = \beta_1 \end{cases}$$

❑

In ähnlicher Weise kann man sich nach dem Verhalten von $f(z)$ fragen, wenn $|z|$ immer größer wird.

2.2.10 Proposition. *Seien $p, q \in \mathbb{C}[x]$, $q \neq 0$, und setze $f(x) = \frac{p(x)}{q(x)}$ für $q(x) \neq 0$. Dann gilt entweder*

- *der Grenzwert $\lim_{|z|\to\infty} f(x)$ existiert in $\mathbb{C}$ (d.h. es gibt eine komplexe Zahl w mit*

$$\forall \epsilon > 0\, \exists R > 0 : \; |f(z) - c| < \epsilon \text{ für alle } z \text{ mit } |z| > R$$

oder

- *es ist $\lim_{|z|\to\infty} |f(x)| = +\infty$.*

Beweis. Ist $p = 0$, so ist die Behauptung trivial. Schreibe $p(z) = a_n z^n + \ldots + a_1 z + a_0$ und $q(z) = b_m z^m + \ldots + b_1 z + b_0$ mit $a_n, b_m \neq 0$. Dann ist

$$f(z) = z^{n-m} \frac{a_n + a_{n-1} z^{-1} + \ldots + a_0 z^{-n}}{b_m + b_{m-1} z^{-1} + \ldots + b_0 z^{-m}},$$

und wir sehen, dass

$$\lim_{|z|\to\infty} f(z) = \begin{cases} 0 & , \; n < m \\ \frac{a_n}{b_m} & , \; n = m \end{cases}$$

und $\lim_{|z|\to\infty} |f(z)| = \infty$ falls $n > m$.

❑

2.3 Gleichmäßige Konvergenz

Wir haben schon gesehen, dass z.B. die geometrische Reihe für jedes q, $0 \leq q < 1$, konvergiert. Betrachtet man diese Reihe nicht nur für ein festes vorgegebenes q, sondern sozusagen für alle q auf einmal, so hat man also eine Reihe deren Summanden Funktionen von q sind (nämlich q^n), und deren Summe ebenfalls eine Funktion von q ist (nämlich $\frac{1}{1-q}$).

Betrachten wir also eine Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen die definiert ist, zum Beispiel, auf einer Menge $E \subseteq \mathbb{R}$ und die, zum Beispiel, reelle Werte annimmt. Wir würden gerne erklären was es bedeutet das diese Folge gegen eine Funktion f konvergiert.

Um einen „vernünftigen“ Grenzwertbegriff zu bekommen, müßten wir jetzt also eine Metrik definieren, und zwar eine Metrik auf einer Menge von Funktionen $f : E \to \mathbb{R}$. Aber könnte man in diesem konkreten Fall nicht auch naiver an die Sache herangehen? Ist $(f_n)_{n\in\mathbb{N}}$, $f_n : E \to \mathbb{R}$, eine Folge von Funktionen, dann ist ja für jedes feste $x \in E$ sicher $(f_n(x))_{n\in\mathbb{N}}$ eine Folge von Zahlen. Und für diese wissen wir was es bedeutet zu konvergieren. Man könnte also definieren:

2.3.1 Definition. Sei E eine Menge und $\langle Y, d_Y \rangle$ ein metrischer Raum. Eine Folge $(f_n)_{n\in\mathbb{N}}$, $f_n : E \to Y$, heißt *punktweise konvergent* gegen die Funktion $f : E \to Y$, wenn für jedes feste $x \in E$ gilt $\lim_{n\to\infty} f_n(x) = f(x)$.

Es entsteht die Frage, ob sich Eigenschaften der Funktionen f_n auf die Grenzfunktion f übertragen. Zum Beispiel die, ganz fundamentale, Eigenschaft der Stetigkeit. Das folgende Beispiel illustriert dass in dieser Angelegenheit etwas schiefgehen kann.

2.3.2 Beispiel. Betrachte die Funktionen $f_n : \mathbb{R} \to \mathbb{R}$ die gegeben sind als $f_n(x) := \frac{x^2}{(1+x^2)^n}$. Dann ist für jedes $x \neq 0$ die Reihe

$$\sum_{n=0}^{\infty} \frac{x^2}{(1+x^2)^n}$$

eine konvergente geometrische Reihe. Ihre Summe $f(x)$ läßt sich berechnen als

$$f(x) = 1 + x^2,\ x \neq 0.$$

Für $x = 0$ sind alle Summanden $= 0$, also auch ihre Summe. Man erhält daher

$$\sum_{n=0}^{\infty} f_n(x) = \begin{cases} 1 + x^2 & x \neq 0 \\ 0 & x = 0 \end{cases} \tag{2.3}$$

Jede der Funktionen f_n, und damit auch alle Partialsummen obiger Reihe sind stetige Funktionen von $x \in \mathbb{R}$. Nicht jedoch die Grenzfunktion (2.3).

Es war also vielleicht doch keine so gute Idee einfach naiv an die Sache heranzugehen.

2.3.3 Definition. Sei E eine Menge und $\langle Y, d_Y \rangle$ ein metrischer Raum. Wir bezeichnen mit $\mathcal{B}(E,Y)$ die Menge aller *beschränkten Funktionen* $f : E \to Y$, d.h.

$$\mathcal{B}(E,Y) := \big\{ f : E \to Y : \exists R > 0, y_0 \in Y : d_Y(f(x), y_0) \leq R, x \in E \big\}.$$

Man definiert nun eine Abbildung

$$d_\infty : \begin{cases} \mathcal{B}(E,Y) \times \mathcal{B}(E,Y) & \to \quad \mathbb{R} \\ (f,g) & \mapsto \quad \sup_{x \in E} d_Y(f(x), g(x)) \end{cases}$$

und spricht von der *Supremumsmetrik*.

2.3.4 Lemma. *Es ist tatsächlich d_∞ eine Metrik.*

Beweis. Es gilt $d_\infty(f,f) = \sup_{x\in E} d_Y(f(x), f(x)) = \sup_{x \in E} 0 = 0$. Ist umgekehrt $d_\infty(f,g) = \sup_{x\in E} d_Y(f(x), g(x)) = 0$, so folgt $d_Y(f(x), g(x)) = 0$ für jedes $x \in E$ und daher $f(x) = g(x)$, $x \in E$. D.h. $f = g$. Wegen $d_Y(f(x), g(x)) = d_Y(g(x), f(x))$ folgt das auch

$$d_\infty(f,g) = \sup_{x \in E} d_Y(f(x), g(x)) = \sup_{x\in E} d_Y(g(x), f(x)) = d_\infty(g,f).$$

Sei nun weiters eine Funktion h gegeben. Es gilt für jedes feste $x \in E$

$$d_Y(f(x), g(x)) \leq d_Y(f(x), h(x)) + d_Y(h(x), g(x)),$$

und daher auch

$$d_Y(f(x), g(x)) \leq \sup_{x \in E} d_Y(f(x), h(x)) + \sup_{x\in E} d_Y(h(x), g(x)) = d_\infty(f,h) + d_\infty(h,g).$$

Da diese Beziehung ja für jedes $x \in E$ gilt, folgt auch $d_\infty(f,g) = \sup_{x\in E} d_Y(f(x), g(x)) \leq d_\infty(f,h) + d_\infty(h,g)$.

❑

Wir wissen also was es bedeutet dass eine Folge von Funktionen in dieser Metrik konvergiert, nämlich gilt $\lim_{n\to\infty} f_n = f$ bezüglich d_∞, wenn

$$\forall \epsilon > 0 \exists N \in \mathbb{N} : d_\infty(f_n, f) < \epsilon, n \geq N \tag{2.4}$$

Man sagt dafür die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiere *gleichmäßig* gegen f.

Der Unterschied zwischen punktweiser und gleichmäßiger Konvergenz liegt (genauso wie bei der gleichmäßigen Stetigkeit) wieder darin begründet, dass man ein N finden muß das für alle $x \in E$ funktioniert. Denn (2.4) ist ja äquivalent zu

$$\forall \epsilon > 0\, \exists N \in \mathbb{N}\, \forall x \in E : d_Y(f_n(x), f(x)) < \epsilon \text{ für } n \geq N\,,$$

wogegen punkteweise Konvergenz bedeutet

$$\forall x \in E\, \forall \epsilon > 0\, \exists N \in \mathbb{N} : d_Y(f_n(x), f(x)) < \epsilon \text{ für } n \geq N\,.$$

Insbesondere sehen wir, dass jede gleichmäßig konvergente Folge auch punktweise konvergiert , und zwar zur gleichen Grenzfunktion.

Gleichmäßige Konvergenz sichert nun dass sich Grenzübergänge „brav“ verhalten.

2.3.5 Satz. *Seien $\langle X, d_x\rangle$ und $\langle Y, d_y\rangle$ metrische Räume, wobei Y vollständig sei. Sei weiters $E \subseteq X$, und seien $f_n : E \to Y$, $n \in \mathbb{N}$, sowie $f : E \to Y$ beschränkte Funktionen. Schliesslich sei x ein Häufungspunkt von E, und $(x_m)_{m\in\mathbb{N}}$ eine Folge von Punkten aus E mit $x_n \to x$.*

Ist die Folge $(f_n)_{n\in\mathbb{N}}$ auf E gleichmäßig konvergent zu f, und gilt für jedes $n \in \mathbb{N}$

$$\lim_{m\to\infty} f_n(x_m) = A_n\,,$$

dann ist die Folge $(A_n)_{n\in\mathbb{N}}$ konvergent, und es gilt

$$\lim_{m\to\infty} f(x_m) = \lim_{n\to\infty} A_n\,.$$

Beweis.

Schritt 1: Wir zeigen, dass die Folge $(A_n)_{n\in\mathbb{N}}$ konvergiert.

Sei $\epsilon > 0$ gegeben. Da die Folge $(f_n)_{n\in\mathbb{N}}$ gleichmäßig konvergiert existiert $N \in \mathbb{N}$, sodaß für $n, k \geq N$ und alle $t \in E$ gilt

$$d_Y\big(f_n(t), f_k(t)\big) \leq \epsilon\,.$$

Hält man n und k fest, so erhält man also für $m \to \infty$ die Beziehung $d_Y(A_n, A_m) \leq \epsilon$. Damit ist $(A_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in Y und daher konvergent, $\lim_{n\to\infty} A_n =: A$.

Schritt 2: Wir zeigen die behauptete Limesbeziehung. Es gilt

$$d_Y(f(t), A) \leq d_Y(f(t), f_n(t)) + d_Y(f_n(t), A_n) + d_Y(A_n, A)\,.$$

Wähle N so groß, daß $d_Y(f(t), f_n(t)) < \epsilon$ für alle $t \in E$, und so daß $d_Y(A_N, A) < \epsilon$. Für dieses N existiert M, sodaß $d_Y(f_N(x_m)), A_N) < \epsilon$, $m \geq M$. Insgesamt erhalten wir

$$d_Y(f(x_m), A) < 3\epsilon,\ m \geq M\,.$$

❑

Etwas schlampiger, aber wahrscheinlich suggestiver, läßt sich die Aussage von Satz 2.3.5 wie folgt formulieren: Ist $f_n \to f$ gleichmäßig und $x_m \to x$, so folgt

$$\lim_{m\to\infty} \lim_{n\to\infty} f_n(x_m) = \lim_{n\to\infty} \lim_{m\to\infty} f_n(x_m) .$$

2.3.6 Korollar. *Seien $\langle X, d_x\rangle$ und $\langle Y, d_y\rangle$ metrische Räume, wobei Y vollständig sei. Weiters seien $f_n : X \to Y$, $n \in \mathbb{N}$, sowie $f : X \to Y$ beschränkte Funktionen, und $x_0 \in X$. Sind alle Funktionen f_n stetig an der Stelle x_0, und ist $(f_n)_{n\in\mathbb{N}}$ gleichmäßig konvergent gegen f, so ist auch f stetig bei x_0.*

Beweis. Sei $(x_m)_{m\in\mathbb{N}}$ eine Folge von Punkten in X mit $x_m \to x_0$. Dann gilt nach dem letzten Satz

$$\lim_{m\to\infty} f(x_m) = \lim_{m\to\infty} \lim_{n\to\infty} f_n(x_m) = \lim_{n\to\infty} \lim_{m\to\infty} f_n(x_m) =$$

$$= \lim_{n\to\infty} f_n(x_0) = f(x_0) .$$

❑

2.3.7 Satz (Cauchy'sches Konvergenzkriterium)**.** *Sei E eine Menge und $\langle Y, d_y\rangle$ ein vollständiger metrischer Raum. Dann ist $\langle \mathcal{B}(E,Y), d_\infty\rangle$ vollständig.*

D.h. eine Folge $(f_n)_{n\in\mathbb{N}}$ von beschränkten Funktionen $f_n : E \to Y$ konvergiert genau dann gleichmäßig zu einer Funktion f, wenn es zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß für alle $n, m \geq N$ und beliebiges $x \in E$ gilt

$$d_Y(f_n(x), f_m(x)) < \epsilon .$$

Beweis. Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\langle \mathcal{B}(E,Y), d_\infty\rangle$. Dann ist wegen $d_Y(f_n(x), f_m(x)) \leq d_\infty(f_n, f_m)$ insbesondere für jedes einzelne x die Folge $(f_n(x))_{n\in\mathbb{N}}$ eine Cauchy-Folge in Y. Daher existiert für jedes feste x der Grenzwert $\lim_{n\to\infty} f_n(x)$. Definiere eine Funktion $f : E \to Y$ als

$$f(x) := \lim_{n\to\infty} f_n(x) .$$

Wir zeigen nun, daß die $f \in \mathcal{B}(E,Y)$ und $f_n \to f$ gleichmäßig.

Sei $\epsilon > 0$ gegeben und wähle $N \in \mathbb{N}$ sodaß $d_Y(f_n(x), f_m(x)) < \epsilon$ für alle $n, m \geq N$ und alle $x \in E$. Hält man x fest und läßt $m \to \infty$ streben, so folgt für $n \geq N$

$$d_Y(f_n(x), f(x)) \leq \epsilon .$$

Da x beliebig war folgt $\sup_{x\in E} d_Y(f_n(x), f(x)) \leq \epsilon$, $n \geq N$.

❑

Wir geben im folgenden ein paar Methoden wie man Folgen bzw. Reihen auf gleichmäßige Konvergenz prüfen kann.

2.3.8 Korollar (Weierstraß Kriterium). *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n : E \to \mathbb{C}$. Es gelte*

$$|f_n(x)| \leq M_n,\ x \in E, n \in \mathbb{N},$$

mit gewissen Zahlen M_n für welche die Reihe $\sum_{n=1}^{\infty} M_n$ konvergiert. Dann ist die Reihe $\sum_{n=1}^{\infty} f_n(x)$ absolut und gleichmäßig konvergent auf E.

Beweis. Es gibt zu $\epsilon > 0$ ein $N \in \mathbb{N}$ so daß für $k, m \geq N$ gilt $\sum_{n=k}^{m} M_n < \epsilon$. Für solche m, k folgt auch

$$\left|\sum_{n=k}^{m} f_n(x)\right| \leq \sum_{n=k}^{m} |f_n(x)| \leq \sum_{n=k}^{m} M_n < \epsilon,\ x \in E,$$

und man kann das Cauchy-Kriterium anwenden.

❑

2.3.9 Satz (Dirichlet Kriterium). *Seien $(a_n(x))_{n\in\mathbb{N}}$ und $(b_n(x))_{n\in\mathbb{N}}$ Folgen von Funktionen $a_n : E \to \mathbb{R}$, $b_n : E \to \mathbb{C}$. Gilt für die Partialsummen $B_N(x) := \sum_{n=1}^{N} b_n(x)$*

$$|B_N(x)| \leq M,\ N \in \mathbb{N}, x \in E,$$

und sind die $a_n(x)$ eine punktweise monotone auf E gleichmäßig gegen 0 konvergierende Folge, dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n(x)b_n(x)$ gleichmäßig auf E.

Beweis. Analog wie für die Reihen mit konstanten Gliedern. Es gilt

$$|b_k(x) + \ldots + b_m(x)| = |B_m - B_{k-1}| \leq 2M,\ x \in E,$$

und für n hinreichend groß ist

$$|a_n(x)| < \epsilon,\ x \in E,$$

wegen der gleichmäßigen Konvergenz der Folge $(a_n(x))_{n\in\mathbb{N}}$.

❑

2.3.10 Korollar (Abelsches Kriterium). *Seien $(a_n(x))_{n\in\mathbb{N}}$ und $(b_n(x))_{n\in\mathbb{N}}$ Folgen von Funktionen $a_n : E \to \mathbb{R}$, $b_n : E \to \mathbb{C}$. Die Reihe $\sum_{n=1}^{\infty} b_n(x)$ konvergiere gleichmäßig auf E und die Funktionen $a_n(x)$ seien gleichmäßig beschränkt und punktweise monoton. Dann konvergiert die Reihe $\sum_{n=1}^{\infty} a_n(x)b_n(x)$ gleichmäßig auf E.*

Beweis. Ebenfalls analog wie bei den Reihen mit konstanten Gliedern.

❑

Ein bedeutendes Beispiel für Reihen von Funktionen sind sogenannte *Potenzreihen*:

$$P(z) = \sum_{n=0}^{\infty} a_n z^n,$$

wobei die Koeffizienten a_n irgendwelche reellen oder komplexen Zahlen sind.

Wir sind solchen Reihen schon begegnet, z.B. der geometrischen Reihe $\sum_{n=0}^{\infty} z^n$.

2.3.11 Satz. *Sei $P(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe. Es gibt eine Zahl R, $0 \leq R \leq \infty$, sodaß gilt:*

(*i*) *Für jedes $r < R$ ist $P(z)$ auf der kompakten Menge $\{z \in \mathbb{C} : |z| \leq r\}$ absolut und gleichmäßig konvergent.*

(*ii*) *Für jedes z mit $|z| > R$ ist $P(z)$ divergent.*

Die Zahl R heißt der Konvergenzradius *von $P(z)$. Er läßt sich aus den Koeffizienten a_n bestimmen als*

$$\frac{1}{\limsup_{n\to\infty} |\frac{a_{n+1}}{a_n}|} \leq R = \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}} \leq \frac{1}{\liminf_{n\to\infty} |\frac{a_{n+1}}{a_n}|}$$

Beweis. Wir zeigen zuerst die folgende Behauptung. Ist $P(z)$ konvergent für ein gewisses z_0, und ist $r < |z_0|$, so ist $P(z)$ auf $\{z : |z| \leq r\}$ absolut und gleichmäßig konvergent. Es gilt ($|a_n z_0^n| \leq C$)

$$|a_n z^n| = |a_n z_0^n| \cdot |\frac{z}{z_0}|^n \leq C \cdot |\frac{r}{z_0}|^n$$

Wegen $|\frac{r}{z_0}| < 1$ ist $\sum_{n=0}^{\infty} C |\frac{r}{z_0}|^n$ konvergent. Wir können also das Weierstraßsche Kriterium anwenden.

Sei nun $R := \sup\{|z| : P(z) \text{ ist konvergent }\}$. Nach Definition gilt (*ii*). Sei also $r < R$ gegeben. Es existiert ein z mit $r < z < R$, sodaß $P(z)$ konvergiert. Nach der obigen Zwischenbehauptung folgt (*i*).

Ist $|z|$ kleiner als einer der beiden linken angeführten Werte, so folgt

$$|z| \cdot \limsup_{n\to\infty} |\frac{a_{n+1}}{a_n}| = \limsup_{n\to\infty} |\frac{a_{n+1} z^{n+1}}{a_n z^n}| < 1,$$

bzw.

$$|z| \cdot \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \limsup_{n\to\infty} \sqrt[n]{|a_n z^n|} < 1,$$

d.h. $P(z)$ konvergiert (entweder nach dem Quotienten- oder nach dem Wurzelkriterium). D.h. jeder der beiden linken Brüche ist $\leq R$. Ist $|z|$ jedoch größer als einer der rechten Brüche, so gilt wieder nach dem Quotienten- bzw. Wurzelkriterium, daß $P(z)$ divergiert.

❑

2.3.12 Beispiel. Das Konvergenzverhalten einer Potenzreihe ist durch den obigen Satz relativ gut abgeklärt. Einzig über Punkte mit $|z| = R$, wo R der Konvergenzradius ist, hat man keine Aussage. Es können hier tatsächlich auch alle Fälle eintreten.

Betrachte die Potenzreihen $P_1(z) = \sum_{n=0}^{\infty} z^n$, $P_2(z) = \sum_{n=0}^{\infty} \frac{z^n}{n}$ und $P_3(z) = \sum_{n=0}^{\infty} \frac{z^n}{n^2}$. Alle haben Konvergenzradius 1. Jedoch ist $P_1(z)$ für $|z| = 1$ divergent, $P_3(z)$ absolut konvergent, und $P_2(z)$ (nicht absolut) konvergent außer bei $z = 1$, wo $P_2(z)$ divergiert.

Sei $P(z)$ eine Potenzreihe mit Konvergenzradius R, $0 < R < \infty$. Dann ist auf Grund der gleichmäßigen Konvergenz die Funktion $P(z)$ auf der Menge $\{z : |z| < R\}$ stetig, d.h. für $|z| < R$ ist $\lim_{t\to z} P(t) = P(z)$.

2.3.13 Satz (Abelscher Grenzwertsatz)**.** *Sei $P(z)$ eine Potenzreihe, R ihr Konvergenzradius und $0 < R < \infty$. Weiters sei $z_0 \in \mathbb{C}$ mit $|z_0| = R$. Ist $P(z_0)$ konvergent, so gilt*

$$\lim_{t \to 1-} P(tz_0) = P(z_0).$$

Beweis. Betrachte den Fall $z_0 = R$ (anderen Fälle analog). Schreibt man

$$\sum_{n=1}^{\infty} a_n z^n = \sum_{n=1}^{\infty} a_n R^n (\frac{z}{R})^n,$$

so ist wegen

$$1 \geq \frac{z}{R} \geq (\frac{z}{R})^2 \geq \ldots \geq 0,\ z \in [0, R]\,,$$

das Abelsche Kriterium anwendbar, und wir erhalten daß die Reihe $\sum_{n=1}^{\infty} a_n z^n$ auf $[0, R]$ gleichmäßig konvergiert.

Die dargestellte Funktion $P(z)$ ist daher auf dem ganzen Intervall $[0, R]$ stetig.

❑

Stellt man sich Potenzreihen als „Polynome mit unendlich vielen Koeffizienten“ vor, so wird ein Quotient von zwei Potenzreihen das Analogon zu einer rationalen Funktion sein. Wir wollen nun, analog zu Proposition 2.2.9, das lokale Verhalten eines Quotienten von zwei Potenzreihen betrachten.

2.3.14 Proposition. *Seien $p(z) = \sum_{n=k}^{\infty} a_n z^n$, $a_k \neq 0$, und $q(z) = \sum_{n=l}^{\infty} b_n z^n$, $b_l \neq 0$, zwei Potenzreihen mit positiven Konvergenzradien R_p bzw. R_q. Dann existiert $R > 0$ sodass $q(z) \neq 0$ für $z \in U_R(0) \setminus \{0\}$. Setze $f(z) := \frac{p(z)}{q(z)}$, $z \in U_R(0) \setminus \{0\}$. Dann gilt*

$$\begin{cases} \lim_{z\to 0} f(z) = 0 & ,\ k > l \\ \lim_{z\to 0} f(z) = \frac{a_k}{b_l} & ,\ k = l \\ \lim_{z\to 0} |f(z)| = +\infty & ,\ k < l \end{cases}$$

Beweis. Setze $p_1(z) := \sum_{n=k}^{\infty} a_n z^{n-k}$, $q_1(z) = \sum_{n=l}^{\infty} b_n z^{n-l}$. Dann ist p_1 bzw. q_1 eine Potenzreihen, und der Konvergenzradius von p_1 ist R_p, der von q_1 ist R_q. Da q_1 stetig auf der Kreisscheibe $U_{R_q}(0)$ ist und $q_1(0) = b_l \neq 0$, existiert $R > 0$ sodass $q_1(z) \neq 0$ für $z \in U_R(0)$. Wegen $q(z) = z^l q_1(z)$ folgt $q(z) \neq 0$ für $z \in U_R(0) \setminus \{0\}$.

Die Funktion f läßt sich schreiben als

$$f(z) = \frac{p(z)}{q(z)} = \frac{z^k p_1(z)}{z^l q_1(z)} = z^{k-l} \frac{p_1(z)}{q_1(z)}\,,$$

und die gewünschte Limesbeziehungen folgen.

❑

2.4 Der Satz von Arzela-Ascoli

Wir haben gesehen, daß jede beschränkte Folge von Zahlen eine konvergente Teilfolge besitzt. Gelten ähnliche Resultate für „beschränkte" Folgen von Funktionen ? Wir beschränken uns auf die Betrachtung reeller Funktionen.

2.4.1 Definition. Sei $(f_n)_{n\in\mathbb{N}}$, $f_n : E \to \mathbb{R}$ eine Folge von Funktionen definiert auf einer Menge $E \subseteq \mathbb{R}$. Die Familie $\{f_n : n \in \mathbb{N}\}$ heißt *punktweise beschränkt*, wenn für jedes $x \in E$ die Menge $\{f_n(x) : n \in \mathbb{N}\}$ beschränkt ist. D.h. wenn es eine Funktion ϕ auf E gibt mit

$$|f_n(x)| \leq \phi(x),\ \ n \in \mathbb{N}, x \in E.$$

Die Familie $\{f_n : n \in \mathbb{N}\}$ heißt *gleichmäßig beschränkt*, wenn es eine Zahl M gibt mit

$$|f_n(x)| \leq M,\ \ n \in \mathbb{N}, x \in E.$$

2.4.2 Beispiel.

(i) Es gibt eine gleichmäßig beschränkte Folge von Funktionen, die keine punktweise konvergente Teilfolge enthält. Betrachte $f_n(x) := \sin nx$ auf $0 \leq x \leq 2\pi$ für jedes $n \in \mathbb{N}$ (Beweis dieser Tatsache später).

(ii) Es gibt eine gleichmäßig beschränkte und punktweise konvergente Folge die keine gleichmäßig konvergente Teilfolge enthält. Betrachte $f_n(x) := \frac{x^2}{x^2+(1-nx)^2}$ auf $x \in [0,1]$ für $n \in \mathbb{N}$. Dann ist $|f_n(x)| \leq 1$ und $\lim_{n\to\infty} f_n(x) = 0$ punktweise. Jedoch gilt $f_n(\frac{1}{n}) = 1$, also kann keine Teilfolge gleichmäßig konvergieren.

Beachte die triviale Tatsache, daß jede gleichmäßig gegen eine beschränkte Funktion konvergente Folge auch gleichmäßig beschränkt ist.

2.4.3 Definition. Eine Familie $\mathcal{F}$ von Funktionen auf einer Menge E heißt *gleichgradig stetig*, wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt mit $(x, y \in E)$

$$|f(x) - f(y)| < \epsilon, |x - y| < \delta,\ \ f \in \mathfrak{F}.$$

Klarerweise ist jedes Element f einer gleichgradig stetigen Familie von Funktionen selbst stetig.

2.4.4 Satz (Arzela[1]-Ascoli[2])**.** *Sei $(f_n)_{n\in\mathbb{N}}$ punktweise beschränkt und gleichgradig stetig auf der kompakten Menge K. Dann ist $\{f_n : n \in \mathbb{N}\}$ gleichmäßig beschränkt und enthält eine gleichmäßig konvergente Teilfolge.*

Beweis. Eine kompakte Menge enthält eine abzählbare dichte Teilmenge E: Sei $J(x,r) = \{y \in K : |x-y| < r\}$. Da für jedes feste r diese Mengen K überdecken und K kompakt ist erhält man $K \subseteq J(x_1, r) \cup \ldots \cup J(x_m, r)$ wobei m von r abhängt. Für $r = 1, \frac{1}{2}, \frac{1}{3}, \ldots$ wähle man aus jeder Menge $J(x_m, r)$ einen Punkt. Die Menge E aller dieser Punkte ist dicht in K.

Seien $x_1, x_2, x_3, \ldots$ die Punkte aus E irgendwie abgezählt. Da $\{f_n(x_1) : n \in \mathbb{N}\}$ beschränkt ist, existiert eine Teilfolge die konvergiert. Wir bezeichnen die

[1]Cesare Arzela. 16.3.1847 St.Stefano di Magra (Italien) - 15.3.1912 St.Stefano di Magra
[1]Giulio Ascoli. 20.11.1843 Triest - 12.7.1896 Mailand

in dieser Teilfolge auftretenden Funktionen mit $f_{1,k}$, d.h. $(f_{1,k}(x_1))_{k\in\mathbb{N}}$ konvergiert. Da $\{f_{1,k}(x_2) : k \in \mathbb{N}\}$ beschränkt ist, gibt es eine konvergente Teilfolge $(f_{1,k_l}(x_2))_{l\in\mathbb{N}}$. Wähle diese so, daß $k_{l+1} > k_l$. Wir bezeichnen die auftretenden Funktionen $\{f_{1,k_l}\}$ mit $\{f_{2,k}\}$. Verfährt man induktiv so weiter, erhält man ein Schema von Folgen $S_1, S_2, S_3, \ldots$

$$\begin{array}{lllll} S_1: & f_{1,1} & f_{1,2} & f_{1,3} & \cdots \\ S_2: & f_{2,1} & f_{2,2} & f_{2,3} & \cdots \\ S_3: & f_{3,1} & f_{3,2} & f_{3,3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{array}$$

mit den folgenden Eigenschaften:

(i) S_{n+1} ist eine Teilfolge von S_n.

(ii) $(f_{n,k}(x_n))_{k\in\mathbb{N}}$ konvergiert.

(iii) Treten zwei Funktionen in S_n und S_{n+1} auf, so in der gleichen Reihenfolge. D.h. Funktionen bewegen sich immer nach links wenn man das Schema nach unten geht.

Betrachte die Diagonalfolge $S : (f_{k,k})_{k\in\mathbb{N}}$. Sei $i \in \mathbb{N}$ fest. Da S bis auf die ersten $i-1$ Terme eine Teilfolge von S_i ist, ist $(f_{k,k}(x_i))_{k\in\mathbb{N}}$ konvergent.

Sei $\epsilon > 0$ gegeben. Da die Familie $\{f_n : n \in \mathbb{N}\}$ gleichgradig stetig ist existiert $\delta > 0$, sodaß

$$|f_n(x) - f_n(y)| < \epsilon,\ n \in \mathbb{N}, |x-y| < \delta.$$

Sei $J(x,\delta)$ wieder die Kugel um x mit Radius δ und sei P so groß, daß $K \subseteq J(x_1,\delta) \cup \ldots \cup J(x_p,\delta)$. Weiters sei N so groß, daß

$$|f_{n,n}(x_i) - f_{m,m}(x_i)| < \epsilon,\ i = 1,\ldots,P, n,m \geq N.$$

Sei $x \in K$. Dann ist $x \in J(x_i,\delta)$ für ein gewisses i, und daher gilt für $n,m \geq N$

$$|f_{n,n}(x) - f_{m,m}(x)| \leq |f_{n,n}(x) - f_{n,n}(x_i)| + |f_{n,n}(x_i) - f_{m,m}(x_i)| +$$

$$+ |f_{m,m}(x_i) - f_{m,m}(x)| < 3\epsilon.$$

Also konvergiert $(f_{k,k})_{k\in\mathbb{N}}$ gleichmäßig auf K.

Setze $\phi(x) := \sup_{x\in K} |f_n(x)|$. Zu $\epsilon > 0$ wähle $\delta > 0$, sodaß $|f_n(x) - f_n(y)| < \epsilon$ für $|x-y| < \delta$, $n \in \mathbb{N}$. Hält man x und y fest, so folgt $||f_n(y)| - |f_n(x)|| < \epsilon$, $n \in \mathbb{N}$, und damit auch

$$|\phi(x) - \phi(y)| \leq \epsilon.$$

Also ist $\phi(x)$ stetig. Da K kompakt ist also auch beschränkt. D.h. $\{f_n : n \in \mathbb{N}\}$ ist gleichmäßig beschränkt.

❑

2.5 Die Exponentialfunktion

Die Exponentialfunktion $\exp : \mathbb{C} \to \mathbb{C}$ ist definiert als

$$\exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}, \; z \in \mathbb{C}. \tag{2.5}$$

Wir haben schon gesehen, dass diese Reihe für alle $z \in \mathbb{C}$ konvergiert. Sie ist also eine Potenzreihe mit Konvergenzradius ∞. Insbesondere ist $\exp : \mathbb{C} \to \mathbb{C}$ stetig. Weiters haben wir schon gesehen, dass sie der Funktionalgleichung

$$\exp(z+w) = \exp(z)\exp(w), \; z, w \in \mathbb{C},$$

genügt.

2.5.1 Satz. *Es gilt:*

(*i*) *Es ist für jedes* $z \in \mathbb{C}$

$$\exp(z) \neq 0.$$

(*ii*) *Es ist für jedes* $z \in \mathbb{C}$

$$\lim_{w \to z} \frac{\exp(w) - \exp(z)}{w - z} = \exp(z).$$

(*iii*) *Die Funktion* exp *eingeschränkt auf die reelle Achse ist eine streng monoton wachsende bijektive Funktion von* $\mathbb{R}$ *auf* $\mathbb{R}^+$. *Insbesondere gilt*

$$\lim_{x \to +\infty} \exp(x) = +\infty, \quad \lim_{x \to -\infty} \exp(x) = 0.$$

(*iv*) *Für jedes* $z \in \mathbb{C}$ *ist* $|\exp(z)| = \exp(\operatorname{Re} z)$. *Insbesondere gilt* $|\exp(z)| = 1$ *genau dann, wenn* $\operatorname{Re} z = 0$.

(*v*) *Es gibt eine eindeutig bestimmte positive reelle Zahl* π, *sodaß*

$$\exp(i\pi) + 1 = 0,$$

und daß

$$\exp(z) = 1 \iff \frac{z}{2\pi i} \in \mathbb{Z}.$$

Die Funktion $\exp(z)$ *ist periodische Funktion mit Periode* $2\pi i$, *d.h. es gilt stets*

$$\exp(z + 2\pi i) = \exp(z).$$

(*vi*) *Ist* $w \in \mathbb{C}$ *eine von* 0 *verschiedene Zahl, so ist* $w = \exp(z)$ *für eine geeignete Zahl* $z \in \mathbb{C}$.

Beweis.
`ad`(i): Wegen $\exp(z) \cdot \exp(-z) = \exp(0) = 1$ kann $\exp(z)$ nicht 0 sein.

`ad`(ii): Schreibe $w = z + h$. Es gilt

$$\lim_{h \to 0} \frac{\exp(z+h) - \exp(z)}{h} = e^z \lim_{h \to 0} \frac{\exp(h) - 1}{h} = \exp(z) \lim_{h \to 0} \sum_{n=1}^{\infty} \frac{h^{n-1}}{n!} = \exp(z),$$

da $\sum_{n=1}^{\infty} \frac{h^{n-1}}{n!}$ eine Potenzreihe mit Konvergenzradius ∞ ist, und daher bei 0 stetig.

`ad`(iii): Für $x \geq 0$ ist nach Definition $\exp(x)$ streng monoton wachsend, denn jeder Summand der Reihe (2.5) hat diese Eigenschaft. Weiters gilt $\lim_{x\to\infty} \exp(x) = +\infty$, denn $\exp(x) \geq 1 + x$, $x > 0$. Wegen $\exp(x)\exp(-x) = 1$ folgt das $\exp(x)$ auch für $x \leq 0$ streng monoton wachsend ist und das $\lim_{x\to-\infty} \exp(x) = 0$.

Es ist $\exp(x)$ als streng monotone Funktion sicher injektiv. Nun ist $\exp(x) : \mathbb{R} \to \mathbb{R}^+$ stetig. Wegen der obigen Limesbeziehungen und dem Zwischenwertsatz ist $\exp(x)$ auch surjektiv.

`ad`(iv): Nach Definition ist $\overline{\exp(z)} = \exp(\overline{z})$. Hier bezeichnet $\overline{z}$ die konjugiert komplexe Zahl. Es ist

$$|\exp(z)|^2 = \exp(z)\overline{\exp(z)} = \exp(z)\exp(\overline{z}) = \exp(z + \overline{z}) = \exp(2\operatorname{Re} z) = \exp(\operatorname{Re} z)^2$$

Da sowohl $|\exp(z)|$ als auch $\exp(\operatorname{Re} z)$ positive reelle Zahlen sind, folgt daraus $|\exp(z)| = \exp(\operatorname{Re} z)$.

`ad`(v): Sei $t \in \mathbb{R}$. Zerlegt man $\exp(it)$ in Real- und Imaginärteil, so erhält man

$$\exp(it) = \sum_{n=0}^{\infty} \frac{z^n}{n!} = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} = \sum_{n=0, n \text{ gerade}}^{\infty} i^n \frac{t^n}{n!} + \sum_{n=0, n \text{ ungerade}}^{\infty} i^n \frac{t^n}{n!} =$$

$$= \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!}\,. \tag{2.6}$$

Also haben wir

$$\operatorname{Re}\exp(it) = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} = 1 - \frac{t^2}{2} + \frac{t^4}{4!} - \frac{t^6}{6!} \pm \ldots\,,$$

$$\operatorname{Im}\exp(it) = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!} = t - \frac{t^3}{3!} + \frac{t^5}{5!} - \frac{t^7}{7!} \pm \ldots\,.$$

Wir haben $\operatorname{Re}\exp(0) = 1$. Betrachten wir nun $\operatorname{Re}\exp(2i)$. Da, für $t = 2$,

$$\frac{t^6}{6!} - \frac{t^8}{8!} \geq 0,\ \frac{t^{10}}{10!} - \frac{t^{12}}{12!} \geq 0, \quad \text{etc.}$$

gilt, folgt das $(t = 2)$

$$\operatorname{Re}\exp(it) = 1 - \frac{t^2}{2} + \frac{t^4}{4!} - \Big(\frac{t^6}{6!} - \frac{t^8}{8!}\Big) - \Big(\frac{t^{10}}{10!} - \frac{t^{12}}{12!}\Big) + \ldots \leq 1 - \frac{t^2}{2} + \frac{t^4}{4!} = -\frac{1}{3}\,,$$

also ist $\operatorname{Re}\exp(2i) < 0$. Da exp stetig ist, ist auch die Funktion

$$t \mapsto \operatorname{Re}\exp(it)$$

stetig. Nach dem Zwischenwertsatz (und noch einmal der Stetigkeit) gibt es eine kleinste Stelle $t_0 \in (0, 2)$ mit $\operatorname{Re}\exp(it_0) = 0$. Definiere

$$\pi := 2t_0\,.$$

Wegen $|\exp(i\frac{\pi}{2})| = 1$ und $\operatorname{Re}\exp(i\frac{\pi}{2}) = 0$, folgt $|\operatorname{Im}\exp(i\frac{\pi}{2})| = 1$. Nun gilt für $t \in [0, 2]$

$$\frac{t^5}{5!} - \frac{t^7}{7!} \geq 0, \quad \frac{t^9}{9!} - \frac{t^{11}}{11!} \geq 0, \quad \text{etc.}$$

also erhalten wir für solche t

$$\operatorname{Im}\exp(it) = t - \frac{t^3}{3!} + \big(\frac{t^5}{5!} - \frac{t^7}{7!}\big) + \big(\frac{t^9}{9!} - \frac{t^{11}}{11!}\big) + \ldots \geq t - \frac{t^3}{3!} \geq 0\,.$$

Wir schliessen, daß $\operatorname{Im}\exp(i\frac{\pi}{2}) = 1$. Insgesamt folgt also

$$\exp(i\frac{\pi}{2}) = i$$

und daher $\exp(i\pi) = -1$ und $\exp(2\pi i) = 1$.

Es folgt, dass für alle $k \in \mathbb{Z}$

$$\exp(z + 2k\pi i) = \exp(z)\exp(2\pi i)^k = \exp(z)\,.$$

Insbesondere ist $\exp(z)$ periodisch mit der Periode $2\pi i$.

Sei nun z mit $\exp(z) = 1$ gegeben. Schreibe $z = x + iy$ in Real- und Imaginärteil. Es ist $|\exp(z)| = \exp(\operatorname{Re} z) = 1$, und daher $\operatorname{Re} z = 0$. Es bleibt zu zeigen, daß $\frac{y}{2\pi} \in \mathbb{Z}$. Wegen der Periodizität genügt es zu zeigen, daß für $0 < y < 2\pi$ stets $\exp(iy) \neq 1$ ist. Sei also $0 < y < 2\pi$ gegeben. Schreibe

$$\exp(i\frac{y}{4}) = u + iv \text{ mit } u, v \in \mathbb{R}\,.$$

Wegen $0 < \frac{y}{4} < \frac{\pi}{2}$, und der Definition von $\frac{\pi}{2}$, ist $u, v > 0$. Weiters gilt

$$\exp(iy) = (u + iv)^4 = \big[u^4 - 6u^2v^2 + v^4\big] + i\big[4uv(u^2 - v^2)\big]\,.$$

Die rechte Zahl ist genau dann reell, wenn $u^2 - v^2 = 0$ ist. Wegen $u^2 + v^2 = 1$ ist das genau dann der Fall wenn $u^2 = v^2 = \frac{1}{2}$ ist. Dann ist aber $u^4 - 6u^2v^2 + v^4 = -1 \neq 1$.

Die Eindeutigkeit folgt nun, denn ist $\pi' \neq \pi$, so können $\frac{\pi'}{\pi}$ und $\frac{\pi}{\pi'}$ nicht beide ganze Zahlen sein.

`ad`(vi): Sei $w \in \mathbb{C}$, $w \neq 0$, gegeben. Da $\exp(x)$ eine Bijektion von $\mathbb{R}$ auf $\mathbb{R}^+$ ist, gibt es ein $x \in \mathbb{R}$ mit $\exp(x) = |w|$. Es folgt dass $|w\exp(-x)| = 1$. Kann man die Zahl $w\exp(-x)$ schreiben in der Form $\exp(z)$, so kann man auch w in dieser Form schreiben.

Es genügt also zu zeigen, dass es zu jedem w mit $|w| = 1$ eine reelle Zahl t mit $w = \exp(it)$ gibt. Sei also $w \in \mathbb{C}$, $|w| = 1$ gegeben. Schreibe $w = u + iv$, $u, v \in \mathbb{R}$. Man betrachte zunächst den Fall $u \geq 0$, $v \geq 0$.

Da sicher $u \leq 1$ ist, gibt es (Zwischenwertsatz) eine Zahl $t, 0 \leq t \leq \frac{\pi}{2}$, mit $\operatorname{Re}\exp(it) = u$. Dann folgt, wegen $|\exp(it)| = 1$,

$$(\operatorname{Im}\exp(it))^2 = 1 - (\operatorname{Re}\exp(it))^2 = 1 - u^2 = v^2\,,$$

und da $\operatorname{Im}\exp(it) \geq 0$ ist, auch $\operatorname{Im}\exp(it) = v$. Insgesamt gilt $e^{it} = w$.

Ist $u < 0, v \geq 0$, so betrachte $-iw$ statt w. Dann ist $-iw = \exp(it)$ für ein geeignetes t und daher $w = \exp(i(t + \frac{\pi}{2}))$. Ist $v < 0$, so betrachte $-w$ statt w. Dann ist für ein gewisses t auch $-w = \exp(it)$ und damit $w = \exp(i(t + \pi))$.

✌

Als eine erste Folgerung dieses Satzes erhalten wir die Existenz von Wurzeln in $\mathbb{C}$.

2.5.2 Korollar. *Sei $w \in \mathbb{C}\backslash\{0\}$ und $n \in \mathbb{N}$. Dann existiert eine Zahl $z_0 \in \mathbb{C}\backslash\{0\}$ mit $z_0^n = w$. Die Gesamtheit aller n-ten Wurzeln von w ist gegeben als*

$$\{z \in \mathbb{C} : z^n = w\} = \{z_0 \exp\big[\frac{2\pi i k}{n}\big] : k = 0, \ldots, n-1\}.$$

Beweis. Es ist $w = \exp(w_0)$ mit einer gewissen Zahl $w_0 \in \mathbb{C}$. Setzt man $z_0 := \exp(\frac{w_0}{n})$, so gilt $z_0^n = w$.

Sei $z \in \mathbb{C}$ mit $z^n = w$. Dann ist sicher $z \neq 0$. Also können wir z anschreiben als $z = \exp(v)$. Nun gilt

$$\exp(nv) = z^n = w = \exp(w_0),$$

und daher $nv - w_0 = 2\pi i k$ mit einem $k \in \mathbb{Z}$. Es folgt $v = \frac{w_0}{n} + \frac{2\pi i k}{n}$, und daher

$$z = z_0 \exp\big(\frac{2\pi i k}{n}\big).$$

Ist $k - k' \in n\mathbb{Z}$, so gilt $\exp\big(\frac{2\pi i k}{n}\big) = \exp\big(\frac{2\pi i k'}{n}\big)$. Also folgt die Inklusion '$\subseteq$'. Die umgekehrte Inklusion ist klar.

❑

2.6 'Elementare' Funktionen

Wir wollen jetzt einige der sogenannten elementaren Funktionen betrachten.

Wir definieren für $z \in \mathbb{C}$ die beiden *trigonometrischen Funktionen* $\cos z$ und $\sin z$ durch

$$\begin{aligned} \cos z &:= \frac{\exp(iz) + \exp(-iz)}{2} = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}, \\ \sin z &:= \frac{\exp(iz) - \exp(-iz)}{2i} = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!}. \end{aligned} \tag{2.7}$$

Cosinus und Sinus sind also auf ganz $\mathbb{C}$ definierte und stetige Funktionen.

2.6.1 Korollar.

(i) Es gilt die Eulersche Identität*:*

$$\exp(iz) = \cos z + i \sin z, \; z \in \mathbb{C},$$

sowie die Formel von de Moivre*:*

$$(\cos z + i \sin z)^n = \cos(nz) + i \sin(nz), \; n \in \mathbb{N}, z \in \mathbb{C}.$$

(ii) Für $t \in \mathbb{R}$ *gilt* $\cos t = \operatorname{Re}\exp(it)$ *und* $\sin t = \operatorname{Im}\exp(it)$*. Es gilt*

$$(\cos t)^2 + (\sin t)^2 = 1, \ t \in \mathbb{R}.$$

(iii) Es sind $\cos z$ *und* $\sin z$ periodisch *mit Periode* 2π*, d.h. es gilt*

$$\cos(z + 2\pi k) = \cos z, \ \sin(z + 2\pi k) = \sin z, \ k \in \mathbb{Z}, z \in \mathbb{C}.$$

Es ist $\cos z = 0$ *genau dann, wenn* $z = \frac{\pi}{2} + \pi k$*,* $k \in \mathbb{Z}$*, und es ist* $\sin z = 0$ *genau dann, wenn* $z = \pi k$*,* $k \in \mathbb{Z}$*.*

Beweis. Die Eulersche Identität folgt unmittelbar aus der Definition (2.7) von Cosinus und Sinus. Die Formel von de Moivre folgt nun aus der Funktionalgleichung der Exponentialfunktion, denn auf der linken Seite steht $\exp(z)^n$ und auf der rechten $\exp(nz)$.

Ist $t \in \mathbb{R}$, so sind, wie man aus der Reihendarstellung in (2.7) sieht, $\cos t$ und $\sin t$ reell. Daher sind sie Real- und Imaginärteil von $\cos t + i \sin t$ und diese Zahl ist nach der Eulerschen Identität gleich $\exp(it)$. Es ist $1 = |\exp(it)|^2 = \operatorname{Re}\exp(it)^2 + \operatorname{Im}\exp(it)^2$.

Wir benützen die Periodizität der Exponentialfunktion:

$$\cos(z + 2\pi k) = \frac{\exp(i(z + 2\pi k)) + \exp(-i(z + 2\pi k))}{2} =$$

$$= \frac{\exp(iz + 2\pi ki)) + \exp(-iz - 2\pi ki))}{2} = \frac{\exp(iz) + \exp(-iz)}{2} = \cos z.$$

Genauso für $\sin z$.

Ist $\sin z = 0$ genau dann, wenn $\exp(iz) = \exp(-iz)$, also wenn $\exp(2iz) = 1$. Dieses ist genau dann der Fall, wenn $\frac{2iz}{2\pi i} \in \mathbb{Z}$, d.h. wenn $z \in \pi\mathbb{Z}$. Analog bestimmt man die Nullstellen des Cosinus: Es ist $\cos z = 0$ genau dann, wenn $\exp(iz) = -\exp(-iz) = \exp(i\pi - iz)$, also wenn $\exp(2iz - i\pi) = 1$. Dieses tritt genau dann ein, wenn $\frac{2iz - i\pi}{2\pi i} \in \mathbb{Z}$, d.h. wenn $z \in \frac{\pi}{2} + \pi\mathbb{Z}$.

❑

2.6.2 Korollar. *Es gelten die* Summensätze *für Sinus und Cosinus:*

$$\cos(x + y) = \cos x \cos y - \sin x \sin y, \ \sin(x + y) = \sin x \cos y + \cos x \sin y.$$

Beweis. Nach der Definition (2.7) gilt

$$\cos x \cos y = \frac{\exp(ix) + \exp(-ix)}{2} \frac{\exp(iy) + \exp(-iy)}{2} =$$

$$= \frac{\exp(ix)\exp(iy) + \exp(-ix)\exp(-iy) + \exp(-ix)\exp(iy) + \exp(ix)\exp(-iy)}{4}.$$

$$\sin x \sin y = \frac{\exp(ix) - \exp(-ix)}{2i} \frac{\exp(iy) - \exp(-iy)}{2i} =$$

$$= \frac{\exp(ix)\exp(iy) + \exp(-ix)\exp(-iy) - \exp(-ix)\exp(iy) - \exp(ix)\exp(-iy)}{-4}.$$

Also folgt die erste Formel. Die zweite Formel sieht man analog.

❑

Wie wir in Satz 2.5.1 festgestellt haben, ist die Funktion $\exp(x) : \mathbb{R} \to \mathbb{R}^+$ bijektiv. Also existiert ihre Inverse, eine Funktion von $\mathbb{R}^+$ auf $\mathbb{R}$. Diese bezeichnen wir mit ln und sprechen vom *natürlichen Logarithmus*.

2.6.3 Korollar. *Die Funktion* $\ln : \mathbb{R}^+ \to \mathbb{R}$ *ist eine stetige und streng monoton wachsende Bijektion. Es gilt*

$$\lim_{x\to 0+} \ln x = -\infty, \ \lim_{x\to\infty} \ln x = +\infty\,,$$

sowie die Funktionalgleichung

$$\ln(xy) = \ln x + \ln y, \ x, y > 0\,.$$

Beweis. Da $\exp(x)$ stetig, streng monoton wachsend und bijektiv ist, hat auch die Umkehrfunktion diese Eigenschaften, vgl. Korollar 2.2.7. Die Limesbeziehungen gelten wegen der entsprechenden Formeln Satz 2.5.1, (*iii*). Die Funktionalgleichung folgt wegen der Funktionalgleichung von exp.

❑

Beachte, dass wir den Logarithmus nur für reelle Werte definiert haben. Dies ist kein Zufall, will man den Logarithmus auch für komplexe Werte definieren, trifft man auf Schwierigkeiten ganz essentieller Natur (vgl. Vorlesung zur Komplexen Analysis).

Sind $a, b \in \mathbb{R}$, $a > 0$, so definieren die *allgemeine Potenz* $\mathrm{pot}(a, b)$ als

$$\mathrm{pot}(a, b) := \exp(b \ln a)$$

Das zweite Argument b nennt man manchmal auch den *Exponenten* der Potenz. Die Rechtfertigung für die Namensgebung „allgemeine Potenz" entnehmen wir der folgenden Aussage.

2.6.4 Korollar. *Die allgemeine Potenz* $\mathrm{pot} : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^+$ *ist stetig. Sie erfüllt die Funktionalgleichungen*

$$\mathrm{pot}(a, b_1 + b_2) = \mathrm{pot}(a, b_1)\,\mathrm{pot}(a, b_2), \ \mathrm{pot}(a_1 a_2, b) = \mathrm{pot}(a_1, b)\,\mathrm{pot}(a_2, b)\,.$$

Für $b \in \mathbb{Q}$, $b = \frac{p}{q}$, *gilt*

$$\mathrm{pot}(a, b) = \sqrt[q]{a^p}\,.$$

Beweis. Als Zusammensetzung der stetigen Funktionen $\exp, \ln$ und $\cdot$, ist pot stetig. Die Funktionalgleichungen folgen aus denen von exp und ln, z.B. ist

$$\mathrm{pot}(a, b_1 + b_2) = \exp((b_1 + b_2) \ln a) = \exp(b_1 \ln a + b_2 \ln a) =$$

$$= \exp(b_1 \ln a) \exp(b_2 \ln a) = \mathrm{pot}(a, b_1)\,\mathrm{pot}(a, b_2)\,.$$

Ist $b = \frac{p}{q} \in \mathbb{Q}$, so ist $\mathrm{pot}(a, b)$ eine positive reelle Zahl mit der Eigenschaft dass

$$\mathrm{pot}(a, b)^q = \underbrace{\mathrm{pot}(a, b) \cdot \ldots \cdot \mathrm{pot}(a, b)}_{q \text{ mal}} = \mathrm{pot}(a, qb) = \mathrm{pot}(a, p) =$$

$$= \mathrm{pot}(a, \underbrace{1 + \ldots + 1}_{p \text{ mal}}) = \underbrace{\mathrm{pot}(a, 1) \cdot \ldots \cdot \mathrm{pot}(a, 1)}_{p \text{ mal}} = \big(\exp(1 \cdot \ln a)\big)^p = a^p\,.$$

Also ist $\operatorname{pot}(a,b) = \sqrt[q]{a^p}$.

❑

Wir können also, ohne dass dies zu einer Bezeichnungskollision führt, die suggestivere Schreibweise

$$\operatorname{pot}(a,b) =: a^b,\ a,b \in \mathbb{R}, a > 0\,,$$

verwenden.

Die Funktion exp kann nun selbst mit dieser Notation als allgemeine Potenz angeschrieben werden. Sei dazu $e := \exp(1)$, die *Eulersche Zahl*. Dann gilt

$$\exp(x) = e^x,\ x \in \mathbb{R}\,.$$

Denn: Es ist nach der Definition der allgemeinen Potenz

$$e^x = \operatorname{pot}(e,x) = \exp(x \ln e) = \exp\big(x \underbrace{\ln\exp(1)}_{=1}\big) = \exp(x)\,.$$

2.6.5 Bemerkung. Wir haben schon festgestellt, und auch demonstriert, dass die Eulersche Exponentialfunktion eine ganz zentrale Rolle spielt. Daher wird auch die reelle Zahl e ein interessantes Objekt sein. Dazu wollen wir hier bemerken, dass man die Eulersche Zahl e, neben ihrer Definition als $e := \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!}$ auch in vielfacher Weise anders charakterisieren kann. Zum Beispiel kann man zeigen, dass

$$e = \lim_{n\to\infty} \big(1 + \frac{1}{n}\big)^n\,,$$

ein oft gebräuchlicher Zugang.

2.7 Der Fundamentalsatz der Algebra

Wir sind nun in der Lage den Fundamentalsatz der Algebra, der besagt dass jedes Polynom mit komplexen Koeffizienten in ein Produkt von Linearfaktoren zerfällt, zu beweisen. Interessanterweise ist der Beweis dieser, eigentlich völlig „algebraischen", Aussage nicht ohne Zuhilfenahme von „topologischen" Mitteln möglich.

2.7.1 Satz (Fundamentalsatz der Algebra)**.** *Sei $p \in \mathbb{C}[x]$, $\operatorname{grad} p = n > 0$, mit Führungskoeffizienten a_n. Dann existieren n (nicht notwendig verschiedene) Zahlen $z_1, \dots, z_n \in \mathbb{C}$ sodaß*

$$p(x) = a_n \prod_{k=1}^{n} (x - z_k)\,.$$

Wir benötigen das folgende Lemma, wo die Stetigkeit von Polynomen eine ganz wesentliche Rolle spielt.

2.7.2 Lemma. *Es gilt:*

(i) Sei $f \in \mathbb{C}[x]$ nicht konstant, und sei $C > 0$. Dann existiert $R > 0$ sodass $|f(x)| \geq C$ für alle $x \in \mathbb{C}$ mit $|x| \geq R$.

(ii) *Ist $f \in \mathbb{C}[x]$, so gibt es eine Zahl $c \in \mathbb{C}$ mit $|f(c)| \leq |f(z)|$ für alle $z \in \mathbb{C}$.*

(iii) *Ist $g \in \mathbb{C}[x]$ mit $g(0) = 0$, so gibt es $r > 0$ sodaß $|g(z)| \leq \frac{1}{2}$ für alle $z \in \mathbb{C}$ mit $|z| < r$.*

Beweis. Sei $f \in \mathbb{C}[x]$ gegeben, f nicht konstant, und schreibe $f(x) = a_n x^n + \ldots + a_0$ mit $a_n \neq 0$. Setze

$$R := \max\left\{1, \frac{2(|a_{n-1}| + \ldots + |a_0|)}{|a_n|}, \frac{2C}{|a_n|}\right\},$$

dann gilt für $|x| \geq R$

$$|f(x)| = |x^n| \cdot \left|a_n + a_{n-1}\frac{1}{x} + \ldots + a_0\frac{1}{x^n}\right| \geq |x^n| \cdot \left(|a_n| - |a_{n-1}|\frac{1}{|x|} - \ldots - |a_0|\frac{1}{|x|^n}\right) \geq$$

$$\geq |x| \cdot \left(|a_n| - \frac{1}{|x|}(|a_{n-1}| + \ldots + |a_0|)\right) \geq |x|\frac{|a_n|}{2} \geq C$$

Um die zweite Behauptung einzusehen sei wieder $f(x) = a_n x^n + \ldots + a_0 \in \mathbb{C}[x]$ mit $a_n \neq 0$ (ist $f = 0$ ist die Behauptung trivialerweise richtig). Wegen (*i*) existiert insbesondere eine Zahl $R > 0$ sodaß $|f(x)| \geq |a_0| = |f(0)|$, $x \in \mathbb{C}$ mit $|x| > R$. Nun ist die Kreisscheibe $K := \{x \in \mathbb{C} : |x| \leq \mathbb{R}\}$ kompakt und $|f(x)|$ ist, als Zusammensetzung der stetigen Funktionen f und $|.|$, stetig auf K. Daher wird ein Minimum angenommen, $\min_{x \in K} |f(x)| = |f(c)|$. Unsere Wahl von R sichert, dass $|f(c)| = \min_{x \in \mathbb{C}} |f(x)|$.

Die dritte Aussage ist einfach die Tatsache, dass das Polynom g stetig an der Stelle 0 ist (mit der Wahl $\epsilon = \frac{1}{2}$).

❑

Beweis. (von Satz 2.7.1)

Schritt 1: Wir zeigen zunächst die folgende Behauptung: Sei $k \in \mathbb{N}$, $h(z) = 1 + bz^k + z^k g(z)$, mit $b \in \mathbb{C} \setminus \{0\}$, g Polynom, $g(0) = 0$. Dann existiert ein $u \in \mathbb{C}$ mit $|h(u)| < 1$.

Denn: Wir wählen eine k-te Wurzel aus $-\frac{1}{b}$, d.h. eine Zahl $d \in \mathbb{C}$ mit $bd^k = -1$. Für $t \in (0, 1]$ gilt dann

$$|h(td)| \leq |1 - t^k| + |t^k d^k g(td)| = 1 - t^k + t^k |d^k g(td)|.$$

Wegen Lemma 2.7.2 gibt es $t \in (0, 1]$ mit $|d^k g(td)| \leq \frac{1}{2}$. Dann gilt $|h(td)| \leq 1 - \frac{1}{2}t_k < 1$.

Schritt 2: Als nächstes zeigen wir: Jedes nichtkonstante Polynom $f \in \mathbb{C}[x]$ hat eine Nullstelle in $\mathbb{C}$.

Denn: Sei c so daß $|f(c)| = \min_{z \in \mathbb{C}} |f(z)|$. Angenommen $f(c) \neq 0$. Betrachte

$$h(z) := \frac{f(z + c)}{f(c)} = 1 + b_k z^k + b_{k+1} z^{k+1} + \ldots + b_n z^n, \ b_k \neq 0.$$

Nach dem ersten Beweisschritt existiert $u \in \mathbb{C}$ mit $|h(u)| < 1$, also mit

$$|f(u + c)| = |h(u)| \cdot |f(c)| < |f(c)|,$$

ein Widerspruch.

Schritt 3: Wir kommen schließlich zum Beweis der behaupteten Faktorisierung. Dazu verwenden wir Induktion nach $\operatorname{grad} p > 0$: Ist $\operatorname{grad} p = 1$, also $p(x) = a_1 x + a_0$ mit $a_1 \neq 0$, so ist $p(x) = a_1(x - \frac{a_0}{a_1})$. Sei nun angenommen das der Satz für alle Polynome vom Grad kleiner als n bereits bewiesen ist, und sei $\operatorname{grad} p = n$. Dann hat nach dem vorigen Beweisschritt p eine Nullstelle z_1. Schreibe mittels Polynomdivision $p(x) = s(x)(x - z_1)$. Das Polynom s hat den gleichen Führungskoeffizienten wie p und lässt sich nach Induktionsvoraussetzung in der angegebenen Weise fatorisieren.

❑

Wir wissen ja, dass ein Polynom mit Koeffizienten in einem Körper K höchstens soviele Nullstellen haben kann wie sein Grad angibt. Der Fundamentalsatz der Algebra besagt nun, salopp gesprochen, dass ein Polynom mit komplexen Koeffizienten in $\mathbb{C}$ tatsächlich genau soviele Nullstellen hat wie sein Grad angibt. Dazu muss man jedoch in der Faktorisierung mehrfach auftretende Nullstellen z_k auch entsprechend oft zählen. Betrachten wir zum Beispiel das Polynom $p(x) = x^4 + x^2$, so gilt $p(x) = (x - 0)(x - 0)(x - i)(x + i)$, es hat also $+i$ und $-i$ als einfache Nullstellen und 0 als zweifache Nullstelle.

Kapitel 3

Differentialrechnung

3.1 Definition der Ableitung

3.1.1 Beispiel. Wir betrachten den freien Fall eines Massepunktes (im Vakuum) im Gravitationsfeld der Erde. Nimmt man der Einfachheit halber an, daß die wirkende Kraft auf der Fallstrecke konstant ist, so berechnet sich der Weg den der Massepunkt nach t Sekunden zurückgelegt hat als (g = 9.81 ms^{-2})

$$s = \frac{g}{2}\, t^2$$

Wir wollen die Momentangeschwindigkeit des Punktes zum Zeitpunkt t berechnen. Wäre die Geschwindigkeit des Punktes konstant, so würde sie sich als Verhältnis von zurückgelegtem Weg und dafür benötigter Zeit berechnen lassen.

Beobachten wir den Punkt im Zeitintervall von t bis $t + \Delta t$, so legt er die Strecke

$$\frac{g}{2}\,(t + \Delta t)^2 - \frac{g}{2}\,t^2 = \frac{g}{2}\,(2t \cdot \Delta t + \Delta t^2)$$

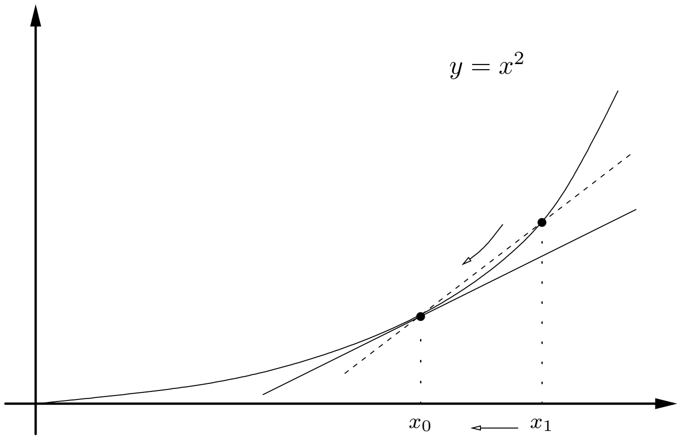
zurück. Ist hierbei Δt sehr klein, so wird der Punkt mit annähernd konstanter Geschwindigkeit unterwegs sein. Diese mittlere Geschwindigkeit im Zeitintervall von t bis $t + \Delta t$ ist also gleich

$$v_{t,t+\Delta t} = \frac{g}{2}\,(2t \cdot \Delta t + \Delta t^2) \cdot \frac{1}{\Delta t} = \frac{g}{2}\,2t + \frac{g}{2}\,\Delta t.$$

Für sehr kleine Zeitintervalle, d.h. sehr kleine Werte von Δt, wird sich nun diese mittlere Geschwindigkeit kaum von der Momentangeschwindigkeit zum Zeitpunkt t unterscheiden. Wir schließen für diese also

$$v_t = \lim_{\Delta t \to 0} v_{t,t+\Delta t} = gt.$$

3.1.2 Beispiel. Wir wollen die Gleichung der Tangente an eine Kurve in einem gewissen Punkt bestimmen. Zum Beispiel betrachte die Parabel $y = x^2$ und den auf ihr liegenden Punkt (x_0, x_0^2).



Zur Konstruktion der Tangente gehen wir folgendermaßen vor: Wir wählen einen zweiten Kurvenpunkt (x_1, x_1^2) und bestimmen die Sekante durch diese beiden Punkte. Deren Gleichung läßt sich zum Beispiel anschreiben als

$$y = x_0^2 + (x - x_0)\frac{x_1^2 - x_0^2}{x_1 - x_0} = x_0^2 + (x - x_0)(x_1 + x_0).$$

Wählt man nun x_1 sehr nahe bei x_0, so wird sich diese Sekante kaum mehr von der Tangente unterscheiden. Man erhält die Gleichung der Tangente durch den Grenzübergang $x_1 \to x_0$

$$y = x_0^2 + (x - x_0)\ (x_0 + x_0) = 2x_0 \cdot x - x_0^2.$$

3.1.3 Beispiel. Wir betrachten wieder die Parabel und suchen die Tangente im Punkt (x_0, x_0^2). Diesmal gehen wir jedoch von der Vorstellung aus, daß die Tangente jene Gerade ist welche die Kurve im Punkt x_0 "am besten" approximiert. Sei dazu eine beliebige Gerade angeschrieben als

$$y = \alpha x + \beta.$$

Soll die Gerade die Kurve im Punkt (x_0, x_0^2) approximieren, so wird zumindestens der Punkt (x_0, x_0^2) selbst auf ihr liegen müssen. D.h.

$$x_0^2 = \alpha x_0 + \beta,$$

oder $\beta = x_0^2 - \alpha x_0$. Die Gleichung der Geraden erhält also die Gestalt $y = x_0^2 + \alpha(x - x_0)$. Für die Differenz der y-Koordinaten von Kurve und Gerade erhält man also

$$x^2 - \ (x_0^2 + \alpha(x - x_0)) = (x - x_0)\ (x + x_0 - \alpha).$$

Die Tatsache, daß der Punkt (x_0, x_0^2) selbst auf beiden liegt, drückt sich darin aus, daß dieser Wert für $x = x_0$ gleich Null ist, da ja die erste Klammer $= 0$ ist. Wählt man für α irgend eine Zahl $\alpha \neq 2x_0$, so wird die zweite Klammer bei x_0 nicht verschwinden, d.h. der Abstand verhält sich für $x \to x_0$ so wie $x - x_0$. Wählt man jedoch $\alpha = 2x_0$, so wird im Punkt x_0 auch die zweite Klammer verschwinden, diese Gerade approximiert die Kurve nahe bei x_0 also noch besser.

Man sieht, daß wir abgesehen von den Bezeichnungsweisen eigentlich immer das selbe getan haben. Diese Gemeinsamkeit drückt sich in der folgenden Definition aus:

3.1.4 Definition. Sei f eine reelle Funktion definiert auf einem Intervall (a, b) und sei $x \in (a, b)$. Dann heißt f *differenzierbar* im Punkt x, falls der Grenzwert

$$\alpha = \lim_{t \to x} \frac{f(t) - f(x)}{t - x}$$

existiert. Der Wert α dieses Grenzwertes heißt im Falle der Existenz die *Ableitung* von f an der Stelle x. Man schreibt auch $\alpha = f'(x)$, oder $\alpha = \frac{df}{dt}(x)$.

3.1.5 Beispiel.

(i) Die konstante Funktion $f(t) = \lambda$ ist an jeder Stelle x differenzierbar. Ihre Ableitung im Punkt x ist gleich 0.

(ii) Die Funktion $f(t) = t$ ist in jedem Punkt x differenzierbar. Ihre Ableitung an der Stelle x ist gleich 1.

(iii) Die Funktion $f(t) = t^2$ ist in jedem Punkt x differenzierbar. Wie wir oben gesehen haben ist ihre Ableitung an der Stelle x gleich $2x$.

(iv) Die Funktion

$$f(t) = \begin{cases} t \sin \frac{1}{t}, & t \neq 0 \\ 0, & t = 0 \end{cases} \tag{3.1}$$

ist im Punkt $x = 0$ nicht differenzierbar. Denn es gilt

$$\frac{f(t) - f(0)}{t - 0} = \frac{t \sin \frac{1}{t} - 0}{t - 0} = \sin \frac{1}{t}.$$

(v) Betrachte die Funktion f definiert für $t \in \mathbb{R}$ durch

$$f : \begin{cases} t^2 \sin \frac{1}{t} \ , & t \neq 0 \\ 0 \ , & t = 0 \end{cases}$$

Die Funktion f ist an der Stelle $x = 0$ differenzierbar, denn es gilt

$$\frac{f(t) - f(0)}{t - 0} = t \sin \frac{1}{t} \to 0 \text{ für } t \to 0.$$

Es ist also $f'(0) = 0$. An einer Stelle $x \neq 0$ ist f (setzt man die Eigenschaften der Sinusfunktion als bekannt voraus) differenzierbar und es gilt (vgl. später)

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}.$$

(vi) Betrachte die Funktion $f(t) = \sqrt{t}, t > 0$. Sei $x > 0$, dann gilt

$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x} = \lim_{t \to x} \frac{\sqrt{t} - \sqrt{x}}{(\sqrt{t} - \sqrt{x})(\sqrt{t} + \sqrt{x})} = \lim_{t \to x} \frac{1}{\sqrt{t} + \sqrt{x}} = \frac{1}{2\sqrt{x}}.$$

D.h. die Funktion f ist an jeder Stelle $x > 0$ differenzierbar und es gilt

$$f'(x) = \frac{1}{2\sqrt{x}}.$$

(vii) Die Funktion $f(t) := |t|$ ist an jeder Stelle $x \neq 0$ differenzierbar, denn es ist

$$\frac{|t| - |x|}{t - x} = \begin{cases} 1 & , \ t, x > 0 \\ -1 & , \ t, x < 0 \end{cases}$$

Also gilt

$$\lim_{t \to x} \frac{|t| - |x|}{t - x} = \begin{cases} 1 & , \ x > 0 \\ -1 & , \ x < 0 \end{cases}$$

An der Stelle $x = 0$ ist sie jedoch nicht differenzierbar, denn es ist

$$\frac{|t|}{t} = \begin{cases} 1 & , \ t > 0 \\ -1 & , \ t < 0 \end{cases}$$

also existiert der Limes $\lim_{t \to x} \frac{f(t) - f(0)}{t - 0}$ nicht. Beachte, dass f an der Stelle 0 trotzdem stetig ist.

3.1.6 Bemerkung. Die Funktion f ist im Punkt x differenzierbar genau dann, wenn sich f schreiben läßt als

$$f(t) = f(x) + (t - x)\alpha + (t - x)\mu(t),$$

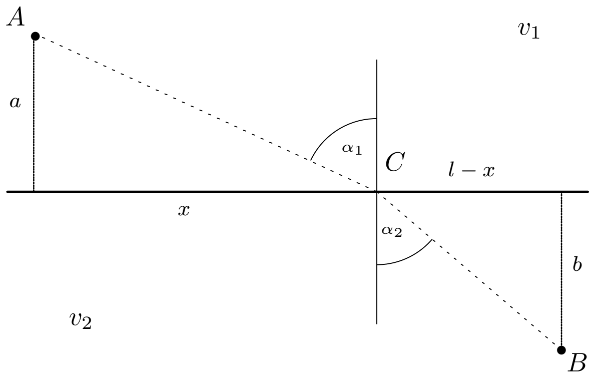
wobei α eine Konstante ist und $\mu(t)$ eine Funktion mit $\lim_{t \to x} \mu(t) = 0$, d.h. μ hat eine stetige Fortsetzung in den Punkt x mit $\mu(x) = 0$. In diesem Fall ist α die Ableitung von f an der Stelle x.

3.1.7 Bemerkung. Die Funktion f ist im Punkt x differenzierbar genau dann, wenn der Differenzenquotient

$$\phi(t) := \frac{f(t) - f(x)}{t - x}$$

eine in x stetige Fortsetzung hat. In diesem Fall gilt $\phi(x) = f'(x)$.

3.1.8 Beispiel. Betrachte das Problem der Brechung eines Lichtstrahles an der Wasseroberfläche:



Im oberen Medium (Luft) bewegt sich das Licht mit einer Geschwindigkeit von v_1, im unteren (Wasser) mit v_2. Das Brechungsgesetz von Snellius besagt, daß sich die Winkel α_1, α_2 bestimmen lassen durch

$$\frac{\sin \alpha_1}{v_1} = \frac{\sin \alpha_2}{v_2}.$$

Wieso?

Um dieses Gesetz zu begründen gehen wir davon aus, daß das Licht den Weg von A nach B nimmt den es am schnellsten zurücklegt. Schreibt man die Zeit die das Licht von A nach B längs eines Weges durch den Punkt C benötigt in Abhängigkeit von den Winkeln α_1, α_2, bzw. äquivalent in Abhängigkeit von x

$$T = T(x),$$

so sucht man also jenen Wert von x für den $T(x)$ minimal ist. Ein solcher Wert von x muß existieren, eindeutig sein, und zwischen 0 und l liegen, denn liegt z.B. am Grund ein Seeigel, so sieht man ja auch einen (und nur einen). Weiters ist $x = 0$ bzw. $x = l$ sicher ausgeschlossen, denn schickt man den Lichtstrahl horizontal bzw. vertikal von A weg, so wird er erfahrungsgemäß nicht B erreichen.

Läßt man x von 0 nach l wandern, so wird $T(x)$ also zunächst (bis zur richtigen Lage von C) abnehmen und danach wieder zunehmen. Die Steigung der Kurve $T(x)$ ist links von der gesuchten Lösung also < 0, rechts davon > 0, im gesuchten Punkt also $= 0$.

Berechnet man $T(x)$, so erhält man

$$T(x) = \frac{\sqrt{a^2 + x^2}}{v_1} + \frac{\sqrt{b^2 + (l - x^2)}}{v_2}.$$

Die Ableitung ist (vgl. später):

$$T'(x) = \frac{1}{v_1} \frac{2x}{2\sqrt{a^2 + x^2}} + \frac{1}{v_2} \frac{-2(l - x)}{2\sqrt{b^2 + (l - x)^2}},$$

und ist offenbar gleich 0 wenn

$$\left(\frac{x}{\sqrt{a^2 + x^2}} \right) : v_1 = \frac{l - x}{\sqrt{b^2 + (l - x)^2}} : v_2.$$

Nun ist $\frac{x}{\sqrt{a^2+x^2}} = \sin \alpha_1$, $\frac{l-x}{\sqrt{b^2+(l-x)^2}} = \sin \alpha_2$.

Man beachte, daß wir in obigem Beispiel viele essentielle Schlüsse aus unserer Anschauung gezogen haben. Z.B. die Existenz eines Minimums, seine Eindeutigkeit, die Stetigkeit von T, die Differenzierbarkeit von T o.ä. .

3.2 Eigenschaften der Ableitung

3.2.1 Lemma. *Sei f differenzierbar im Punkt x. Dann ist f stetig bei x.*

Beweis. Sei $\lim_{t \to x} \frac{f(t)-f(x)}{t-x} = \alpha$. Dann folgt

$$\lim_{t \to x} [f(t) - f(x)] = \lim_{t \to x} \left[\frac{f(t) - f(x)}{t - x} (t - x) \right] = \alpha \cdot 0 = 0.$$

❑

3.2.2 Bemerkung. Wie man am Beispiel der Funktion f aus (3.1) sieht, gilt die Umkehrung von Lemma 3.2.1 nicht. Siehe auch Beispiel 3.4.9.

3.2.3 Satz. *Seien f, g definiert auf (a, b) und differenzierbar im Punkt $x \in (a, b)$, und sei $\lambda \in \mathbb{R}$. Dann ist auch $f + g, \lambda f, fg$ und (falls $g(x) \neq 0$) auch $\frac{f}{g}$ an der Stelle x differenzierbar und es gilt*

(i) $(f + g)'(x) = f'(x) + g'(x)$, *und auch* $(\lambda f)'(x) = \lambda f'(x)$.

(ii) Produktregel*:* $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$

(iii) Quotientenregel*:* $(\frac{f}{g})'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$.

Beweis. Es gilt

$$\lim_{t \to x} \frac{(f + g)(t) - (f + g)(x)}{t - x} = \lim_{t \to x} \frac{f(t) - f(x)}{t - x} + \lim_{t \to x} \frac{g(t) - g(x)}{t - x},$$

also folgt die erste Aussage in (i). Die zweite folgt genauso, oder mit Hilfe der Produktregel (ii). Setzt man $h = fg$, so ist

$$\frac{h(t) - h(x)}{t - x} = \frac{f(t)g(t) - f(x)g(x)}{t - x} = f(t) \frac{g(t) - g(x)}{t - x} + \frac{f(t) - f(x)}{t - x} g(x).$$

Da f bei x stetig ist folgt (ii). Um (iii) einzusehen setze $h = \frac{f}{g}$ und berechne

$$\frac{h(t) - h(x)}{t - x} = \frac{\frac{f(t)}{g(t)} - \frac{f(x)}{g(x)}}{t - x} =$$

$$\frac{1}{g(t)g(x)} \left[g(x) \frac{f(t) - f(x)}{t - x} - f(x) \frac{g(t) - g(x)}{t - x} \right].$$

Die Behauptung folgt da g bei x stetig ist.

❑

3.2.4 Beispiel. Durch wiederholte Anwendung dieser Rechenregeln folgt, daß $f(t) = t^n$, $n \in \mathbb{Z}$, $n \neq 0$, differenzierbar ist und

$$f'(x) = nx^{n-1}$$

gilt. Weiters ist eine rationale Funktion in jedem Punkt wo der Nenner nicht verschwindet differenzierbar.

Durch wiederholte Anwendung der Produktregel erhält man die Formel

$$(f_1 \cdot \ldots \cdot f_n)' = f_1' f_2 \cdots f_n + f_1 f_2' f_3 \cdots f_n + \ldots + f_1 \cdots f_{n-1} f_n'.$$

3.2.5 Satz (Kettenregel)**.** *Seien f, g reelle Funktionen,* $\operatorname{dom} f = (a, b)$, $\operatorname{dom} g = (c, d) \supseteq \operatorname{ran} f$. *Sei $x \in (a, b), f$ differenzierbar an der Stelle x, und g differenzierbar an der Stelle $f(x)$. Dann ist $g \circ f$ differenzierbar an der Stelle x und es gilt*

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

Beweis. Schreibe

$$f(t) = f(x) + (t-x)\left[f'(x) + \mu(t)\right], \ \lim_{t\to x} \mu(t) = 0,$$

$$g(u) = g(f(x)) + (u - f(x))\left[g'(f(x)) + \nu(u)\right], \ \lim_{u\to x} \nu(u) = 0.$$

Dann folgt

$$g(f(t)) = g(f(x)) + (f(t) - f(x))\left[g'(f(x)) + \nu(f(t))\right] =$$

$$= g(f(x)) + (t-x)\left[f'(x) + \mu(t)\right]\left[g'(f(x)) + \nu(f(t))\right] =$$

$$= g(f(x)) + (t-x)f'(x)g'(f(x)) +$$

$$+(t-x)\big[f'(x)\nu(f(t)) + \mu(t)g'(f(x)) + \mu(t)\nu(f(t))\big].$$

Der Ausdruck in der eckigen Klammer strebt für $t \to x$ gegen 0.

❑

Beweis. (#2 von Satz 3.2.5)

$$\frac{(g\circ f)(t) - (g\circ f)(x)}{t-x} = \frac{g(f(t)) - g(f(x))}{f(t) - f(x)} \cdot \frac{f(t) - f(x)}{t-x}.$$

Für $t \to x$ strebt der erste Faktor wegen $f(t) \to f(x)$ gegen $g'(f(x))$. Der zweite Faktor strebt gegen $f'(x)$.

❑

3.2.6 Beispiel. Sei $f(x) = x^2 \sin \frac{1}{x}$. Durch Anwendung der Kettenregel (unter Benützung von $[\sin x]' = \cos x$) ergibt sich

$$f'(x) = 2x \sin \frac{1}{x} + x^2 \cos \frac{1}{x}(-\frac{1}{x^2}).$$

3.2.7 Satz. *Sei $f : [a, b] \to [c, d]$ stetig und bijektiv. Bezeichne mit $g : [c, d] \to [a, b]$ die Umkehrfunktion von f. Ist f an einer Stelle x differenzierbar und gilt $f'(x) \neq 0$, so ist g an der Stelle $f(x)$ differenzierbar, und es gilt*

$$g'(f(x)) = \frac{1}{f'(x)}.$$

Beweis. Wir betrachten den Limes

$$\lim_{t\to f(x)} \frac{g(t) - g(f(x))}{t - f(x)}.$$

Da f stetig ist, ist auch die Inverse g stetig. Führen wir in obigem Grenzwert die neue Variable $u = g(t)$ ein, so erhält man

$$\lim_{t\to f(x)} \frac{g(t) - g(f(x))}{t - f(x)} = \lim_{u\to x} \frac{u - x}{f(u) - f(x)} = \frac{1}{f'(x)}.$$

❑

3.2.8 Bemerkung. Weiss man in der Situation des obigen Satzes schon, daß g an der Stelle $f(x)$ differenzierbar ist, so folgt die Regel zur Berechnung der Ableitung auch aus der Kettenregel:

$$1 = [t]'(f(x)) = [f(g(t))]'(f(x)) = f'(g(f(x))) \cdot g'(f(x)) = f'(x) \cdot g'(f(x)).$$

3.2.9 Beispiel. Betrachte die Funktion $f(t) = t^2 : \mathbb{R}^+ \to \mathbb{R}^+$. Diese ist stetig und bijektiv, ihre Umkehrfunktion ist $g(u) = \sqrt{u} : \mathbb{R}^+ \to \mathbb{R}^+$. Dann gilt also $(y = f(x) = x^2)$

$$[\sqrt{u}]'(y) = [\sqrt{y}]'(f(x)) = \frac{1}{f'(x)} = \frac{1}{2x} = \frac{1}{2\sqrt{y}}.$$

3.2.10 Definition. Sei f eine reelle Funktion definiert auf (a,b). Wir definieren eine Funktion f' durch

$$\operatorname{dom} f' := \{x \in (a,b) : f \text{ ist differenzierbar an der Stelle } x\}$$

$$f'(x) := \lim_{t \to x} \frac{f(t) - f(x)}{t - x}, \; x \in \operatorname{dom} f'.$$

f' heißt *Ableitung* von f.

D.h. einer Funktion f wird eine weitere Funktion zugeordnet, die die aus f abgeleitete Funktion f' genannt wird. Ihr Wert an einer Stelle x ist gerade der Limes des Differenzenquotienten von f bei x. Diese Sichtweise erklärt auch die Schreibweise "$\alpha = f'(x)$" aus Definition 3.1.4. Die Schreibweise $\frac{df}{dt}(x)$ erklärt sich aus der Interpretation der Ableitung als Grenzfall des Zuwachses von f dividiert durch den Zuwachs von t (vgl. Beispiel 3.1.1).

Es ist also sinnvoll von Eigenschaften der Funktion f', wie zum Beispiel Stetigkeit oder auch wieder Differenzierbarkeit zu sprechen. Wie wir in Beispiel 3.1.5, (v), gesehen haben, muß die Ableitung f' nicht notwendig stetig sein.

3.2.11 Definition. Sei f definiert und differenzierbar für jedes $x \in (a,b)$. Ist die Ableitung f' an einer Stelle x differenzierbar, so bezeichnet man $(f')'(x)$ mit $f''(x)$ und spricht von der zweiten Ableitung von f an der Stelle x. Allgemeiner definiert man *höhere Ableitungen* rekursiv durch

$$f^{(n)}(x) := (f^{(n-1)})'(x), \; n \in \mathbb{N},$$

wann immer die auftretenden Ableitungen existieren.

Existiert $f^{(n)}$ an der Stelle x, so sagt man f ist bei x *n-mal differenzierbar*. Ist $f^{(n)}$ sogar stetig, so spricht man von einer n-mal *stetig differenzierbaren* Funktion.

Aus der Produktregel erhält man mittels vollständiger Induktion die, oft nützliche, *Leibnizsche Formel*

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)},$$

Beweis in der Übung.

3.2.12 Beispiel.

(*i*) Sei $f(x) = x^3 - 2x$. Dann gilt

$$f'(x) = 3x^2 - 2,\ f''(x) = 6x,\ f'''(x) = 6,\ f''''(x) = 0,\ f^{(5)}(x) = 0, \dots$$

Man sieht genauso, daß jedes Polynom p beliebig oft differenzierbar ist und, wenn n der Grad von p ist, $p^{(n+1)}(x) = p^{(n+2)}(x) = \ldots = 0$ gilt.

(*ii*) Sei f die Funktion

$$f(x) = \begin{cases} x^2 & ,\ x \geq 0 \\ -x^2 & ,\ x \leq 0 \end{cases}$$

Die Ableitung von f ist $f'(x) = 2|x|$. Die zweite Ableitung existiert also an der Stelle $x = 0$ nicht.

3.3 Mittelwertsätze

3.3.1 Definition. Sei f definiert auf einer Menge E. Wir sagen, f hat ein *lokales Maximum* in einem Punkt $x \in E$, falls gilt

$$\exists\, \delta > 0 : f(x) \geq f(t) \text{ für } t \in E, |t - x| < \delta.$$

Analog definiert man ein *lokales Minimum*. Will man sich nicht festlegen ob x ein Minimum oder Maximum ist, so spricht man zusammenfassend von einem *lokalen Extremum*.

Beachte den Unterschied zum Begriff des Maximums. Das ist eine Stelle $x \in E$, sodaß für jedes $t \in E$ gilt $f(x) \geq f(t)$, also nicht nur lokal bei x. Man spricht dann von einem *absoluten Maximum*. Analog für *absolute Minima* bzw. zusammenfassend *absolute Extrema*. Natürlich ist ein absolutes Extremum stets auch ein lokales.

3.3.2 Satz. *Sei f definiert auf $[a, b]$. Hat f an der Stelle $x \in (a, b)$ ein lokales Extremum und ist f bei x differenzierbar, so gilt $f'(x) = 0$.*

Beweis. Betrachte zum Beispiel den Fall, daß x ein lokales Maximum ist, der Fall des Minimums wird analog behandelt.

Wähle $\delta > 0$ wie in der Definition des Maximums. Es gilt also $f(x) \geq f(t)$ für alle $|t - x| < \delta$. Ist $t > x$, so folgt

$$\frac{f(t) - f(x)}{t - x} \leq 0,$$

und daher ist $f'(x) = \lim_{t \to x+} \frac{f(t)-f(x)}{t-x} \leq 0$. Ist jedoch $t < x$, so ist

$$\frac{f(t) - f(x)}{t - x} \geq 0,$$

und damit folgt $f'(x) = \lim_{t \to x-} \frac{f(t)-f(x)}{t-x} \geq 0$.

❑

3.3.3 Satz (Mittelwertsatz)**.** *Sei f stetig auf $[a,b]$ und differenzierbar auf (a,b). Dann existiert ein Punkt $x \in (a,b)$ mit*

$$\frac{f(b)-f(a)}{b-a} = f'(x).$$

Beweis. Betrachte die Funktion ($t \in [a,b]$)

$$F(t) := f(t) - f(a) - \frac{f(b)-f(a)}{b-a}(t-a).$$

Dann ist F stetig auf $[a,b]$, differenzierbar auf (a,b) und $F(a) = F(b) = 0$. Es gilt

$$F'(x) = f'(x) - \frac{f(b)-f(a)}{b-a},$$

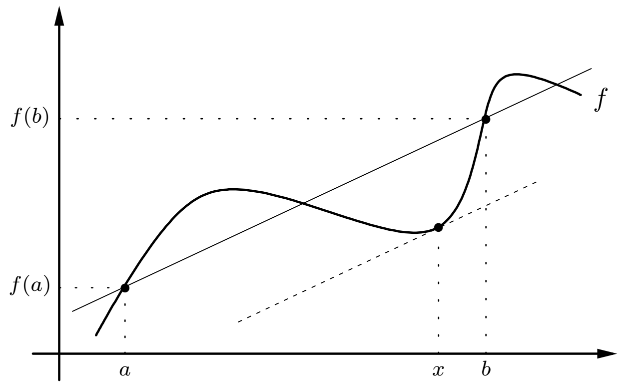
also genügt es zu zeigen, daß eine Stelle $x \in (a,b)$ existiert mit $F'(x) = 0$.

Ist $F(t) = 0$ für alle $t \in [a,b]$, so ist auch $F'(x) = 0$ für alle $x \in (a,b)$, dieser Fall ist also trivial. Betrachte nun den Fall daß ein $t_0 \in (a,b)$ mit $F(t_0) > 0$ existiert. Als stetige Funktion auf der kompakten Menge $[a,b]$ besitzt F ein Maximum x. Wegen $F(a) = F(b) = 0 < F(t_0)$ muß dieses im Inneren des Intervalls liegen. Natürlich ist x insbesondere ein lokales Maximum, also gilt $F'(x) = 0$.

Sollte stets $F(t) \leq 0$ gelten, so muß ein $t_0 \in (a,b)$ mit $F(t_0) < 0$ existieren (da sonst der Trivialfall $F \equiv 0$ vorliegt). Dann argumentiert man genauso mit dem Minimum von F wie oben mit dem Maximum.

❑

3.3.4 Bemerkung. Geometrisch bedeutet Satz 3.3.2, daß an einem lokalen Extremum die Tangente an die Kurve $y = f(x)$, falls eine solche existiert, waagrecht liegen muß. Satz 3.3.3 besagt, daß man (falls durchwegs Tangenten existieren) stets eine Tangente findet, welche parallel zur Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ liegt.



Offensichtlich ist die Stelle x an der die Steigung der Kurve gerade gleich der mittleren Steigung im Intervall $[a,b]$ ist nicht eindeutig bestimmt.

Der Mittelwertsatz wird manchmal auch als *1.Mittelwertsatz der Differentialrechnung* bezeichnet. Der Spezialfall $f(a) = f(b)$ heißt auch *Satz von Rolle*[1].

Eine unmittelbare Folgerung ist:

[1]Michel Rolle, geb.21.4.1652 Ambert (Frankreich), gest.8.11.1719 Paris

3.3.5 Korollar. *Sei $f : [a, b] \to \mathbb{R}$ stetig und n-mal differenzierbar auf (a, b). Weiters habe f mindestens $n + 1$ Nullstellen in $[a, b]$. Dann existiert $\xi \in (a, b)$ mit $f^{(n)}(\xi) = 0$.*

Beweis. Der Fall $n = 1$ ist gerade der Satz von Rolle. Zwischen je zwei Nullstellen von f liegt mindestens eine Nullstelle von f', also hat f' (mindestens) n Nullstellen. Nach Induktionsvoraussetzung existiert ξ mit $(f')^{(n-1)}(\xi) = 0$.

❑

Satz 3.3.3 läßt sich in der folgenden Weise verallgemeinern:

3.3.6 Satz (Verallgemeinerter Mittelwertsatz)**.** *Seien f und g stetig auf $[a, b]$ und differenzierbar auf (a, b). Weiters sei $g'(t) \neq 0$ für alle $t \in (a, b)$. Dann existiert eine Stelle $x \in (a, b)$ mit*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(x)}{g'(x)}.$$

Beweis. Es existiert eine Stelle $t \in (a, b)$ mit $g(b) - g(a) = g'(t)(b - a)$, also ist $g(b) - g(a) \neq 0$. Betrachte die Funktion

$$F(t) = f(t) - f(a) - \frac{f(b) - f(a)}{g(b) - g(a)} \, (g(t) - g(a)).$$

Es gilt

$$F'(t) = f'(t) - \frac{f(b) - f(a)}{g(b) - g(a)} \, g'(t),$$

also genügt es zu zeigen, daß ein $x \in (a, b)$ mit $F'(x) = 0$ existiert. Diese Aussage folgt mit Satz 3.3.3, denn es gilt $F(a) = F(b) = 0$.

❑

Beachte, daß man für $g(t) = t$ gerade den Mittelwertsatz erhält. Satz 3.3.6 heißt auch *2.Mittelwertsatz der Differentialrechnung* .

3.3.7 Korollar. *Sei f differenzierbar auf (a, b). Dann gilt:*

(*i*) *Ist $f'(x) \geq 0$, für alle $x \in (a, b)$, so ist f monoton wachsend.*

(*ii*) *Ist $f'(x) \leq 0$, für alle $x \in (a, b)$, so ist f monoton fallend.*

(*iii*) *Ist $f'(x) = 0$, für alle $x \in (a, b)$, so ist f konstant.*

Beweis. Sei $a < x_1 < x_2 < b$. Dann existiert eine Stelle x mit

$$f(x_2) - f(x_1) = (x_2 - x_1) f'(x).$$

❑

[2] Guillaume Francois L'Hospital, Marquis de Saint-Mesme, geb.1661 Paris, gest.3.2.1704 Paris

3.3.8 Korollar (Regel von de L'Hospital[2]). *Seien* $a, b, \in \mathbb{R} \cup \{\pm\infty\}, -\infty \leq a < b \leq +\infty$. *Weiters seien* f *und* g *differenzierbar auf* (a, b) *und sei* $g'(x) \neq 0, x \in (a, b)$. *Es gelte*

$$\lim_{x \to a+} \frac{f'(x)}{g'(x)} = A, \tag{3.2}$$

mit $A \in \mathbb{R} \cup \{\pm\infty\}$. *Gilt*

$$\lim_{x \to a+} f(x) = \lim_{x \to a+} g(x) = 0, \tag{3.3}$$

oder

$$\lim_{x \to a+} g(x) = +\infty, \tag{3.4}$$

so folgt

$$\lim_{x \to a+} \frac{f(x)}{g(x)} = A.$$

Die analoge Aussage ist richtig, wenn man überall $x \to a+$ *durch* $x \to b-$ *bzw. in* (3.4) $+\infty$ *durch* $-\infty$ *ersetzt.*

Beweis. Wir betrachten zunächst den Fall $A < +\infty$. Sei $q > A$ und wähle r mit $A < r < q$. Wegen (3.2) existiert ein Punkt $c \in (a, b)$, sodaß für $a < x < c$ gilt

$$\frac{f'(x)}{g'(x)} < r.$$

Für y mit $a < x < y < c$ existiert $t \in (x, y)$ mit

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(t)}{g'(t)} < r. \tag{3.5}$$

Angenommen die Bedingung (3.3) ist erfüllt. Läßt man in obiger Beziehung x gegen $a+$ streben, so folgt

$$\frac{f(y)}{g(y)} \leq r < q \ (a < y < c).$$

Sei nun angenommen die Bedingung (3.4) ist erfüllt. Hält man y in (3.5) fest, so existiert ein Punkt $c_1 \in (a, y)$ mit $g(x) > g(y)$ und $g(x) > 0$ für $a < x < c_1$. Es folgt

$$\frac{f(x)}{g(x)} < \frac{f'(t)}{g'(t)} \frac{g(x) - g(y)}{g(x)} + \frac{f(y)}{g(x)} < r - r\frac{g(y)}{g(x)} + \frac{f(y)}{g(x)}, \ a < x < c_1.$$

Läßt man hier $x \to a+$ streben, so erhält man die Existenz eines $c_2, a < c_2 < c_1$, sodaß

$$\frac{f(x)}{g(x)} < q, \ a < x < c_2.$$

Wir haben also die folgende Aussage bewiesen: Ist $q > A$, so existiert $C > a$ mit $\frac{f(x)}{g(x)} < q$ für alle $x \in (a, C)$.

Genau so erhält man (falls $A > -\infty$) die Aussage: Ist $p < A$, so existiert $C' > a$ mit $\frac{f(x)}{g(x)} > p$ für alle $x \in (a, C')$. Insgesamt folgt $\lim_{x \to a+} \frac{f(x)}{g(x)} = A$.

❑

3.3.9 Bemerkung. Oft findet man mit dem folgenden Spezialfall der Regel von de L'Hospital, welcher wesentlich einfacher und instruktiver zu beweisen ist, das Auslangen: Seien f, g differenzierbar auf einem Intervall (c, d), sei $x \in (c, d)$, $f(x) = g(x) = 0$ und $g'(x) \neq 0$. Dann gilt

$$\lim_{t \to x} \frac{f(t)}{g(t)} = \frac{f'(x)}{g'(x)}.$$

Diese Aussage sieht man wie folgt ein: Da f und g differenzierbar sind, kann man sie anschreiben als

$$f(t) = f(x) + f'(x)(t - x) + \mu(t)(t - x), \; g(t) = g(x) + g'(x)(t - x) + \nu(x)(t - x),$$

wobei $\lim_{t \to x} \mu(t) = \lim_{t \to x} \nu(t) = 0$. Dann ist also wegen unserer Voraussetzungen an f und g für $t \neq x$

$$\frac{f(t)}{g(t)} = \frac{f'(x) + \mu(t)}{g'(x) + \nu(t)},$$

und damit $\lim_{t \to x} \frac{f(t)}{g(t)} = \frac{f'(x)}{g'(x)}$.

3.3.10 Beispiel. Wir setzen wieder voraus dass die Ableitungen von e^x, $\sin x$, etc. schon bekannt sind.

(*i*) Wir wollen untersuchen ob der Grenzwert $\lim_{x \to 0+} x \ln x$ existiert und welchen Wert er in diesem Fall hat. Dazu betrachte die Funktionen $f(x) = \ln x$, $g(x) = \frac{1}{x}$, für $x \in (0, \infty)$. Sie sind auf diesem Intervall differenzierbar und es ist $f'(x) = \frac{1}{x}$, $g'(x) = -\frac{1}{x^2}$. Wir sehen dass $g'(x) \neq 0$, $x \in (0, \infty)$, und $\lim_{x \to 0+} g(x) = \infty$. Weiters gilt

$$\lim_{x \to 0+} \frac{f'(x)}{g'(x)} = \lim_{x \to 0+} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x \to 0+} (-x) = 0\,.$$

Also folgt nach der Regel von de L'Hospital, dass auch der Limes $\lim_{x \to 0+} \frac{f(x)}{g(x)}$ existiert und gleich Null ist.

(*ii*) Sei $n \in \mathbb{N} \cup \{0\}$. Betrachtet man die e^x definierende Potenzreihe, so sieht man dass für $x > 0$ sicher

$$e^x \geq \frac{x^{n+1}}{(n+1)!}\,.$$

Es folgt das $\lim_{x \to +\infty} \frac{x^n}{e^x} = 0$ gilt. Wir wollen diese Aussage auch mittels der Regel von de L'Hospital einsehen. Dazu verwenden wir Induktion nach n. Für $n = 0$ haben wir diese Aussage gleich nach der Definition von e^x bewiesen. Sei angenommen sie sei richtig für ein $n \in \mathbb{N} \cup \{0\}$, und betrachte den Limes $\lim -x \to +\infty \frac{x^{n+1}}{e^x}$. Setze $f(x) = x^{n+1}$, $g(x) = e^x$, für $x \in \mathbb{R}$. Dann sind f und g differenzierbar, $f'(x) = (n+1)x^n$, $g'(x) = e^x$, und es ist $g'(x) \neq 0$ für alle x. Weiters gilt (nach Induktionsvoraussetzung)

$$\lim_{x \to +\infty} \frac{f'(x)}{g'(x)} = 0$$

und, nach dem Fall $n = 0$, $\lim_{x \to +\infty} g(x) = +\infty$. Also folgt dass auch der Limes $\lim_{x \to +\infty} \frac{f(x)}{g(x)}$ existiert und gleich 0 ist.

(*iii*) Wir wollen die Existenz und, gegebenfalls, den Wert des Limes $\lim_{x\to 0}\frac{\cos x-1}{x^2}$ untersuchen. Setze $f(x)=\cos x-1$, $g(x)=x^2$, für $x\in\mathbb{R}$. Dann sind f,g differenzierbar auf $\mathbb{R}$, $f'(x)=-\sin x$, $g'(x)=2x$, und es gilt $g'(x)\neq 0$ für $x\neq 0$. Sollte also der Limes $\lim_{x\to 0}\frac{f'(x)}{g'(x)}$ existieren und einen Wert A haben, so könnten wir die Regel von de L'Hosiptal anwenden (auf die Funktionen f,g einmal am Intervall $(0,\infty)$ und einmal auf $(-\infty,0)$), und schliessen, dass auch der Limes $\lim_{x\to 0}\frac{f(x)}{g(x)}$ existiert und den selben Wert A hat.

Nun ist $\frac{f'(x)}{g'(x)}=\frac{-\sin x}{2x}$. Um die Existenz und den Wert dieses Limes zu untersuchen, wenden wir nochmals die Regel von de L'Hospital an. Setze also $f_1(x)=-\sin x$, $g_1(x)=2x$. Dann sind f_1,g_1 auf ganz $\mathbb{R}$ differenzierbar, $f_1'(x)=-\cos x$, $g_1'(x)=2$, und es gilt $g_1'(0)=0$. Wir können also die Variante der Regel von de L'Hospital aus Bemerkung 3.3.9 anwenden, und erhalten

$$\lim_{x\to 0}\frac{f_1(x)}{g_1(x)}=\lim_{x\to 0}\frac{f_1'(x)}{g_1'(x)}=\lim_{x\to 0}\frac{-\cos x}{2}=-\frac{1}{2}\,.$$

Nach der Argumentation des obigen ersten Teils, folgt damit auch $\lim_{x\to 0}\frac{\cos x-1}{x^2}=-\frac{1}{2}$.

Dieses Ergebnis könnte man auch unmittelbarer ersehen: Dazu bemerke dass nach Definition

$$\cos x-1=\big(1-\frac{x^2}{2}+\frac{x^4}{4!}\mp\ldots\big)-1=-\frac{x^2}{2}+\frac{x^4}{4!}\mp\ldots$$

gilt. Daher ist

$$\frac{\cos x-1}{x^2}=-\frac{1}{2}+\frac{x^2}{4!}\mp\ldots\,.$$

Die Reihe auf der rechten Seiten ist eine Potenzreihe mit Konvergenzradius ∞, sie ist also z.B. für $|x|\leq 1$ gleichmässig konvergent. Daher gilt

$$\lim_{x\to 0}\frac{\cos x-1}{x^2}=\lim_{x\to 0}\big(-\frac{1}{2}+\frac{x^2}{4!}\mp\ldots\big)=$$

$$=\lim_{x\to 0}\big(-\frac{1}{2}\big)+\lim_{x\to 0}\big(\frac{x^2}{4!}\big)\mp\ldots=-\frac{1}{2}\,.$$

Sei nun $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktion welche gegen eine Funktion f konvergiert. Überträgt sich diese Eigenschaft auf die abgeleiteten Funktionen?

3.3.11 Beispiel. Sei $f_n(x)=\frac{\sin\ nx}{\sqrt{n}}$ für $x\in\mathbb{R}$. Dann gilt

$$\lim_{n\to\infty}f_n(x)=0,$$

und das sogar gleichmäßig auf ganz $\mathbb{R}$. Man berechnet (wir benützen wieder die Eigenschaften der Sinusfunktion, vgl. später)

$$f_n'(x)=\sqrt{n}\ \cos\ nx,$$

also gilt z.B. $\lim_{n\to\infty}f_n'(0)=+\infty$.

3.3.12 Satz. *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen, definiert und differenzierbar auf $[a,b]$. Es gelte:*

(i) Es existiert ein Punkt $x_0 \in [a,b]$ sodaß $(f_n(x_0))_{n\in\mathbb{N}}$ konvergiert.

(ii) Die Folge $(f_n')_{n\in\mathbb{N}}$ ist gleichmäßig konvergent auf $[a,b]$.

Dann ist die Folge $(f_n)_{n\in\mathbb{N}}$ gleichmäßig konvergent auf $[a,b]$ und es gilt

$$\frac{d}{dx}\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty}\frac{d}{dx}f_n(x),\ x\in[a,b].$$

Beweis. Sei $\epsilon > 0$ gegeben. Wähle $N \in \mathbb{N}$ sodaß für $n, m \geq N$ stets gilt

$$|f_n(x_0) - f_m(x_0)| < \epsilon,$$

und

$$|f_n'(t) - f_m'(t)| < \epsilon,\ t \in [a,b].$$

Nach dem Mittelwertsatz gilt ($x \in [a,b]$)

$$f_n(x) - f_m(x) = [f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)] + [f_n(x_0) - f_m(x_0)] =$$

$$= (x - x_0)\,[f_n'(t) - f_m'(t)] + [f_n(x_0) - f_m(x_0)],$$

mit einer geeigneten Zwischenstelle t. Es folgt

$$|f_n(x) - f_m(x)| < (b-a)\epsilon + \epsilon.$$

Also konvergiert $(f_n)_{n\in\mathbb{N}}$ gleichmäßig auf $[a,b]$. Sei $f = \lim_{n\to\infty} f_n$.

Sei nun $x \in (a,b)$ festgehalten. Setze

$$\phi_n(t) := \frac{f_n(t) - f_n(x)}{t - x},\ \phi(t) := \frac{f(t) - f(x)}{t - x},$$

dann gilt für n fest aber beliebig

$$\lim_{t\to x}\phi_n(t) = f_n'(x).$$

Weiters ist

$$\phi_n(t) - \phi_m(t) = \frac{f_n(t) - f_m(t) - f_n(x) + f_m(x)}{t - x} = f_n'(s) - f_m'(s)$$

mit einer geeigneten Stelle s zwischen t und x. Da (f_n') gleichmäßig konvergiert, existiert auch der Limes $\lim_{n\to\infty}\phi_n(t)$ gleichmäßig bzgl. $t \neq x$. Man erhält daher

$$\lim_{t\to x}\lim_{n\to\infty}\phi_n(t) = \lim_{n\to\infty}\lim_{t\to x}\phi_n(t).$$

Die linke Seite ist gleich $f'(x)$, die rechte ist $\lim_{n\to\infty} f_n'(x)$.

❑

3.3.13 Beispiel. Sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit positivem Konvergenzradius R. Dann gilt für jedes x mit $|x| < R$

$$\frac{d}{dx}\sum_{n=0}^{\infty} a_n x^n = \sum_{n=1}^{\infty} n a_n x^{n-1}.$$

Denn die rechte Reihe hat auch den Konvergenzradius R.

3.3.14 Beispiel. Wir können nun die Ableitungen von durch Potenzreihen definierten Funktionen berechnen. Es gilt

$$(e^x)' = e^x,\ (\cos x)' = -\sin x,\ (\sin x)' = \cos x\,.$$

Denn nach dem letzten Beispiel ist

$$(e^x)' = \Big(\sum_{n=0}^{\infty} \frac{x^n}{n!}\Big)' = \sum_{n=1}^{\infty} \frac{nx^{n-1}}{n!} =$$

$$= \sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x\,,$$

sowie

$$(\cos x)' = \Big(\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}\Big)' = \sum_{n=1}^{\infty} (-1)^n \frac{(2n)x^{2n-1}}{(2n)!} =$$

$$= -\sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = -\sin x\,,$$

und

$$(\sin x)' = \Big(\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}\Big)' = \sum_{n=0}^{\infty} (-1)^n \frac{(2n+1)x^{2n}}{(2n+1)!} =$$

$$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \cos x\,.$$

Obwohl die Ableitung f' einer auf einem Intervall (a, b) differenzierbaren Funktion nicht notwendig stetig sein muß, so gilt trotzdem stets die Zwischenwerteigenschaft.

3.3.15 Satz. *Sei f differenzierbar auf $[a, b]$ und sei $c \in \mathbb{R}$ mit $f'(a) < c < f'(b)$. Dann existiert eine Stelle $x \in (a, b)$ mit $f'(x) = c$.*

Beweis. Betrachte die Funktion

$$g(t) = f(t) - ct,\ a \le t \le b.$$

Dann gilt

$$g'(a) = f'(a) - c < 0,\ g'(b) = f'(b) - c > 0.$$

Sei x eine Stelle an der g sein Minimum annimmt. Wegen $0 = g'(x) = f'(x) - c$, genügt es zu zeigen, daß $x \in (a, b)$ liegt.

Da $g'(a) < 0$ ist, existiert $\delta > 0$ mit

$$\frac{g(t) - g(a)}{t - a} < 0,\ a < t < a + \delta.$$

Also folgt $g(t) < g(a)$ für solche Werte von t. Der Punkt a scheidet als Kanditat für das Minimum also aus. Wegen $g'(b) > 0$ folgt die Existenz eines $\delta' > 0$ sodaß

$$\frac{g(t) - g(b)}{t - b} > 0,\ b - \delta' < t < b.$$

Also ist $g(t) < g(b)$ für solche t, und der Punkt b kommt daher auch nicht in Frage.

❑

3.3.16 Korollar. *Sei f differenzierbar auf $[a, b]$. Dann hat f' keine Sprungstelle in $[a, b]$.*

Beweis. An einer Sprungstelle existiert $\lim_{t \to x+} f(t)$ und $\lim_{t \to x-} f(t)$, es sind jedoch nicht beide gleich $f(x)$. Also kann die Zwischenwerteigenschaft nicht gelten.

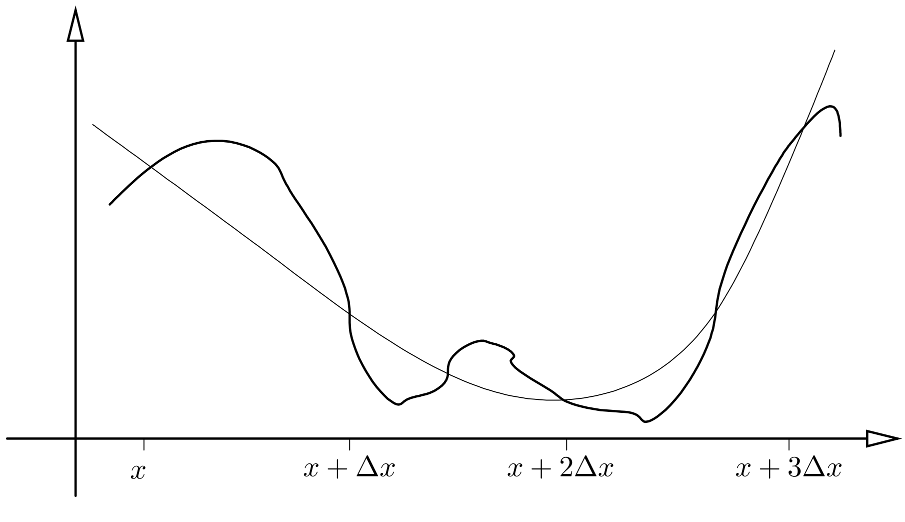
❑

3.4 Der Taylorsche Lehrsatz

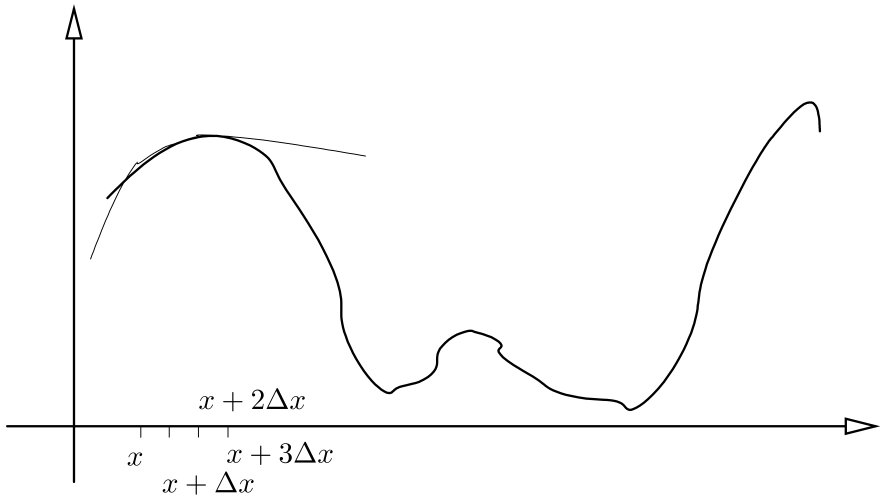
Die Gerade welche eine Kurve in einem Punkt am besten approximiert ist die Tangente (falls sie existiert). Approximiert man die Kurve mit einem Polynom höheren Grades, so hofft man, daß die Approximation genauer wird.

Wir haben die Tangente gefunden (eigentlich definiert) als die Grenzlage von Sekanten durch die Punkte $(x, f(x))$ und $(x + \triangle x, f(x + \triangle x))$. Da eine Gerade durch zwei Punkte eindeutig betimmt ist, sind diese Sekanten wohldefinierte Objekte.

Ein Polynom n-ten Grades ist eindeutig festgelegt durch die Vorgabe der Werte an $n+1$ verschiedenen Stellen. Betrachten wir also die $n+1$ Punkte der Kurve f mit den x-Koordinaten $x, x+\triangle x, \ldots, x+n\triangle x$, und legen ein Polynom p durch diese Punkte.



Für große Schrittweite $\triangle x$ wird das erhaltene Polynom nicht viel mit der Kurve zu tun haben, läßt man jedoch $\triangle x \to 0$ streben, so erhofft man sich eine gute Approximation.



Um diese Idee zu formalisieren, müssen wir eine explizite Darstellung für ein Polynom erhalten, welches die Funktion f an den Stellen $x, x+\Delta x, \ldots, x+n\Delta x$ interpoliert.

Hat man allgemein paarweise verschiedene Stützstellen $x_0, \ldots, x_n$ und zugehörige Werte $y_0, \ldots, y_n$ gegeben, so liefert der folgenden Satz eine Lösung des Interpolationsproblems.

3.4.1 Satz (Newtonsche Interpolationsformel)**.** *Aus den Stützstellen $x_0, \ldots, x_n$ und zugehörigen Werten $y_0, \ldots, y_n$ definieren wir rekursiv*

(i) $[x_i] := y_i$, $i = 0, \ldots, n$.

(ii) $[x_i x_{i+1} \ldots x_k] := \frac{[x_{i+1} \ldots x_k] - [x_i \ldots x_{k-1}]}{x_k - x_i}$, $0 \leq i < k \leq n$.

Dann gilt für das Polynom

$$P(x) := [x_0] + [x_0 x_1](x - x_0) + \ldots + [x_0 x_1 \ldots x_n](x - x_0)(x - x_1) \cdots (x - x_{n-1}) \tag{3.6}$$

dass

$$P(x_k) = y_k, \;\; k = 0, \ldots, n\,.$$

Zum Beweis benützen wir die folgende Aussage.

3.4.2 Lemma. *Seien $x_0, x_1, \ldots$ und $y_0, y_1, \ldots$ gegeben. Setzt man*

$$\omega_{i,k}(x) := \prod_{l=i}^{k} (x - x_l), \;\; 0 \leq i < k\,,$$

so gilt

$$[x_i \ldots x_k] = \sum_{j=i}^{k} \frac{y_j}{\omega'_{i,k}(x_j)}\,.$$

Beweis. Als erstes bemerken wir, dass für $0 \leq i < k$ und $i \leq j \leq k$ gilt

$$\omega'_{i,k}(x_j) = \prod_{\substack{l=i \\ l \neq j}}^{k} (x_j - x_l)\,.$$

Die Behauptung beweisen wir nun mittels Induktion nach der Länge $k-i$.

Betrachte den Fall dass $k-i=1$ ist. Sei also $0 \le i$ und $k=i+1$. Dann gilt

$$[x_i x_{i+1}] = \frac{[x_{i+1}]-[x_i]}{x_{i+1}-x_i} = \frac{y_{i+1}}{x_{i+1}-x_i} + \frac{y_i}{x_i - x_{i+1}} \,.$$

Sei nun angenommen die Behauptung ist bewiesen für alle i,k mit $k-i \le n$ und sei i,k mit $0 \le i < k, k-i=n$ gegeben. Dann gilt

$$[x_i \dots x_{k+1}] = \frac{[x_{i+1}\dots x_{k+1}] - [x_i \dots x_k]}{x_{k+1}-x_i} =$$

$$= \frac{1}{x_{k+1}-x_i}\Big(\sum_{j=i+1}^{k+1} \frac{y_j}{\prod_{\substack{l=i+1\\ l\neq j}}^{k+1}(x_j - x_l)} - \sum_{j=i}^{k} \frac{y_j}{\prod_{\substack{l=i\\ l\neq j}}^{k}(x_j - x_l)} \Big) =$$

$$= \frac{1}{x_{k+1}-x_i} \sum_{j=i+1}^{k} \frac{1}{\prod_{\substack{l=i+1\\ l\neq j}}^{k}(x_j-x_l)} \cdot \Big(\frac{y_j}{x_j - x_{k+1}} - \frac{y_j}{x_j - x_i} \Big) +$$

$$+ \frac{y_{k+1}}{\prod_{l=i}^{k}(x_{k+1}-x_l)} + \frac{y_i}{\prod_{l=i+1}^{k+1}(x_i - x_l)} \,.$$

Nun gilt

$$\frac{1}{x_{k+1}-x_i}\Big(\frac{y_j}{x_j - x_{k+1}} - \frac{y_j}{x_j - x_i} \Big) = \frac{y_j}{(x_j - x_{k+1})(x_j - x_i)} \,,$$

und damit folgt weiter dass

$$[x_i \dots x_{k+1}] = \sum_{j=i}^{k+1} \frac{y_j}{\prod_{\substack{l=i\\ l\neq j}}^{k+1}(x_j - x_l)} \,.$$

❑

Insbesondere zeigt uns dieses Lemma, dass der Ausdruck $[x_i \dots x_k]$ unabhängig von der Reihenfolge der Werte von $x_i, \dots, x_k$ ist.

Beweis. (Satz 3.4.1) Für gegebene $x_0, \dots, x_n$ und $y_0, \dots, y_n$ bezeichne das Polynom aus (3.6) mit $P((x_0; y_0), \dots, (x_n; y_n))$. Wir verwenden Induktion nach der Anzahl $n+1$ der Stützstellen. Für $n=0$ ist

$$P((x_0;y_0)) = [x_0] = y_0 \,.$$

Sei nun angenommen, dass für je n Stützstellen bereits bewiesen ist dass das Polynom (3.6) interpoliert, und seien $n+1$ Stützstellen $x_0, \dots, x_n$ mit zugeörigen Werten $y_0, \dots, y_n$ gegeben. Für $0 \le k < n$ gilt nach Induktionsvoraussetzung

$$P((x_0;y_0), \dots, (x_n;y_n))(x_k) = P((x_0;y_0), \dots, (x_k;y_k))(x_k) = y_k \,.$$

Es ist

$$[x_0 \dots x_n] = [x_n x_0 \dots x_{n-1}] = \frac{[x_0 \dots x_{n-1}] - [x_n x_0 \dots x_{n-2}]}{x_{n-1} - x_n} =$$

$$= \frac{[x_0 \dots x_{n-1}] - [x_0 \dots x_{n-2} x_n]}{x_{n-1} - x_n},$$

und wir berechnen

$$P((x_0; y_0), \dots, (x_n; y_n))(x_n) = [x_0] + \dots + [x_0 \dots x_{n-1}](x_n - x_0) \cdots (x_n - x_{n-2}) +$$
$$+ [x_0 \dots x_n](x_n - x_0) \cdots (x_n - x_{n-2})(x_n - x_{n-1}) =$$
$$= [x_0] + \dots + [x_0 \dots x_{n-1}](x_n - x_0) \cdots (x_n - x_{n-2}) -$$
$$-\big([x_0 \dots x_{n-1}] - [x_n x_0 \dots x_{n-2}]\big)(x_n - x_0) \cdots (x_n - x_{n-2}) =$$
$$= [x_0] + \dots + [x_0 \dots x_{n-2}](x_n - x_0) \cdots (x_n - x_{n-3}) +$$
$$+ [x_0 \dots x_{n-2} x_n](x_n - x_0) \cdots (x_n - x_{n-2}) =$$
$$= P((x_0; y_0), \dots (x_{n-2}; y_{n-2}), (x_n, y_n))(x_n) = y_n .$$

❑

Im Zusammenhang mit der Taylorschen Formel ist der Spezialfall der Newtonsche Interpolationsformel von Bedeutung wo die Stützstellen gleiche Abstände haben, d.h. wo $x_j = x + j\Delta x$ ist.

3.4.3 Korollar. *Sei $x, \triangle x$ und Werte $y_0, \dots, y_n$ gegeben. Das Polynom welches durch die Punkte $(x, y_0), (x + \triangle x, y_1), \dots, (x + n\triangle x, y_n)$ geht ist gleich*

$$p(x) = y_0 + \frac{(x - x_0)}{1!} \frac{\triangle y_0}{\triangle x} + \frac{(x - x_0)(x - x_1)}{2!} \frac{\triangle^2 y_0}{\triangle x^2} + \cdots$$
$$\dots + \frac{(x - x_0)(x - x_1) \dots (x - x_n)}{(n+1)!} \frac{\triangle^{n+1} y_0}{\triangle x^{n+1}},$$

wobei wir $x_j = x + j\triangle x$ gesetzt haben und $\triangle^j y_0$ rekursiv definiert als

$$\triangle y_i = y_{i+1} - y_i, \ \triangle^{k+1} y_i = \triangle^k y_{i+1} - \triangle^k y_i .$$

Beweis. Wir zeigen dass $\Delta^k y_i = k!(\Delta x)^k [x_i \dots x_{i+k}]$ gilt. Dazu zeigen wir dass diese Zahlen den gleichen rekursiven Beziehung genügen. Für $k = 1$ haben wir

$$\Delta x [x_i x_{i+1}] = [x_{i+1}] - [x_i] = y_{i+1} - y_i .$$

Weiter ist

$$(k+1)!(\Delta x)^{k+1} [x_i \dots x_{i+k+1}] = (k+1)!(\Delta x)^{k+1} \frac{[x_{i+1} \dots x_{i+k+1}] - [x_i \dots x_{i+k}]}{x_{i+k+1} - x_i} =$$
$$= k!(\Delta x)^k [x_{i+1} \dots x_{i+k+1}] - k!(\Delta x)^k [x_i \dots x_{i+k}] .$$

Setzt man in die Formel (3.6) ein, so erhält man das gewünschte Ergebnis.

❑

Ist f an der Stelle x differenzierbar, so gilt $\lim_{\triangle x \to 0} \frac{\triangle y_0}{\triangle x} = \lim_{\triangle x \to 0} \frac{f(x + \triangle x) - f(x)}{\triangle x} = f'(x)$. Allgemein gilt:

3.4.4 Lemma. *Sei $f : [a,b] \to \mathbb{R}$ stetig und n-mal stetig differenzierbar auf (a,b). Ist $x \in (a,b)$, so gilt ($y_j = f(x + j\triangle x)$)*

$$\lim_{\triangle x \to 0} \frac{\triangle^n y_0}{\triangle x^n} = f^{(n)}(x).$$

Beweis. Sei $p(x) = y_0 + \frac{(x-x_0)}{1!}\frac{\triangle y_0}{\triangle x} + \ldots + \frac{(x-x_0)\cdots(x-x_{n-1})}{n!}\frac{\triangle^n y_0}{\triangle x^n}$. Die Funktion $h(x) := f(x) - p(x)$ hat die $n+1$ Nullstellen $x_0, \cdots, x_n$ ($\in (a,b)$ für $\triangle x$ hinreichend klein). Mit Korollar 3.3.5 folgt die Existenz von $\xi \in (x_0, x_n)$ mit $h^{(n)}(\xi) = 0$. Nun gilt

$$0 = h^{(n)}(\xi) = f^{(n)}(\xi) - p^{(n)}(\xi) = f^{(n)}(\xi) - \frac{\triangle^n y_0}{\triangle x^n}.$$

Für $\triangle x \to 0$ folgt wegen der Stetigkeit von $f^{(n)}$ auch $\frac{\triangle^n y_0}{\triangle x^n} \to f^{(n)}(x)$.

❏

Man erhält also als Grenzfall des in einem Punkt x_0 approximierenden Polynoms gerade

$$p(x) = f(x_0) + (x-x_0)f'(x_0) + \frac{(x-x_0)^2}{2}f''(x_0) + \cdots + \frac{(x-x_0)}{n!}f^{(n)}(x_0).$$

3.4.5 Satz (Taylorscher Lehrsatz)**.** *Sei f definiert auf $[a,b]$, sei $f^{(n-1)}$ stetig auf $[a,b]$ und existiere $f^{(n)}$ auf (a,b). Dann existiert eine Stelle $\xi \in (a,b)$ mit*

$$f(b) = \sum_{k=0}^{n-1} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^n}{n!} f^{(n)}(\xi).$$

Beweis. Sei die Zahl M definiert durch

$$f(b) - \sum_{k=0}^{n-1} \frac{(b-a)^k}{k!} f^{(k)}(a) = \frac{(b-a)^n}{n!} M.$$

Wir müßen zeigen $\exists\, \xi \in (a,b) : M = f^{(n)}(\xi)$. Betrachte die Funktion ($a \le t \le b$)

$$F(t) = f(b) - \sum_{k=0}^{n-1} \frac{(b-t)^k}{k!} f^{(k)}(t) - M\frac{(b-t)^n}{n!}.$$

Es gibt $F(b) = F(a) = 0$, also existiert ein Punkt $\xi \in (a,b)$ mit $F'(\xi) = 0$. Nun ist

$$F'(t) = -\sum_{k=0}^{n-1} \frac{(b-t)^k}{k!} f^{(k+1)}(t) + \sum_{k=1}^{n-1} \frac{k(b-t)^{k-1}}{k!} f^{(k)}(t) + M\frac{n(b-t)^{n-1}}{n!} =$$

$$= \frac{(b-t)^{n-1}}{(n-1)!}(-f^{(n)}(t) + M).$$

❏

Der Taylorsche Lehrsatz (wie auch sein Beweis) ist gerade eine Verallgemeinerung des Mittelwertsatzes.

Wir sind davon ausgegangen jenes Polynom vom Grad n zu suchen welches die Kurve am besten approximiert. Läßt man Polynome höheren Grades zu, so wird die Approximation wohl besser werden. In „guten“ Fällen wird man also erhoffen, daß man die Funktion f als Potenzreihe darstellen kann.

3.4.6 Definition. Sei die Funktion f an der Stelle x_0 beliebig oft differenzierbar. Dann heisst die Reihe

$$\mathrm{TR}[f,x_0](x) := \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

die *Taylorreihe* von f mit Anschlussstelle x_0.

3.4.7 Beispiel.

(*i*) Sei $f(t) = e^t$. Dann gilt $f^{(n)}(t) = e^t$, also $f^{(n)}(0) = 1$. Wir erhalten als Taylorreihe also $\mathrm{TR}[e^t, 0](x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$. Diese Reihe konvergiert für alle $x \in \mathbb{R}$ und zwar tatsächlich gegen die Funktion f.

Betrachtet man den Taylorschen Lehrsatz, so hat man

$$e^x = \sum_{k=0}^{n-1} \frac{x^k}{k!} + \frac{x^n}{n!} e^{\xi}$$

mit $\xi \in (0, x)$. Man sieht also auch von diesem Standpunkt, dass der Fehlerterm punktweise (sogar gleichmäßig auf jeder kompakten Menge) gegen Null strebt.

(*ii*) Sei $\sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit positiven Konvergenzradius R. Wir betrachten die von ihr dargestellte Funktion $f(x) := \sum_{k=0}^{\infty} a_k x^k$, $|x| < R$. Wegen Beispiel 3.3.13 gilt für jedes $n \in \mathbb{N}$

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1)\cdots(k-n+1) a_k x^{k-n}, \ |x| < R\,,$$

und damit $f^{(n)}(0) = n! a_n$. Wir sehen, dass die Taylorreihe von f mit Anschlussstelle 0 gegeben ist als

$$\mathrm{TR}\Big[\sum_{k=0}^{\infty} a_k t^k, 0\Big](x) = \sum_{n=0}^{\infty} a_n x^n\,.$$

Sie ist also genau die gleiche Potenzreihe wie die mit der wir gestartet sind. Es gilt also: Wird eine Funktion f durch eine Potenzreihe mit positiven Konvergenzradius dargestellt, so ist ihre Taylorreihe konvergent und zwar gegen f.

(*iii*) Betrachte die Funktion

$$f(t) = \sum_{k=1}^{\infty} \frac{\cos(2^k t)}{k!}.$$

Differenziert man diese Reihe gliedweise, so erhält man

$$\sum_{k=1}^{\infty} \frac{-2^k \sin(2^k t)}{k!}, \sum_{k=1}^{\infty} \frac{-2^{2k} \cos(2^k t)}{k!}, \dots$$

Da $\sum_{k=1}^{\infty} \frac{(2^k)^l}{k!}$ für jedes $l \in \mathbb{N}$ konvergiert, sind sämtliche dieser Reihen gleichmäßig konvergent auf $\mathbb{R}$. Die Funktion f ist also in jedem Punkt beliebig oft differenzierbar und ihre Ableitungen werden durch obige Reihen dargestellt. Es gilt daher

$$f'(0) = f'''(0) = \dots = f^{(2k+1)}(0) = \dots = 0,$$

und

$$f^{(2n)}(0) = (-1)^n \sum_{k=1}^{\infty} \frac{2^{2nk}}{k!} = (-1)^n (e^{4^n} - 1).$$

die Taylorreihe von f bei 0 ist also gleich

$$\mathrm{TR}[f, 0](x) = \sum_{n=0}^{\infty} \frac{(-1)^n (e^{4^n} - 1)}{(2n)!} x^{2n}.$$

Wendet man das Quotientenkriterium an, so erhält man $(a_n = \frac{(-1)^n (e^{4^n} - 1)}{(2n)!} x^{2n})$.

$$|\frac{a_{n+1}}{a_n}| = \frac{(e^{4^{n+1}} - 1)}{(e^{4^n} - 1)(2n+2)(2n+1)} x^2 =$$

$$= \frac{(e^{3 \cdot 4^n} + e^{2 \cdot 4^n} + e^{4^n} + 1)}{(2n+2)(2n+1)} x^2 \longrightarrow \infty, \; x \neq 0\,.$$

Also ist diese Reihe für kein x (außer im Trivialfall $x = 0$) konvergent.

(*iv*) Sei

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Wir überlegen uns, dass diese Funktion beliebig oft differenzierbar ist, und dass stets $f^{(n)}(0) = 0$ gilt. Dazu zeigen wir zunächst: Sei $x \neq 0$, dann ist f beliebig oft differenzierbar und $f^{(n)}(x)$ lässt sich anschreiben in der Form

$$f^{(n)}(x) = \sum_{k=0}^{3n} a_{n,k} \frac{1}{x^k} \cdot e^{-\frac{1}{x^2}}$$

mit geeigneten Konstanten $a_{n,k}$.

Im Fall $n = 0$ wähle $a_{0,0} := 1$. Im Fall $n = 1$ haben wir $f'(x) = \frac{2}{x^3} e^{-\frac{1}{x^2}}$, wähle also $a_{1,0} := 0$, $a_{1,1} := 0$, $a_{1,2} := 0$, $a_{1,3} := 2$. Angenommen die Behauptung ist schon gezeigt für ein n, dann folgt

$$f^{(n+1)}(x) = \Big[\sum_{k=0}^{3n} a_{n,k} \frac{1}{x^k}\Big]' \cdot e^{-\frac{1}{x^2}} + \Big[\sum_{k=0}^{3n} a_{n,k} \frac{1}{x^k}\Big] \cdot \frac{2}{x^3} e^{-\frac{1}{x^2}}$$

und die gewünschte Darstellung folgt auch für $n+1$.

Nun betrachten wir die Differenzierbarkeit von f an der Stelle 0. Sei $n=1$, dann haben wir den Limes

$$\lim_{x\to 0}\frac{f(x)-f(0)}{x}=\lim_{x\to 0}\frac{1}{x}e^{-\frac{1}{x^2}}$$

zu untersuchen. Dieser existiert und ist gleich 0, vgl. Beispiel 3.3.10. Sei nun vorausgesetzt, dass für ein $n\in\mathbb{N}$ schon gezeigt ist dass $f^{(n)}(0)$ existiert und gleich 0 ist. Für $f^{(n+1)}(0)$ haben wir den Limes

$$\lim_{x\to 0}\frac{f^{(n)}(x)-f^{(n)}(0)}{x}=\lim_{x\to 0}\frac{1}{x}\cdot\sum_{k=0}^{3n}a_{n,k}\frac{1}{x^k}\cdot e^{-\frac{1}{x^2}}$$

zu betrachten. Wieder sehen wir, dass dieser existiert und gleich 0 ist.

Die Taylorreihe von f ist also gegeben als $\mathrm{TR}[f,0](x)=\sum_{n=0}^{\infty}0\cdot x^n$. Man sieht dass diese Reihe für jedes $x\in\mathbb{R}$ konvergiert, ihre Grenzfunktion (nämlich 0) aber nichts mit der Funktion f zu tun hat.

(v) Manchmal ist es zur Berechnung von Taylorreihen nützlich Symmetrieeigenschaften auszunützen. Eine Funktion $f:(-a,a)\to\mathbb{R}$ heißt *gerade*, wenn $f(-x)=f(x)$, $x\in(-a,a)$, gilt. Sie heißt *ungerade*, wenn $f(-x)=-f(x)$, $x\in(-a,a)$.

Sei f differenzierbar und gerade. Dann gilt

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=-\lim_{h\to 0}\frac{f(-x+(-h))-f(-x)}{-h}=-f'(-x)\,,$$

also ist f' ungerade. Genauso sieht man, dass die Ableitung einer ungeraden Funktion, soferne sie existiert, eine gerade Funktion ist.

Ist g irgendeine ungerade Funktion, so ist also $g(0)=-g(-0)=-g(0)$, und daher $g(0)=0$. Hat man nun eine n-mal differenzierbare Funktion gerade Funktion f gegeben, so folgt $f^{(2k-1)}(0)=0$, für alle k mit $2k-1\leq n$. Analog ist für eine n-mal differenzierbare ungerade Funktion f der Wert $f^{(2k)}(0)$ für alle k mit $2k\leq n$ gleich 0.

Insbesondere treten in der Taylorreihe (in einem Taylorpolynom) einer geraden Funktion nur die Summanden zu geraden Indizes auf, und in der Taylorreihe (in einem Taylorpolynom) einer ungeraden Funktion nur die Summanden zu ungeraden Indizes.

Wir haben gesehen, daß für eine differenzierbare Funktion f, welche an einer Stelle x ein lokales Extremum besitzt, $f'(x)=0$ gelten muß, daß jedoch die Umkehrung im allgemeinen nicht gilt (Beispiel: $f(t)=t^3$ an der Stelle $x=0$). Aus dem Taylorschen Satz erhält man unmittelbar eine hinreichende Bedingung für ein lokales Extremum.

3.4.8 Korollar. *Sei $f:(c,d)\to\mathbb{R}$ eine n-mal stetig differenzierbare Funktion, $x\in(c,d)$, und gelte*

$$f'(x)=f''(x)=\ldots=f^{(n-1)}(x)=0,\ f^{(n)}(x)\neq 0.$$

Ist n gerade, so liegt ein lokales Extremum vor, und zwar ein lokales Minimum falls $f^{(n)}(x)>0$ und ein lokales Maximum falls $f^{(n)}(x)<0$. Ist n ungerade, so liegt kein lokales Extremum vor.

Beweis. Nach der Taylorschen Formel gilt für $t \in (c,d)$ mit einer geeigneten Zwischenstelle ξ zwischen t und x

$$f(t) = f(x) + \frac{(t-x)^n}{n!} f^{(n)}(\xi).$$

Ist nun z.B. $f^{(n)}(x) > 0$, so gilt für ξ in einer hinreichend kleinen Umgebung von x ebenfalls $f^{(n)}(\xi) > 0$. Ist n gerade, so ist für t hinreichend nahe bei x (dann muß ja auch ξ nahe bei x sein) sicher $f(t) > f(x)$, es liegt also ein lokales Minimum vor. Die anderen Fälle folgen genauso.

❑

Wir wissen, dass eine Funktion die an einer Stelle x differenzierbar ist, dort auch stetig ist, und wir haben auch am Beispiel von $f(t) := |t|$ und der Stelle 0 gesehen das die Umkehrung nicht gilt. Zum Abschluß wollen wir noch ein in diesem Kontext wesentlich stärkeres Beispiel angeben.

3.4.9 Beispiel (Weierstraß). Sei $0 < c < 1$ und sei p eine ungerade ganze Zahl mit $pc > 1 + \frac{3\pi}{2}$. Betrachte die Funktion

$$f(x) = \sum_{n=0}^{\infty} c^n \cos(p^n \pi x), \; x \in \mathbb{R}.$$

Diese Reihe konvergiert gleichmäßig, also ist f stetig auf $\mathbb{R}$. Wir werden zeigen, daß f in keinem Punkt differenzierbar ist. Von ähnlichem Typ ist auch die Funktion $g(x) = \sum_{n=1}^{\infty} \frac{\sin(n^2 x)}{n^2}$ von der Riemann noch dachte, daß sie nirgends differenzierbar ist, es jedoch nicht beweisen konnte. Tatsächlich ist g nämlich z.B. bei 0 differenzierbar.

Für $m \in \mathbb{N}$ setze

$$s_m(x) = \sum_{n=0}^{m-1} c^n \cos(p^n \pi x), \; r_m(x) = \sum_{n=m}^{\infty} c^n \cos(p^n \pi x).$$

Sei im folgenden x festgehalten. Wir betrachten die Quotienten

$$\phi_m(h) = \frac{s_m(x+h) - s_m(x)}{h}, \; \psi_m(h) = \frac{r_m(x+h) - r_m(x)}{h}.$$

Nach dem Mittelwertsatz existiert $t \in (x, x+h)$ sodaß

$$\phi_m(h) = s'_m(t) = -\sum_{n=0}^{m-1} c^n p^n \pi \sin(p^n \pi t).$$

Daher ist

$$|\phi_m(h)| \le \pi \frac{1-(cp)^m}{1-cp} < \frac{\pi (cp)^m}{cp-1}.$$

Ist x und m gegeben, so existiert eine Zahl $k \in \mathbb{Z}$ sodaß

$$|p^m x - k| \le \frac{1}{2}.$$

Setze $q = p^m x - k$ und $h_m = \frac{1-q}{p^m}$. Dann gilt $|h_m| \le \frac{3}{2}\frac{1}{p^m}$. Es gilt

$$p^n(x + h_m) = p^{n-m}(p^m x + (1-q)) = p^{n-m}(k+1),$$

also folgt der p ungerade ist.

$$\cos\left[p^n\pi(x+h_m)\right] = (-1)^{k+1},\ n \geq m.$$

Weiters gilt

$$\cos(p^n\pi x) = \cos\left[p^{n-m}\pi(k+q)\right] = \cos(p^{n-m}\pi k)\cos(p^{n-m}\pi q) = (-1)^k\cos(p^{n-m}\pi q)$$

Man erhält

$$\psi_m(h_m) = \frac{(-1)^{k+1}}{h_m}\sum_{n=m}^{\infty} c^n(1+\cos(p^{n-m}\pi q)).$$

Jeder Summand der Reihe ist ≥ 0, also gilt

$$|\psi_m(h_m)| \geq \frac{c^m}{h_m}(1+\cos(\pi q)) \geq \frac{c^m 2p^m}{3}(1+\cos(\pi q)) = \frac{2}{3}(cp)^m.$$

Es folgt

$$\left|\frac{f(x+h_m)-f(x)}{h_m}\right| = |\phi_m(h_m)+\psi_m(h_m)| \geq |\psi_m(h_m)| -$$

$$-|\phi_m(h_m)| \geq (cp)^m\left(\frac{2}{3}-\frac{\pi}{cp-1}\right).$$

Daher ist

$$\left|\frac{f(x+h_m)-f(x)}{h_m}\right| \to +\infty,\ m\to\infty,$$

und es kann f bei x nicht differenzierbar sein.

3.5 Ein Fixpunktsatz

Betrachtet man Beispiele von metrischen Räumen, so fällt auf daß die Metrik häufig definiert ist als $d(p,q) = \|p-q\|$, wo $\|.\|$ irgendein „Betrag“ ist.

3.5.1 Definition. Sei X ein Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$. Eine Abbildung $\|.\| :\ X \to \mathbb{R}^+ \cup \{0\}$ heißt eine *Norm* auf X, falls sie den folgenden Axiomen genügt:

(N1) Ist $\|x\| = 0$ so folgt $x = 0$

(N2) Für $\lambda \in \mathbb{R}$ (oder $\mathbb{C}$) und $x \in X$ gilt $\|\lambda x\| = |\lambda|\cdot\|x\|$.

(N3) Für $x, y \in X$ gilt *Dreiecksungleichung* $\|x+y\| \leq \|x\| + \|y\|$.

Man spricht in diesem Fall von dem *normierten Raum* $(X, \|.\|)$.

Jeder normierte Raum ist insbesondere ein metrischer Raum:

3.5.2 Lemma. *Sei* $(X,\|.\|)$ *ein normierter Raum. Dann ist*

$$d : \begin{cases} X\times X & \to \quad \mathbb{R}^+\cup\{0\} \\ (p,q) & \mapsto \quad \|p-q\| \end{cases}$$

eine Metrik auf X.

Beweis. Seien $p, q \in X$, $d(p,q) = 0$. D.h. es ist $\|p-q\| = 0$, und daher $p-q=0$. Wegen (N2) gilt

$$d(p,q) = \|p-q\| = \|(-1)\cdot(q-p)\| = |-1|\cdot\|q-p\| = \|q-p\| = d(q,p).$$

Schliesslich erhält man aus (N3) die Dreiecksungleichung für d: Seien $p, q, r \in X$, dann ist

$$d(p,q) = \|p-q\| = \|(p-r)+(r-q)\| \leq \|p-r\| + \|r-q\| = d(p,r) + d(r,q).$$

❑

In einem normierten Raum weiß man also was Konvergenz zu bedeuten hat, was offene Mengen sind, was Cauchy-Folgen sind, etc. Ist $(X, \|.\|)$ vollständig, d.h. ist jede Cauchy-Folge konvergent, so spricht man von einem *Banachraum*.

3.5.3 Beispiel.

(i) Betrachte den $\mathbb{R}^n$ mit einer der Normen

$$\|(x_1,\ldots,x_n)\|_\infty := \max_{i=1,\ldots,n} |x_i|,$$

$$\|(x_1,\ldots,x_n)\|_2 := \left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}},$$

$$\|(x_2,\ldots,x_n)\|_1 := \sum_{i=1}^n |x_i|.$$

Die zu diesen Normen gehörigen Metriken sind gerade die uns schon bekannten d_∞, d_2 bzw. d_1. Es gilt

$$\max_{i=1,\ldots,n} |x_i| \leq (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}} \leq \sum_{i=1}^n |x_i| \leq n \cdot \max_{i=1,\ldots,n} |x_i|\,.$$

Ist eine Folge $(a_n)_{n\in\mathbb{N}}$, $a_n \in \mathbb{R}^n$, konvergent gegen $a \in \mathbb{R}^n$ bezüglich einer dieser Normen, so auch bezüglich der anderen. Genauso haben diese Normen die gleichen Cauchy-Folgen. Wir sehen also, dass $\mathbb{R}^n$ bezüglich jeder dieser Normen vollständig ist.

(ii) Betrachte den Raum X aller Folgen $(x_n)_{n\in\mathbb{N}}$ mit $\sum_{i=1}^\infty |x_i|^2 < \infty$. Die Abbildung

$$\|(x_n)_{n\in\mathbb{N}}\|_2 := \left(\sum_{i=1}^\infty |x_i|^2\right)^{\frac{1}{2}}$$

ist eine Norm. Diesen normierten Raum bezeichnet man auch als ℓ^2. Man kann zeigen, dass ℓ^2 vollständig ist.

(iii) Betrachte den Raum $C([a,b])$ aller stetigen Funktionen auf $[a,b]$. Die *Supremumsnorm*

$$\|f\|_\infty = \sup_{t\in[a,b]} |f(t)|$$

induziert die Metrik d_∞ der gleichmäßigen Konvergenz.

(*iv*) Betrachte auf $X = C([a,b])$ eine *gewichtete Supremumsnorm*. Sei $g : [a,b] \to \mathbb{R}$ eine beschränkte Funktion ist, die auf keinem Intervall $(\alpha, \beta) \subseteq [a,b]$ verschwindet (diese Forderung dient dazu im folgenden die Eigenschaft (N1) zu sichern), und definiere

$$\|f\|_{\infty,g} := \sup_{t\in[a,b]} \big(|f(t)| \cdot |g(t)|\big),\ f \in X.$$

Ist $|g(t)| \geq \delta > 0$ für alle $t \in [a,b]$, so gilt

$$\big(\inf_{t\in[a,b]} |g(t)|\big) \cdot \|f\|_\infty \leq \|f\|_{\infty,g} \leq \big(\sup_{t\in[a,b]} |g(t)|\big) \cdot \|f\|_\infty.$$

Dann haben $\|.\|_\infty$ und $\|.\|_{\infty,g}$ also die gleichen konvergenten Folgen, Cauchy-Folgen, etc. Ist $g(t)$ zum Beispiel $g(t) = (t-a)$, so kann eine solche Ungleichungskette nicht mehr gelten. Denn betrachtet man z.B. die Funktion

$$f_n(t) = \begin{cases} 1 - n\,(t-a) & ,\quad t \in [a, a+\frac{1}{n}] \\ 0 & ,\quad t \in [a+\frac{1}{n},\ b] \end{cases}$$

so ist $\|f_n\|_\infty = 1$ aber sicher $\|f_n\|_{\infty,g} \leq \frac{1}{n}$. Weiters wissen wir dass der Raum $(C([a,b]), \|.\|_\infty)$ vollständig ist, man kann aber zeigen dass $(C([a,b]), \|.\|_{\infty,g})$ diese Eigenschaft nicht hat. Dieses Beispiel spielt auch in der Theorie der Differentialgleichungen eine wichtige Rolle.

Ist X eine Menge, $D \subseteq X$, $f : D \to X$ eine Abbildung und $x \in X$, so heißt x ein *Fixpunkt* von f, falls $f(x) = x$ ist. Genauso wie man oft die Nullstellen einer Funktion zu untersuchen hat, hat man auch oft eine Fixpunktgleichung zu lösen. Tatsächlich sind diese Probleme äquivalent, denn eine Lösung ξ von $g(x) = 0$ ist eine Lösung von $f(x) = x$ für die Funktion $f(x) := g(x) + x$, und umgekehrt. Der folgende Satz gibt eine Bedingung für die Existenz eines Fixpunktes für spezielle Abbildungen f, sogenannte Kontraktionen.

3.5.4 Definition. Sei $(X, \|.\|)$ ein normierter Raum, $D \subseteq X$, and $f : D \to X$. Dann heißt f eine *Kontraktion*, wenn eine *Lipschitz-Bedingung*

$$\|f(x) - f(y)\| \leq q\|x - y\|,\ x, y \in D\,,$$

mit einer Konstanten $q < 1$ gilt.

3.5.5 Satz (Fixpunktsatz von Banach)**.** *Sei $(X, \|.\|)$ ein normierter Raum, $D \subseteq X$ sei eine mit der von $\|.\|$ induzierten Metrik vollständige Teilmenge. Ist $f : D \to D$ eine Kontraktion, so existiert genau ein Fixpunkt von f in D.*

Beweis. Sei $y_0 \in D$ und betrachte die Folge $y_n := f^n y_0$. Wir zeigen zunächst mittels Induktion die Ungleichung

$$\|y_{n+1} - y_n\| \leq q^n \|y_1 - y_0\|. \tag{3.7}$$

Für $n = 0$, ist die Beziehung klar. Nun gilt

$$\|y_{n+2} - y_{n+1}\| = \|fy_{n+1} - fy_n\| \leq q\|y_{n+1} - y_n\| \leq$$

$$\leq q^{n+1}\|y_1 - y_0\|,$$

also folgt (3.7) für alle $n \in \mathbb{N}$. Sei nun $n \leq m$ gegeben. Dann folgt

$$\|y_m - y_n\| \leq \|y_m - y_{m-1}\| + \ldots + \|y_{n+1} - y_n\| \leq$$

$$\leq (q^{m-1} + \ldots + q^n)\|y_1 - y_0\| = \left(\sum_{k=n}^{m-1} q^k\right) \|y_1 - y_0\|.$$

Wegen $q < 1$ ist die geometrische Reihe $\sum_{k=0}^{\infty} q^k$ konvergent. Also ist die Folge $(y_n)_{n\in\mathbb{N}}$ bzgl. $\|.\|$ eine Cauchy-Folge. Daher existiert der Grenzwert $\lim_{n\to\infty} y_n$ in D.

Da f einer Lipschitz-Bedingung genügt ist f stetig, d.h. aus $\lim_{n\to\infty} x_n = x$ folgt $\lim_{n\to\infty} fx_n = fx$. Setzt man $y := \lim_{n\to\infty} y_n$, so erhält man also auch

$$fy = f(\lim_{n\to\infty} y_n) = \lim_{n\to\infty} fy_n.$$

Wegen $fy_n = y_{n+1}$ gilt

$$\lim_{n\to\infty} (fy_n - y_n) = \lim_{n\to\infty} (y_{n+1} - y_n) = 0.$$

Insgesamt gilt also $fy = y$.

Die Eindeutigkeit des Fixpunktes folgt unmittelbar aus der Lipschitzbedingung. Denn ist $fy = y$ und $f\tilde{y} = \tilde{y}$, so folgt

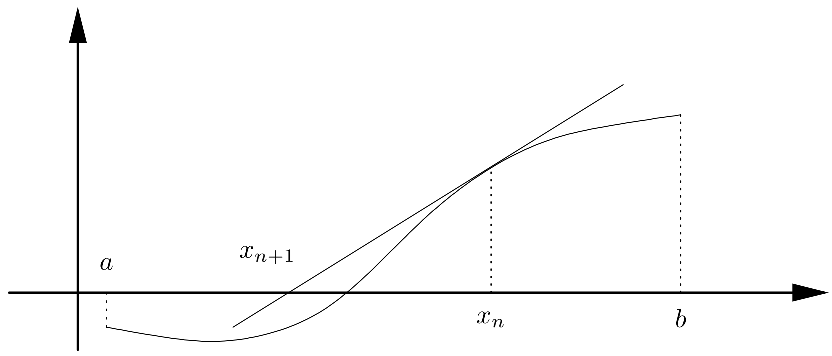
$$\|y - \tilde{y}\| = \|fy - f\tilde{y}\| \leq q\|y - \tilde{y}\|.$$

Wegen $q < 1$ folgt $\|y - \tilde{y}\| = 0$, d.h. $y = \tilde{y}$.

❑

3.5.6 Bemerkung. Das im Beweis angegebene Verfahren den Fixpunkt y von f zu berechnen, heißt auch Verfahren der *sukzessiven Approximation*: Startet man mit $y_0 \in D$, und definiert man rekursiv $y_{n+1} := fy_n$, so gilt $\lim_{n\to\infty} y_n = y$ und die Konvergenzgüte kann durch $\|y_n - y\| \leq \frac{\|y_1 - y_0\|}{1-q} \cdot q^n$ abgeschätzt werden.

Als ein Beispiel wo der Fixpunktsatzes von Banach gewinnbringend eingesetzt werden kann, wollen wir ein Näherungsverfahren zur Nullstellenberechnung angeben, das *Newton-Verfahren* (oder *Tangentenverfahren*). Hat man eine Funktion f gegeben, und sucht man eine Nullstelle, so geht man von der folgenden Idee aus: Man wähle zunächst einen Punkt x_0 der eine erste Näherung der gesuchten Nullstelle sein soll, also schon möglichst nahe an ihr liegen soll. Die Funktion f wird in der Nähe des Punktes x_0 durch die Tangente in diesem Punkt approximiert. Eine zweite, hoffentlich bessere, Näherung für die gesuchte Nullstelle könnte also die Nullstelle der Tangente sein, und diese lässt sich ja leicht berechnen. Mit dieser zweiten Näherung verfährt man genauso weiter, die Kurve wird durch die Tangente approximiert und die Nullstelle der Tangente stellt die nächste Näherung dar.



Da die Gleichung der Tangente an f in einem Punkt a die Gestalt $y = f(a) + f'(a)(x-a)$ hat, schreibt sich dieses Verfahren in Formeln wie folgt an.

3.5.7 Definition. Sei $f : [a,b] \to \mathbb{R}$ stetig auf $[a,b]$ und differenzierbar in (a,b), $f'(x) \neq 0$ für $x \in (a,b)$. Weiters sei $x_0 \in (a,b)$ gegeben. Definiere eine (endliche oder unendliche) Folge $(x_n)_n$ rekursiv durch

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

Dabei bricht das Verfahren ab, wenn für ein n der Punkt x_n nicht in (a,b) liegt.

Natürlich ist es von vornherein überhaupt nicht klar ob dieses Verfahren einmal abbricht oder, falls es nicht abbricht, ob es zum Ziel führt, ob also die Werte x_n zu einer Nullstelle von f konvergieren. Der Fixpunktsatz von Banach impliziert nun unmittelbar:

3.5.8 Korollar. *Sei $f : [a,b] \to \mathbb{R}$ stetig, zweimal differenzierbar auf (a,b), sei f' noch stetig auf $[a,b]$ und sei $f'(x) \neq 0$, $x \in [a,b]$. Weiters sei für $x \in [a,b]$ stets auch $x - \frac{f(x)}{f'(x)} \in [a,b]$, und es gelte*

$$\gamma := \max_{x \in (a,b)} \left| \frac{f(x) \cdot f''(x)}{(f'(x))^2} \right| < 1 \,.$$

Dann gilt: Es hat f auf $[a,b]$ genau eine Nullstelle ξ. Ist $x_0 \in [a,b]$ so konvergiert die durch das Newton-Verfahren definierte Folge $(x_n)_{n\in\mathbb{N}}$ gegen ξ. Die Konvergenzgeschwindigkeit kann dabei abgeschätzt werden durch

$$|x_n - \xi| \leq \gamma^n |x_0 - \xi|, \ n \in \mathbb{N} \,.$$

Beweis. Die Abbildung $g : x \mapsto x - \frac{f(x)}{f'(x)}$ bildet das Intervall $[a,b]$ in sich ab. Sie ist stetig auf $[a,b]$ und differenzierbar in (a,b). Nach dem Mittelwertsatz existiert für alle $x,y \in [a,b]$, $x < y$ eine Stelle $x < t < y$ mit $g(x) - g(y) = g'(t)(x-y)$. Nun ist

$$g'(t) = 1 - \frac{f'(t)^2 - f(t)f''(t)}{f'(t)^2} = \frac{f(t) \cdot f''(t)}{(f'(t))^2} \,.$$

Es folgt das für je zwei $x,y \in [a,b]$

$$|g(x) - g(y)| \leq \gamma |x-y|$$

gilt, d.h. g ist eine Kontraktion.

❑

Für das Newton-Verfahren kann man durch explizite Betrachtung eine wesentlich bessere Aussage über die Konvergenzgeschwindigkeit erzielen.

3.5.9 Satz. *Sei $f : [a,b] \to \mathbb{R}$ stetig, zweimal differenzierbar auf (a,b), sei f' noch stetig auf $[a,b]$ und sei $f'(x) \neq 0$, $x \in [a,b]$. Weiters sei $x_0 \in [a,b]$ so dass das Newton-Verfahren nicht abbricht, und sei $\xi \in (a,b)$ eine Nullstelle von f. Setze*

$$M := \frac{\max_{x\in(a,b)} |f'(x)|^2 \cdot \max_{x\in(a,b)} |f''(x)|}{2\min_{x\in(a,b)} |f'(x)|^3} .$$

Dann gilt in jedem Schritt $|x_{n+1} - \xi| \leq M|x_n - \xi|^2$, und damit allgemein

$$|x_n - \xi| \leq M^{2^n-1}|x_0 - \xi|^{2^n}, \ n \in \mathbb{N} . \tag{3.8}$$

Beweis. Wir wenden die Taylorsche Formel an und erhalten für eine geeignete Zwischenstelle t_1

$$f(x_{n+1}) = \underbrace{f(x_n) + f'(x_n)(x_{n+1} - x_n)}_{=0} + f''(t_1)\frac{(x_{n+1} - x_n)^2}{2} =$$

$$= f''(t_1)\frac{1}{2}\Big(\frac{f(x_n)}{f'(x_n)}\Big)^2 .$$

Weiters ist, da $f(\xi) = 0$ ist,

$$f(x_n) = f'(t_2)(x_n - \xi),\ f(x_{n+1}) = f'(t_3)(x_{n+1} - \xi) ,$$

mit gewissen Zwischenstellen t_2, t_3. Es folgt

$$x_{n+1} - \xi = \frac{f(x_{n+1})}{f'(t_3)} = \frac{1}{2}\Big(\frac{f(x_n)}{f'(x_n)}\Big)^2 \frac{f''(t_1)}{f'(t_3)} =$$

$$= \frac{1}{2}\Big(\frac{f'(t_2)}{f'(x_n)}\Big)^2 \frac{f''(t_1)}{f'(t_3)} \cdot (x_n - \xi)^2 .$$

Wir schliessen, dass

$$|x_{n+1} - \xi| \leq M|x_n - \xi|^2 .$$

Die Formel (3.8) ergibt sich nun mittels vollständiger Induktion.

❑

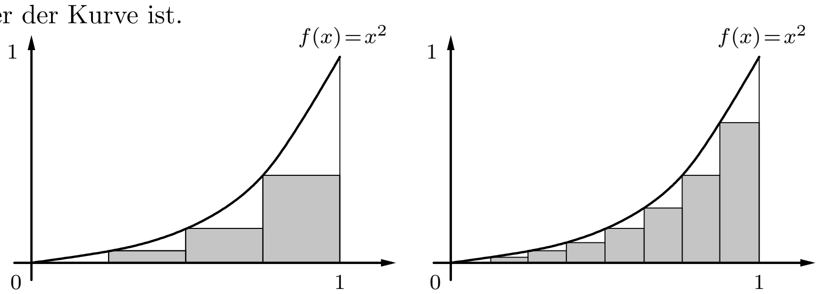
Kapitel 4

Integralrechnung

4.1 Flächenberechnungen

Es war schon in der Mathematik des Altertums eine verbreitete Problemstellung die Fläche gewisser Figuren zu berechnen. Bei Polygonen ist dies durch Zerlegung in Dreiecke unmittelbar möglich, bei krummlinigen Figuren ist dagegen nicht einmal so klar was „Fläche“ überhaupt bedeuten soll.

4.1.1 Beispiel. Betrachte die Parabel gegeben durch $f(x) = x^2$. Wir sind interessiert an der Fläche die von der x-Achse, der Parabel und der Geraden $x = 1$ begrenzt wird. Dazu könnte man, unserer intuitiven Vorstellung von „Fläche“ folgend, wie folgt vorgehen: Man zerlegt die Fläche in Streifen der Breite $\triangle x$, z.B. $\triangle x = \frac{1}{n}$, und approximiert die Fläche eines Streifens durch das Rechteck mit Breite $\triangle x$ und Höhe $\min f(x)$ wobei das Minimum über die im betrachteten Streifen liegenden x-Koordinaten genommen wird. Ist $\triangle x$ sehr klein, so wird man hoffen daß die Fläche des Rechtecks fast gleich jener des Streifens unter der Kurve ist.



In unserem Fall erhält man in Formeln also die folgende Näherung für die Gesamtfläche:

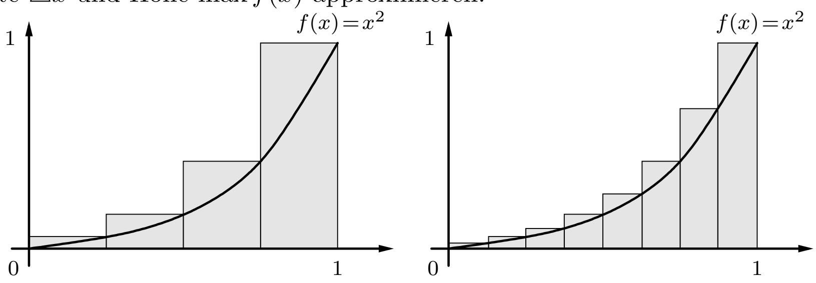
$$A_n = \sum_{k=1}^{n} \min\{f(x) : (k-1)\frac{1}{n} \le f(x) \le k\frac{1}{n}\} \cdot \frac{1}{n} =$$

$$= \sum_{k=1}^{n} \left(\frac{k-1}{n}\right)^2 \cdot \frac{1}{n} = \frac{1}{n^3} \sum_{k=0}^{n-1} k^2 = \frac{1}{n^3} \frac{(n-1)n(2n-1)}{6}.$$

Läßt man in dieser Formel n immer größer werden, so erhält man

$$A = \lim_{n\to\infty} A_n = \frac{1}{3}$$

Genauso könnte man natürlich die Fläche des Streifens durch das Rechteck mit Breite $\triangle x$ und Höhe $\max f(x)$ approximieren.



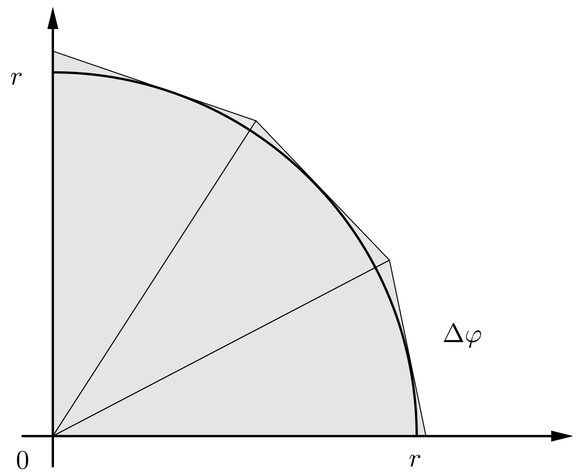
Ist unsere Vorstellung von „Fläche“ vernünftig, so sollte das Selbe herauskommen. Tatsächlich gilt

$$\hat{A}_n = \sum_{k=1}^{n} \max\{f(x) : (k-1)\frac{1}{n} \leq x \leq k\frac{1}{n}\} \cdot \frac{1}{n} = \sum_{k=1}^{n} \frac{k^2}{n^2}\frac{1}{n} =$$

$$= \frac{1}{n^3}\sum_{k=1}^{n} k^2 = \frac{1}{n^3}\frac{n(n+1)(2n+1)}{6},$$

und für $n \to \infty$ erhält man wieder $A = \lim_{n\to\infty} \hat{A}_n = \frac{1}{3}$.

4.1.2 Beispiel. Betrachte einen Viertelkreis mit Radius r und approximiere diesen durch Dreiecke:



Die Fläche eines Dreiecks ist gleich $\frac{r\cdot\triangle\varphi}{2}$, es ergibt sich also

$$A_{\triangle\varphi} = \sum \frac{r\cdot\triangle\varphi}{2} = \frac{r}{2}\cdot\sum\triangle\varphi.$$

Wählt man $\triangle\varphi$ sehr klein, so wird die Summe $\sum\triangle\varphi$ ungefähr gleich der Länge l des Viertelkreisbogens sein. Man erhält also für die Fläche A des Viertelkreises

$$A = \frac{r}{2}l.$$

Bezeichnet also F die Fläche, U den Umfang und r den Radius eines Kreises, so gilt

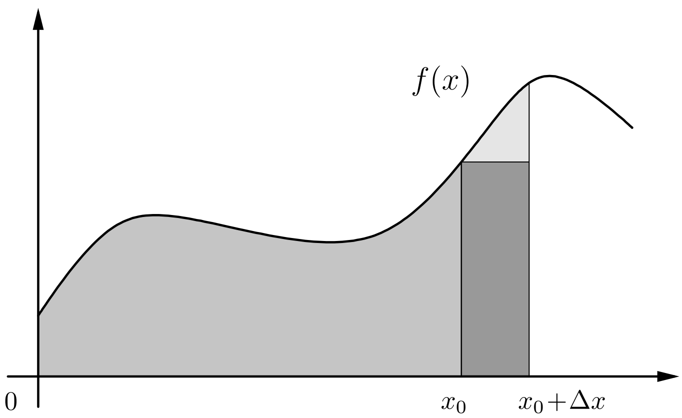
$$F : U = \frac{r}{2}.$$

Solche Flächenberechnungen waren schon zu Zeiten von Archimedes bekannt. Man sieht jedoch, daß man sich für jede Kurve einen speziellen Trick einfallen lassen mußte. Der wesentliche Beitrag der Mathematiker des 17. Jahrhunderts (Leibnitz, Newton, aber auch andere) war nun:

(i) Die „Zerlegung einer Fläche in Rechtecke“ zu formalisieren, was auf den Integralbegriff führt und gleichermaßen eine (vorläufige) Definition des Begriffs der „Fläche“ erlaubt.

(ii) Zu erkennen, daß die Bildung des Integrals in gewissem Sinne die Umkehroperation des Differenzierens ist (vgl. Satz 4.4.1) und damit eine weitläufige Anwendbarkeit des Integrals zu gewährleisten. Denn damit ist es möglich Integrale tatsächlich systematisch auszurechnen.

4.1.3 Beispiel. Betrachte eine Kurve gegeben durch $y = f(x)$:



Bezeichne die Fläche die von den Koordinatenachsen, der Kurve $f(x)$ und der Geraden $x = x_0$ begrenzt wird mit $F(x_0)$. Der Zuwachs der Fläche F wenn man sogar bis $x_0 + \triangle x$ geht wird in erster Näherung ungefähr gleich der Fläche des dunklen Rechtecks sein:

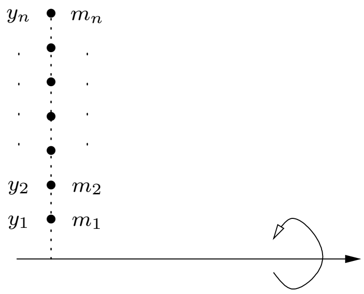
$$F(x_0 + \triangle x) - F(x_0) = f(x_0) \cdot \triangle x.$$

Läßt man hier $\triangle x$ immer kleiner werden, so erhält man

$$F'(x_0) = f(x_0).$$

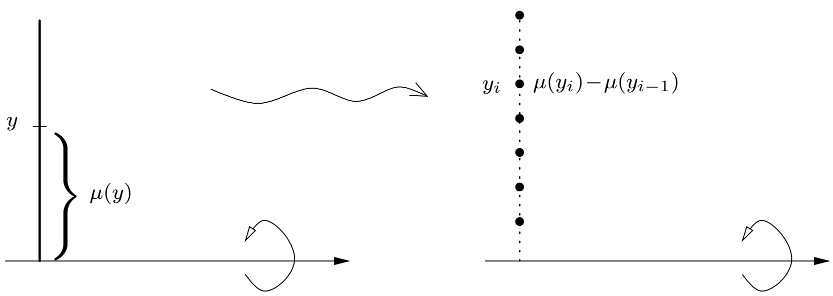
4.1.4 Beispiel. Wir berechnen das Drehmoment eines Stabes mit nichthomogener Massenverteilung.

Das Drehmoment eines einzelnen Massenpunktes mit Masse m der sich in einem Abstand y von der Drehachse befindet, ist gleich my. Betrachte nun ein System endlich vieler Massenpunkte m_i, $i = 1, \dots, n$, in Abständen y_i, $i = 1, \dots, n$:



Dann erhält man für das Drehmoment des gesamten Systems $\sum_{i=1}^{n} m_i y_i$.

Um nun das Drehmoment eines beliebigen Stabes zu berechnen, gehen wir von der folgenden Vorstellung aus: Die Masse des Stabes von 0 bis zu einer Höhe y bezeichne mit $\mu(y)$. Wir zerlegen den Stab in kleine Abschnitte, von $y_0 = 0$ bis y_1, von y_1 bis y_2 usw. und denken uns die Masse des kleinen Abschnittes von y_{i-1} bis y_i im Punkt y_i konzentriert.



Als Näherung für das Moment des Stabes ergibt sich also

$$\sum_{i=1}^{n} y_i \big(\mu(y_i) - \mu(y_{i-1})\big),$$

und wir werden hoffen das, wenn die Zerlung in kleine Abschnitte immer feiner wird, diese Näherung das wahre Moment des Stabes immer besser approximiert.

Sollte zufällig die Massenverteilung des Stabes homogen sein, d.h. $\mu(y) = y$, so passiert genau das Gleiche wie schon bei der Berechnung der Fläche der Parabel. Sonst ist dieser Zugang allgemeiner.

4.2 Definition des Riemannschen Integrals

Betrachte ein Intervall $[a, b]$. Eine *Partition* dieses Intervalles ist eine endliche Menge $P = \{x_0, \dots, x_n\}$, sodass

$$a = x_0 < x_1 < \ldots < x_n = b$$

gilt. Die Menge aller Partitionen von $[a, b]$ bezeichne mit $\mathcal{P}[a, b]$.

Sind $P, P^* \in \mathcal{P}[a, b]$, so heißt P^* eine *Verfeinerung* von P, falls $P \subseteq P^*$, d.h. falls jeder Punkt x_i von P auch in P^* vorkommt. Die Relation $\subseteq$ auf $\mathcal{P}[a, b]$ ist klarerweise reflexiv und transitiv, tatsächlich ist sie sogar eine Ordnungsrelation. Sind P_1, P_2 zwei Partitionen, so gibt es stets eine gemeinsame Verfeinerung, d.h. eine Partition P^* mit $P_1 \subseteq P^*$ und $P_2 \subseteq P^*$. Zum Beispiel kann man $P^* = P_1 \cup P_2$ wählen.

Sei nun eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ gegeben. Ist $P \in \mathcal{P}[a, b]$ eine Partition, so definiert man ($\triangle x_i = x_i - x_{i-1}, i = 1, \dots, n$)

$$U(P, f) := \sum_{i=1}^{n} \big(\sup_{x_{i-1} \le \ x \le \ x_i} f(x)\big) \triangle x_i \,,$$

$$L(P, f) := \sum_{i=1}^{n} \big(\inf_{x_{i=1} \le \ x \le \ x_i} f(x)\big) \triangle x_i \,.$$

die *Riemannsche*[1] *Obersumme* bzw. *Untersumme* der Funktion f zur Partition P.

4.2.1 Lemma. *Sei f eine beschränkte Funktion auf $[a, b]$, und seien P, P^* Partitionen von $[a, b]$. Ist $P^* \supseteq P$, so gilt*

$$L(P, f) \le L(P^*, f),$$

und

$$U(P^*, f) \le U(P, f). \tag{4.1}$$

Weiters gilt stets $L(P, f) \le U(P, f)$.

Beweis. Betrachte zunächst den Fall, daß P^* nur einen Punkt mehr hat als P, $P^* \setminus P = \{x^*\}$, und sei $x_{j-1} \le x^* \le x_j$. Dann gilt

$$L(P, f) = \sum \big(\inf_{x_{i-1} \le x \le x_i} \ f(x)\big) \triangle x_i =$$

$$= \sum_{x_i < x_{j-1}} \big(\inf_{x_{i-1} \le x \le x_i} \ f(x)\big) \triangle x_i + \big(\inf_{x_{j-1} \le x \le x_j} \ f(x)\big) (x_j - x_{j-1}) +$$

$$+ \sum_{x_i > x_j} \big(\inf_{x_{i-1} \le x \le x_i} \ f(x)\big) \triangle x_i \le \sum_{x_i < x_{j-1}} \big(\inf_{x_{i-1} \le x \le x_i} \ f(x)\big) \triangle x_i +$$

$$+ \big(\inf_{x_{j-1} \le x \le x^*} \ f(x)\big) (x^* - x_{j-1}) + \big(\inf_{x^* \le x \le x_j} \ f(x)\big) (x_j - x^*) +$$

$$+ \sum_{x_i > x_j} \big(\inf_{x_{i-1} \le x \le x_j} \ f(x)\big) \triangle x_i = L(P^*, f).$$

Denn es gilt sicher

$$\inf_{x_{j-1} \le x \le x_j} f(x) \le \ \inf_{x_{j-1} \le x \le x^*} \ f(x), \ \inf_{x^* \le x \le x_j} f(x).$$

Die Beziehung (4.1) folgt genauso aus der Tatsache

$$\sup_{x_{j-1} \le x \le x_j} f(x) \ge \ \sup_{x_{j-1} \le x \le x^*} \ f(x), \ \sup_{x^* \le x \le x_j} \ f(x).$$

Ist nun P^* eine beliebige Verfeinerung von P, so kann man diese durch sukzessives ergänzen einzelner Punkte in endlich vielen Schritten erhalten. Also folgt die Behauptung aus dem bereits bewiesenen.

[1]Bernhard Riemann, 17.9.1826-20.7.1866

Die letzte Aussage folgt, da stets $\inf_{x_{i-1}\le x\le x_i} \le \sup_{x_{i-1}\le x\le x_i}$ gilt.

❑

Betrachte nun die Menge $\{L(P,f) : P \in \mathcal{P}[a,b]\}$. Diese ist sicher nach oben beschränkt, z.B. durch $(a-b)\sup_{x\in[a,b]} f(x)$, besitzt also ein Supremum, $L := \sup_{P\in\mathcal{P}[a,b]} L(P,f)$. Wegen der Monotonieeigenschaft die wir im obigen Lemma gezeigt haben, gilt die folgende Eigenschaft:

$$\forall\epsilon>0\,\exists P_0\in\mathcal{P}[a,b]\,\forall P\supseteq P_0:\ |L(P,f)-L|<\epsilon\,.$$

Aus diesem Grund werden wir im folgenden die Schreibweise $L =: \lim_{P\in\mathcal{P}[a,b]} L(P,f)$ verwenden.

Ganz genauso gilt für $U := \inf_{P\in\mathcal{P}[a,b]} U(P,f)$ dass

$$\forall\epsilon>0\,\exists P_0\in\mathcal{P}[a,b]\,\forall P\supseteq P_0:\ |U(P,f)-U|<\epsilon\,.$$

und wir schreiben $U =: \lim_{P\in\mathcal{P}[a,b]} U(P,f)$.

4.2.2 Lemma. *Es gilt* $\lim_{P\in\mathcal{P}[a,b]} L(P,f) \le \lim_{P\in\mathcal{P}[a,b]} U(P,f)$.

Beweis. Sei $\epsilon>0$ gegeben, dann wähle $P_1,P_2\in\mathcal{P}[a,b]$ sodass

$$\lim_{P\in\mathcal{P}[a,b]} L(P,f)-L(P,f)<\epsilon,\ P\supseteq P_1$$

$$U(P,f)-\lim_{P\in\mathcal{P}[a,b]} U(P,f)<\epsilon,\ P\supseteq P_2\,.$$

Ist P^* eine gemeinsame Verfeinerung von P_1 und P_2, so erhalten wir

$$\lim_{P\in\mathcal{P}[a,b]} L(P,f)-\lim_{P\in\mathcal{P}[a,b]} U(P,f)\le$$

$$\le \lim_{P\in\mathcal{P}[a,b]} L(P,f)\underbrace{-L(P,f)+U(P,f)}_{\ge 0}-\lim_{P\in\mathcal{P}[a,b]} U(P,f)<2\epsilon\,.$$

Da ϵ beliebig war folgt die Behauptung.

❑

4.2.3 Definition. Sei f eine auf $[a,b]$ definierte und beschränkte Funktion. Dann heißt f *Riemann-integrierbar* auf $[a,b]$, falls

$$\lim_{P\in\mathcal{P}[a,b]} L(P,f)=\lim_{P\in\mathcal{P}[a,b]} U(P,f)\,. \qquad (4.2)$$

In diesem Fall schreiben wir $f\in\mathcal{R}$ und bezeichnen den gemeinsamen Wert aus (4.2) als das *Riemann Integral* $\int_a^b f\,dx$ von f über $[a,b]$.

4.2.4 Satz. *Sei* $f:[a,b]\to\mathbb{R}$ *eine beschränkte Funktion. Dann sind äquivalent:*

(*i*) $f\in\mathcal{R}$.

(*ii*) *Zu jedem* $\epsilon>0$ *existiert eine Partition* P_0 *von* $[a,b]$ *mit*

$$U(P^*,f)-L(P^*,f)<\epsilon,\ P^*\supseteq P_0\,.$$

(iii) Zu jedem $\epsilon > 0$ existiert eine Partition P_0 von $[a, b]$ mit

$$U(P_0, f) - L(P_0, f) < \epsilon\,.$$

Beweis. Wir zeigen $(i) \Rightarrow (ii)$. Sei $f \in \mathcal{R}$ und sei $\epsilon > 0$ gegeben. Wähle P_1 und P_2 sodass

$$\lim_{P\in\mathcal{P}[a,b]} L(P, f) - L(P, f) < \frac{\epsilon}{2},\ P \supseteq P_1$$

$$U(P, f) - \lim_{P\in\mathcal{P}[a,b]} U(P, f) < \frac{\epsilon}{2},\ P \supseteq P_2\,,$$

und setze $P^* := P_1 \cup P_2$.

Die Implikation $(ii) \Rightarrow (iii)$ ist trivial. Sei (iii) erfüllt. Sei $\epsilon > 0$ und wähle eine Partition P sodass $U(P, f) - L(P, f) < \epsilon$ gilt. Es ist

$$L(P_0, f) \le \lim_{P\in\mathcal{P}[a,b]} L(P, f) \le \lim_{P\in\mathcal{P}[a,b]} U(P_0, f) \le U(P, f)\,,$$

Und daher

$$\lim_{P\in\mathcal{P}[a,b]} U(P, f) - \lim_{P\in\mathcal{P}[a,b]} L(P, f) \le U(P_0, f) - L(P_0, f) < \epsilon\,.$$

Da ϵ beliebig war folgt $\lim_{P\in\mathcal{P}[a,b]} U(P, f) = \lim_{P\in\mathcal{P}[a,b]} L(P, f)$.

❑

4.2.5 Beispiel. Betrachte die Funktion $f(x) := x^2$, $x \in [0, 1]$. Wie wir in der Rechnung in Beispiel 4.1.1 gesehen haben, gibt es zu $\epsilon > 0$ ein Partitionen P mit $U(P, f) - \frac{1}{3} < \epsilon$ und $\frac{1}{3} - L(P, f) < \epsilon$. Also folgt mit Hilfe von Satz 4.2.4, $(iii) \Rightarrow (i)$, dass f Riemann-integrierbar ist. Man beachte, dass die Rechnung aus Beispiel 4.1.1 alleine dieses Ergebniss nicht gegeben hätte, da wir dort ja nur für eine ganz spezielle Folge von Partitionen (nämlich die mit äquidistanten Teilungspunkten) gezeigt haben dass $U(P_n, f) \to \frac{1}{3}$ und $L(P_n, f) \to \frac{1}{3}$. In der Definition der Riemann-integrierbarkeit wird aber verlangt dass man dieses für den Limes zeigt der sich über alle (!) Partitionen erstreckt.

4.2.6 Beispiel. Betrachte die Funktion $f : [0, 1] \to \mathbb{R}$ definiert als

$$f(x) := \begin{cases} 0 & ,\ x \text{ irrational} \\ 1 & ,\ x \text{ rational} \end{cases}.$$

Sei $P = \{x_0, \ldots, x_n\}$ eine Partition von $[0, 1]$. Da in jedem Intervall sowohl rationale wie auch irrationale Zahlen liegen gilt stets

$$\sup_{x_{i-1}\le x\le x_i} f(x) = 1, \quad \inf_{x_{i-1}\le x\le x_i} f(x) = 0\,.$$

Es folgt

$$L(P, f) = 0, U(P, f) = 1.$$

Also ist $f \notin \mathcal{R}$.

4.2.7 Beispiel. Betrachte die Funktion $f : [0,1] \to \mathbb{R}$ definiert als

$$f(x) := \begin{cases} 0 & , \ x \text{ irrational, oder } x = 0 \\ \frac{1}{q} & , \ x = \frac{p}{q}, p, q \text{ teilerfremd} \end{cases} .$$

Sei $P = \{x_0, \dots, x_n\}$ eine Partition von $[0,1]$. Da in jedem Intervall irrationale Zahlen liegen, gilt wieder $L(P,f) = 0$.

Sei $\epsilon > 0$ gegeben. Wähle $N \in \mathbb{N}$, sodass $\frac{1}{N} < \frac{\epsilon}{2}$. Weiters wähle $M \in \mathbb{N}$, sodass $\frac{1}{M} < \frac{\epsilon}{4 \cdot N!}$. Betrachte die Partition $P = \{0, \frac{1}{M}, \frac{2}{M}, \dots, \frac{M-1}{M}, 1\}$. Es gibt höchstens $N!$ rationale Zahlen in $[0,1]$ sodass $x = \frac{p}{q}$ mit $q < N$. Also können auch nur höchstens $2N!$ viele der Intervalle $[x_{i-1}, x_i]$ eine solche Zahl enthalten. Solche wollen wir vom Typ (a) bezeichnen. Alle anderen bezeichnen wir vom Typ (b). Nun gilt für ein Intervall welches vom Typ (b) ist $\sup_{x_{i-1} \le x \le x_i} f(x) \le \frac{1}{N}$. Es folgt

$$U(P,f) = \sum_{i=1}^{M} (\sup_{x_{i-1} \le x \le x_i} f(x)) \frac{1}{M} \le \sum_{\text{Typ (b)}} \frac{1}{N} \frac{1}{M} + \sum_{\text{Typ (a)}} 1 \frac{1}{M} \le$$

$$\le \frac{1}{N} M \frac{1}{M} + 2N! \frac{1}{M} < \epsilon .$$

Nach Satz 4.2.4 ist $f \in \mathcal{R}$ und es gilt $\int_a^b f \, dx = 0$.

4.2.8 Beispiel. Betrachte die Funktion $f : [0,1] \to \mathbb{R}$ definiert als

$$f(x) := \begin{cases} 1 & , \ x \in \{\frac{1}{n} : n \in \mathbb{N}\} \\ 0 & , \ \text{sonst} \end{cases} .$$

Sei $P = \{x_0, \dots, x_n\}$ eine Partition von $[0,1]$. Da in jedem Intervall Zahlen liegen die nicht von der Gestalt $\frac{1}{n}$ sind, gilt wieder $L(P,f) = 0$.

Wir zeigen, dass f Riemann-integrierbar ist. Dazu genügt es nach Satz 4.2.4 zu jedem $\epsilon > 0$ eine Partition P finden mit $U(P,f) - L(P,f) < \epsilon$. Sei uns also $\epsilon > 0$ gegeben. Wähle $\delta > 0$ sodaß

$$\delta < \frac{1}{3} \min \big\{ \frac{\epsilon}{2N}, \frac{1}{N} - \frac{\epsilon}{2}, \frac{1}{N-1} - \frac{1}{N}, \dots, 1 - \frac{1}{2} \big\} .$$

Durch diese Wahl von δ haben wir erreicht, dass die Intervalle

$$[0, \epsilon], [\frac{1}{N} - \delta, \frac{1}{N} + \delta], [\frac{1}{N-1} - \delta, \frac{1}{N-1} + \delta], \dots, [1 - \delta, 1] \tag{4.3}$$

sich nicht überlappen. Betrachte die Partition $P = \{x_0, \dots, x_n\}$ die aus genau den als Randpunkte dieser Intervalle auftretenden Punkten besteht. Ein Intervall $[x_{i-1}, x_i]$ nennen wir vom Typ (a), wenn es in (4.3) auftritt, andernfalls heißt es vom Typ (b). Dann gilt

$$U(P,f) = \sum_{i=1}^{n} \sup_{x \in [x_{i-1}, x_i]} f(x) \cdot (x_i - x_{i-1}) =$$

$$= \sum_{[x_{i-1}, x_i] \ \text{Typ } (a)} \sup_{x \in [x_{i-1}, x_i]} f(x) \cdot (x_i - x_{i-1}) +$$

$$+ \sum_{[x_{i-1},x_i] \text{ Typ } (b)} \sup_{x\in[x_{i-1},x_i]} f(x)\cdot(x_i - x_{i-1})\,.$$

Die zweite Summe ist gleich Null, denn ausserhalb der Intervalle (4.3) ist f stets gleich Null. Die erste Summe können wir abschätzen durch das Supremum von f am ganzen Intervall mal Gesamtlänge aller Intervalle vom Typ (a). Nach unserer Wahl von δ erhalten wir dann

$$\sum_{[x_{i-1},x_i] \text{ Typ } (a)} \sup_{x\in[x_{i-1},x_i]} f(x)\cdot(x_i - x_{i-1}) \le 1\cdot(\frac{\epsilon}{2} + (N-1)2\delta + \delta) < \epsilon\,.$$

Die Frage, ob eine gegebene Funktion f tatsächlich Riemann-integrierbar ist oder nicht, ist nicht so einfach zu beantworten. Das Kriterium aus Satz 4.2.4 ist ja nicht wirklich praktisch. Im folgenden geben wir zwei hinreichende Bedingungen für $f \in \mathcal{R}$ an.

4.2.9 Satz. *Ist f stetig auf $[a,b]$ oder ist f monoton auf $[a,b]$, dann ist $f \in \mathcal{R}$.*

Beweis. Wir betrachten zuerst den Fall dass f stetig ist. Sei $\epsilon > 0$ gegeben. Wähle $\eta > 0$ sodaß

$$\eta < \frac{\epsilon}{b-a},$$

d.h. $(b-a)\eta < \epsilon$. Da f stetig ist und $[a,b]$ kompakt, ist f sogar gleichmäßig stetig. Es existiert also $\delta > 0$ mit

$$|f(x) - f(t)| < \eta \text{ für alle } x,t \in [a,b] \text{ mit } |x-t| < \delta\,.$$

Sei $P = \{x_0,\dots,x_n\}$ die Partition $x_0 = a, x_1 = x_0 + \frac{\delta}{2},\dots,x_n = b$ von $[a,b]$. Dann gilt für $x,t \in [x_{i-1},x_i]$ sicher

$$|f(x) - f(t)| < \eta,$$

also ist

$$\left[\sup_{x_{i-1}\le x\le x_i} f(x)\right] - \left[\inf_{x_{i-1}\le x\le x_i} f(x)\right] \le \eta.$$

Es folgt

$$U(P,f) - L(P,f) = \sum_{i=1}^{n}\left[\sup_{x_{i-1}\le x\le x_i} f(x) \;-\; \inf_{x_{i-1}\le x\le x_i} f(x)\right]\triangle x_i \le$$

$$\le \eta\sum_{1=1}^{n}\triangle x_i = \eta(b-a) < \epsilon.$$

Wegen Satz 4.2.4 folgt also $f \in \mathcal{R}$.

Betrachte nun den Fall dass f monoton wachsend ist. Ist f monoton fallend, so geht man genauso vor. Sei wieder $\epsilon > 0$ gegeben. Wähle $n \in \mathbb{N}$ sodass

$$\frac{b-a}{n}(f(b) - f(a)) < \epsilon\,,$$

und betrachte die Partition $P = \{a, a+\frac{b-a}{n}, a+2\frac{b-a}{n},\dots,a+(n-1)\frac{b-a}{n}, b\}$. Da f monoton ist gilt

$$\sup_{x_{i-1}\le x\le x_i} f(x) = f(x_i), \qquad \inf_{x_{i-1}\le x\le x_i} f(x) = f(x_{i-1}).$$

Man berechnet

$$U(P,f) - L(P,f) = \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] \, \frac{b-a}{n} =$$

$$= \frac{b-a}{n} \sum_{i=1}^{n} [f(x_i) - f(x_{i-1})] = \frac{b-a}{n} \, [f(b) - f(a)] < \epsilon \, .$$

Wir erhalten, wieder aus Satz 4.2.4, daß $f \in \mathcal{R}$

❑

Genauso wie wir in Beispiel 4.2.8 vorgegangen sind, können wir nun vorgehen um diesen Satz noch ein bischen weiter zu verallgemeinern. Dazu definieren wir: Eine Funktion $f : [a,b] \to \mathbb{R}$ heißt *stückweise stetig*, wenn es eine Partition $P = \{c_0, c_1, \ldots, c_k\}$ von $[a,b]$ gibt sodaß $f|_{(c_i, c_{i-1})}$ für jedes $i = 1, \ldots, k$ stetig ist. Wir erhalten nun aus Satz 4.2.9:

4.2.10 Korollar. *Sei $f : [a,b] \to \mathbb{R}$ beschränkt und stückweise stetig. Dann ist $f \in \mathcal{R}$.*

Beweis. Die Funktion f ist beschränkt, d.h. es existiert eine Zahl M mit $|f(x)| \leq M$, $x \in [a,b]$. Sei uns nun $\epsilon > 0$ gegeben. Wir müssen eine Partition finden mit $U(P,f) - L(P,f) < \epsilon$. Dazu wähle $\delta > 0$ sodaß

$$\delta < \frac{1}{3} \min \Big\{ \frac{\epsilon}{4M(k+1)}, c_1 - c_0, \ldots, c_k - c_{k-1} \Big\} \, .$$

Dann haben erreicht dass die Intervalle

$$[c_0, c_0 + \delta], [c_1 - \delta, c_1 + \delta], \ldots, [c_k - \delta, c_k] \tag{4.4}$$

sich nicht überlappen und das ihre Gesamtlänge kleiner als $\frac{\epsilon}{4M}$ ist. Die Partition deren Elemente genau die in den Intervallen (4.4) auftretenden Randpunkte sind, bezeichne mit P_0.

Auf jedem der Intervalle $[c_{i-1} + \delta, c_i - \delta]$, $i = 1, \ldots k$, ist die Funktion f stetig. Nach Satz 4.2.9 angewandt auf $f|_{[c_{i-1}+\delta, c_i - \delta]}$ erhalten wir eine Partition P_i von $[c_{i-1} + \delta, c_i - \delta]$ sodaß

$$U\big(P_i, f|_{[c_{i-1}+\delta, c_i-\delta]}\big) - L\big(P_i, f|_{[c_{i-1}+\delta, c_i-\delta]}\big) < \frac{\epsilon}{2k} \, .$$

Sei $P := P_0 \cup P_1 \cup \ldots \cup P_k$. Dann ist $P = \{x_0, x_1, \ldots, x_n\}$ eine Partition von $[a,b]$. Wir nennen ein Intervall $[x_{i-1}, x_i]$ vom Typ (0) wenn es in (4.4) vorkommt, wir nennen es vom Typ (l), $l = 1, \ldots, k$, wenn es in $[c_{l-1} + \delta, c_l - \delta]$ liegt, also von der Partition P_l kommt. Dann können wir schreiben

$$U(P,f) - L(P,f) = \sum_{i=1}^{n} \big(\sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x) \big)(x_i - x_{i-1}) =$$

$$= \sum_{[x_{i-1}, x_i] \text{ Typ (0)}} \big(\sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x) \big)(x_i - x_{i-1}) +$$

$$+\sum_{l=1}^{k} \sum_{[x_{i-1},x_i] \text{ Typ } (l)} \big(\sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x)\big)(x_i - x_{i-1}) =$$

$$= \sum_{[x_{i-1},x_i] \text{ Typ } (0)} \big(\sup_{x\in[x_{i-1},x_i]} f(x) - \inf_{x\in[x_{i-1},x_i]} f(x)\big)(x_i - x_{i-1}) +$$

$$+\sum_{l=1}^{k} U\big(P_i, f|_{[c_{i-1}+\delta, c_i-\delta]}\big) - L\big(P_i, f|_{[c_{i-1}+\delta, c_i-\delta]}\big) .$$

Die Summe über die Intervalle vom Typ (0) können wir abschätzen mit $2M \cdot \frac{\epsilon}{4M}$, jeden der anderen Summanden mit $\frac{\epsilon}{2k}$. Insgesamt folgt $U(P,f) - L(P,f) < \epsilon$.

❑

4.2.11 Bemerkung. Die Riemann-integrierbarkeit einer Funktion ist tatsächlich sehr eng mit der Stetigkeit verküpft. Tatsächlich gilt (ohne Beweis): Sei $f : [a,b] \to \mathbb{R}$ beschränkt. Dann ist $f \in \mathcal{R}$, genau dann wenn die Menge der Unstetigkeitsstellen von f eine Lebesgue-Nullmenge ist. Dabei heißt eine Menge $A \subseteq \mathbb{R}$ Lebesgue-Nullmenge, wenn es zu jedem $\epsilon > 0$ eine Folge Intervallen gibt deren Gesamtlänge kleiner als ϵ ist und die A überdecken.

Manchmal ist es praktisch zur Berechnung von Integralen auch *Riemannsche Zwischensummen* heranzuziehen: Sei $P = \{x_0, \ldots, x_n\} \in \mathcal{P}[a,b]$ und sei Zwischenstellen $Z = \{t_1, \ldots, t_n\}$ gegeben, d.h. seien die Punkte t_i sodass

$$a = x_0 \le t_1 \le x_1 \le t_2 \le x_2 \le \cdots \le x_{n-1} \le t_n \le x_n = b .$$

Dann definieren wir

$$S(f,P,Z) := \sum_{i=1}^{n} f(t_i) \triangle x_i .$$

Der folgende Satz besagt, dass die Eigenschaft Riemann-integrierbar zu sein bedeutet dass der Limes $\lim_{p\in\mathcal{P}[a,b]} S(f,P,Z)$ unabhängig von der Wahl der Zwischenstellen, und noch dazu in einem stärkeren Sinne als in dem der verallgemeinerten Folgen, existiert. Ist P eine Partition, bezeichne mit $\mu(P)$ die Zahl

$$\mu(P) := \max \triangle x_i .$$

4.2.12 Satz. *Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Dann sind äquivalent:*

(*i*) *Es ist $f \in \mathcal{R}$, d.h. die Limiten $\lim_{P\in\mathcal{P}[a,b]} U(P,f)$ und $\lim_{P\in\mathcal{P}[a,b]} L(P,f)$ stimmen überein.*

(*ii*) *Der Limes $\lim_{P\in\mathcal{P}[a,b]} S(f,P,Z)$ existiert gleichmäßig bezüglich der Wahl von Z, d.h. es gibt eine Zahl $A \in \mathbb{R}$ sodaß für jedes $\epsilon > 0$ ein $P_0 \in \mathcal{P}[a,b]$ existiert mit der Eigenschaft dass für alle $P \in \mathcal{P}[a,b]$, $P \supseteq P_0$, und alle Zwischenstellenmengen Z gilt $|S(f,P,Z) - A| \le \epsilon$.*

(*iii*) *Für jedes $\epsilon > 0$ gibt es $\delta > 0$ mit*

$$U(P,f) - L(P,f) \le \epsilon, \text{ für alle } P \text{ mit } \mu(P) < \delta .$$

(*iv*) *Für jedes* $\epsilon > 0$ *gibt es* $\delta > 0$ *sodass für alle* P *mit* $\mu(P) < \delta$ *und alle Zwischenstellenmengen* Z_1, Z_2 *gilt*

$$S(f, P, Z_1) - S(f, P, Z_2) \leq \epsilon\,.$$

(*v*) *Es gibt eine Zahl* $A \in \mathbb{R}$*, sodass es zu jedem* $\epsilon > 0$ *ein* $\delta > 0$ *gibt mit der Eigenschaft dass* $|U(f, P) - A| \leq \epsilon$ *und* $|L(f, P) - A| \leq \epsilon$ *gilt, falls* $\mu(P) < \delta$.

(*vi*) *Es gibt eine Zahl* $A \in \mathbb{R}$*, sodass*

$$\forall \epsilon > 0 \, \exists \delta > 0 : \forall P \in \mathcal{P}[a, b], \mu(P) < \delta \, \forall Z : \, |S(f, P, Z) - A| \leq \epsilon\,. \quad (4.5)$$

In diesem Fall ist $\int_a^b f\,dx = A$.

Beweis. Als erstes bemerken wir, dass für eine Partition P die folgenden äquivalent sind:

(a) Es ist $|S(f, P, Z) - A| \leq \epsilon$ für alle Zwischenstellenmengen Z.

(b) Es ist $|U(P, f) - A| \leq \epsilon$ und $|L(P, f) - A| \leq \epsilon$.

Die Implikation (b) $\Rightarrow$ (a) folgt, da stets $L(P, f) \leq S(f, P, Z) \leq U(P, f)$ gilt. Umgekehrt sei (a) erfüllt, d.h. gelte

$$A - \epsilon \leq S(f, P, Z) \leq A + \epsilon$$

für alle Z. Wegen $\sup_Z S(f, P, Z) = U(P, f)$ und $\inf_Z S(f, P, Z) = L(P, f)$, folgt ($b$).

Damit ist klar dass (i) $\iff$ (ii) und dass (v) $\iff$ (vi). Genau die gleiche Argumentation zeigt dass auch (iii) $\iff$ (iv). Weiters ist klar dass (v) $\Rightarrow$ (iii), mit Hilfe von Satz 4.2.4 folgt auch (iii) $\Rightarrow$ (i).

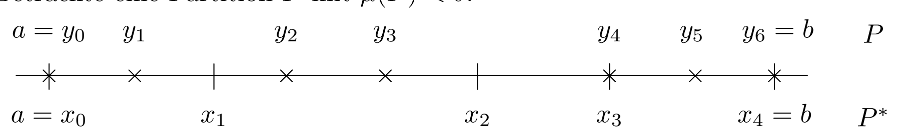
Es bleibt zu zeigen, dass (i) $\Rightarrow$ (vi). Sei also vorausgesetzt dass $f \in \mathcal{R}$ ist, und sei $\epsilon > 0$ gegeben. Wähle eine Partition P^* mit

$$U(P^*, f) \leq \int_a^b f\,dx + \epsilon, \; L(P^*, f) \geq \int_a^b f\,dx - \epsilon\,.$$

Sei $M = \sup_{x\in[a,b]} |f(x)|$, bezeichne mit n die Anzahl der in P^* vorkommenden Punkte, und setze

$$\delta = \frac{\epsilon}{n}\,.$$

Betrachte eine Partition P mit $\mu(P) < \delta$.



Jene Summanden in $U(P, f)$, die zu einem Intervall gehören welches ganz in einem Intervall von P^* enthalten ist, kann man durch den entsprechenden Summanden von $U(P^*, f)$ abschätzen. Die Anzahl jener Intervalle von P die einen Punkt von P^* im Inneren enthalten ist höchstens gleich $n-1$. Also ist ihr Beitrag zu $U(P, f)$ höchstens gleich

$$(n - 1)\mu(P)M < \frac{(n - 1)M\epsilon}{n} < M\epsilon\,.$$

Insgesamt erhält man

$$U(P,f) < U(P^*,f) + M\epsilon < \int_a^b f\,dx + (1+M)\epsilon\,.$$

Genauso geht man vor um zu zeigen dass

$$L(P,f) > \int_a^b f\,dx - (1+M)\epsilon\,.$$

Da klarerweise stets $L(P,f) \le S(f,P,Z) \le U(P,f)$ gilt, folgt für $\mu(P) < \delta$

$$|S(f,P,Z) - \int_a^b f\,dx| < \epsilon(1+M).$$

Also gilt (vi) mit der Zahl $A = \int_a^b f\,dx$.

❑

4.2.13 Bemerkung. Möchte man ein Integral $\int_a^b f(x)\,dx$ ausrechnen, kann es praktisch sein nur einige geschickt gewählte Partitionen zu benützen, so wie zum Beispiel in Beispiel 4.1.1. Dazu überlegt man sich zuerst, dass die Funktion f Riemann-integrierbar ist. Denn dann gilt nach dem letzten Satz: Ist $P_n \in \mathcal{P}[a,b]$ irgendeine Folge von Partitionen mit $\mu(P_n) \to 0$, und Z_n irgendeine Menge von Zwischenstellen, so gilt

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} L(P_n,f) = \lim_{n\to\infty} U(P_n,f) = \lim_{n\to\infty} S(f,P_n,Z_n)\,.$$

4.3 Eigenschaften des Integrals

4.3.1 Satz. *Es gilt*

(i) *Seien $f_1, f_2 \in \mathcal{R}$ auf $[a,b]$ und $c \in \mathbb{R}$. Dann ist $f_1 + f_2 \in \mathcal{R}, cf_1 \in \mathcal{R}$, und es gilt*

$$\int_a^b (f_1+f_2)dx = \int_a^b f_1 dx + \int_a^b f_2 dx,\quad \int_a^b (cf_1)dx = c\int_a^b f_1 dx\,.$$

Ist $f_1(x) \le f_2(x)$ für alle $x \in [a,b]$, so gilt

$$\int_a^b f_1 dx \le \int_a^b f_2 dx.$$

(ii) *Ist $f \in \mathcal{R}$ auf $[a,b]$ und $a < c < b$, so ist $f \in \mathcal{R}$ auf $[a,b]$ und auf $[c,b]$, und es gilt*

$$\int_a^b f dx = \int_a^c f dx + \int_c^b f dx.$$

(iii) Ist $f \in \mathcal{R}$ auf $[a,b]$, und $|f(x)| \leq M$ für alle $x \in [a,b]$, so gilt

$$\Big|\int\limits_a^b f dx\Big| \leq M(b-a).$$

Beweis. Sei $\epsilon > 0$ gegeben. Wähle $\delta > 0$ sodass für jede Partition $P \in \mathcal{P}[a,b]$ mit $\mu(P) < \delta$ gilt dass

$$|L(P,f_1) - \int_a^b f_1\,dx| < \epsilon, \;\; |U(P,f_1) - \int_a^b f_1\,dx| < \epsilon$$

$$|L(P,f_2) - \int_a^b f_2\,dx| < \epsilon, \;\; |U(P,f_2) - \int_a^b f_2\,dx| < \epsilon$$

Nun gilt

$$\sup_{x\in[x_{i-1},x_i]}(f_1(x)+f_2(x)) \leq \sup_{x\in[x_{i-1},x_i]} f_1(x) + \sup_{x\in[x_{i-1},x_i]} f_2(x)$$

$$\inf_{x\in[x_{i-1},x_i]}(f_1(x)+f_2(x)) \geq \inf_{x\in[x_{i-1},x_i]} f_1(x) + \inf_{x\in[x_{i-1},x_i]} f_2(x)$$

also ist

$$L(P,f_1)+L(P,f_2) \leq L(P,f_1+f_2) \leq U(P,f_1+f_2) \leq U(P,f_1)+U(P,f_2)$$

Wir sehen, dass

$$-2\epsilon < L(P,f_1)+L(P,f_2) - \big[\int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx\big] \leq$$

$$\leq L(P,f_1+f_2) - \big[\int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx\big] \leq$$

$$\leq U(P,f_1+f_2) - \big[\int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx\big] \leq$$

$$\leq U(P,f_1)+U(P,f_2) - \big[\int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx\big] < 2\epsilon$$

und daher auch

$$\big|L(P,f_1+f_2) - \big[\int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx\big]\big| < 2\epsilon$$

$$\big|U(P,f_1+f_2) - \big[\int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx\big]\big| < 2\epsilon$$

Also ist die Funktion f_1+f_2 Riemann-integrierbar und es gilt

$$\int_a^b (f_1(x)+f_2(x))\,dx = \int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx\,.$$

Die Tatsache dass $cf_1 \in \mathcal{R}$ und $\int_a^b (cf_1(x))\,dx = c\int_a^b f_1(x)\,dx$ ist, sieht man genauso. Ist $f_1 \le f_2$, so folgt offenbar $L(P, f_1) \le L(P, f_2)$ für jede Partition P. Daher ist

$$\int_a^b f_1(x)\,dx = \sup_{P \in \mathcal{P}[a,b]} L(P, f_1) \le \sup_{P \in \mathcal{P}[a,b]} L(P, f_2) = \int_a^b f_2(x)\,dx\,.$$

Für (ii) sei $\epsilon > 0$ gegeben. Wähle eine Partition P mit $U(P, f) - L(P, f) < \epsilon$. Geht man nötigenfalls zu einer Verfeinerung von P über, so kann man oBdA annehmen, dass $c \in P$ ist. Dann ist $P \cap [a, c]$ eine Partition von $[a, c]$, und wir haben

$$U(P \cap [a,c], f|_{[a,c]}) - L(P \cap [a,c], f|_{[a,c]}) =$$

$$= \sum_{x_i \in P \cap [a,c]} \Big(\sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x) \Big)(x_i - x_{i-1}) \le$$

$$\le \sum_{x_i \in P} \Big(\sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x) \Big)(x_i - x_{i-1}) = U(P, f) - L(P, f) < \epsilon$$

also ist $f|_{[a,c]} \in \mathbb{R}$. Genauso erhält man $f|_{[c,b]} \in \mathcal{R}$. Wähle nun irgendeine Folge von Partitionen P_n mit $\mu(P_n) \to 0$ die alle den Punkt c enthalten. Dann ist

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} L(P_n, f) = \lim_{n\to\infty} \big(L(P_n \cap [a,c], f|_{[a,c]}) + L(P_n \cap [c,b], f|_{[c,b]}\big) =$$

$$= \lim_{n\to\infty} L(P_n \cap [a,c], f|_{[a,c]}) + \lim_{n\to\infty} L(P_n \cap [c,b], f|_{[c,b]}) = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx\,.$$

Für (iii) wähle eine Folge P_n mit $\mu(P_n) \to 0$, sodass also

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} S(f, P_n, P_n \setminus \{b\}) = \lim_{n\to\infty} \sum_{x_i \in P_n} f(x_i)(x_i - x_{i-1})\,.$$

Dann folgt

$$\Big|\int_a^b f(x)\,dx\Big| = \lim_{n\to\infty} \Big| \sum_{x_i \in P_n} f(x_i)(x_i - x_{i-1}) \Big| \le$$

$$\le \lim_{n\to\infty} \sum_{x_i \in P_n} |f(x_i)|(x_i - x_{i-1}) \le M(b-a)\,.$$

❑

Etwas komplizierter ist der Beweis der folgenden Aussage (*Dreiecksungleichung* für Integrale):

4.3.2 Satz. *Sei $f \in \mathcal{R}$ auf $[a, b]$. Dann ist auch $|f| \in \mathcal{R}$ und es gilt*

$$\Big|\int_a^b f\,dx\Big| \le \int_a^b |f|\,dx.$$

Beweis. Sei $\epsilon > 0$ gegeben. Dann existiert eine Partition P von $[a, b]$ sodaß

$$U(P, f) - L(P, f) < \epsilon.$$

Setze

$$M_i = \sup_{x_{i-1} \le x \le x_i} f(x), \; m_i = \inf_{x_{i-1} \le x \le x_i} f(x),$$
$$M_i^* = \sup_{x_{i-1} \le x \le x_i} | f(x) |, \; m_i^* = \inf_{x_{i-1} \le x \le x_i} |f(x)|.$$

Wegen $\big||f(x)| - |f(y)|\big| \le |f(x) - f(y)|$ ist

$$M_i^* - m_i^* \le M_i - m_i.$$

Es folgt

$$U(P, |f|) - L(P, |f|) = \sum_{i=1}^{n} (M_i^* - m_i^*) \triangle x_i \le$$

$$\le \sum_{i=1}^{n} (M_i - m_i) \triangle x_i = U(P, f) - L(P, f) < \epsilon,$$

und daher ist $|f| \in \mathcal{R}$.

Betrachte den positiven bzw. negativen Teil von f

$$f^+ := \max\{f(x), 0\}, f^- := -\min\{f(x), 0\}.$$

Wegen

$$f^+ = \frac{|f| + f}{2}, \; f^- = \frac{|f| - f}{2},$$

ist $f^+, f^- \in \mathcal{R}$. Es gilt $f = f^+ - f^-, |f| = f^+ + f^-$.
Also ist $(f^+, f^- \ge 0)$

$$|\int_a^b f dx| = |\int_a^b f^+ dx - \int_a^b f^- dx| \le |\int_a^b f^+ dx| +$$

$$+ |\int_a^b f^- dx| = \int_a^b f^+ dx + \int_a^b f^- dx = \int_a^b |f| dx.$$

❑

4.3.3 Korollar. *Ist $f, g \in \mathcal{R}$, so ist auch $fg \in \mathcal{R}$.*

Beweis. Ist P eine Partition und sei M_i bzw. m_i daß Supremum bzw. Infimum von $|f(x)|$ auf $[x_{i-1}, x_i]$, und M_i^*, m_i^* die gleichen Größen für $f^2(x)$. Weiters sei $M = \sup_{x \in [a,b]} |f(x)|$. Dann gilt

$$M_i^* - m_i^* = M_i^2 - m_i^2 = (M_i + m_i)(M_i - m_i) \le 2M(M_i - m_i),$$

also folgt

$$U(P, f^2) - L(P, f^2) \le 2M \left[U(P, |f|) - L(P, |f|)\right]$$

und wir erhalten $f^2 \in \mathcal{R}$.

Die Behauptung folgt nun aus der Beziehung

$$fg = \frac{1}{2}\left[(f+g)^2 - f^2 - g^2\right].$$

❑

Wie auch schon bei der Differentiation stellt sich auch hier die Frage nach der Vertauschbarkeit der Integration mit einem Grenzübergang.

4.3.4 Beispiel. Sei $(\alpha_n)_{n\in\mathbb{N}}$ eine Abzählung der rationalen Zahlen im Intervall $[0,1]$, d.h. seien α_n paarweise verschieden und sei $\{\alpha_n : n \in \mathbb{N}\} = \mathbb{Q} \cap [0,1]$. Betrachte die Funktionen $f_n : [0,1] \to \mathbb{R}$ die definiert sind als

$$f_n(x) := \begin{cases} 1 & , \ x \in \{\alpha_1, \dots, \alpha_n\} \\ 0 & , \ \text{sonst} \end{cases}$$

Dann ist jedes f_n stückweise stetig und daher in $\mathbb{R}$. Es gilt für jedes feste $x \in [0,1]$ dass $\lim_{n\to\infty} f_n(x) = f(x)$ wobei f die Funktion aus Beispiel 4.2.6 ist. Wir sehen also dass sich die Eigenschaft Riemann-integrierbar zu sein nicht notwendig auf punktweise Grenzwerte überträgt.

4.3.5 Satz. *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen auf $[a,b]$, $f_n \in \mathcal{R}, n \in \mathbb{N}$. Gilt $\lim_{n\to\infty} f_n = f$ gleichmäßig auf $[a,b]$, so ist $f \in \mathcal{R}$ und es gilt*

$$\lim_{n\to\infty} \int_a^b f_n \, dx = \int_a^b f \, dx.$$

Beweis. Sei $\epsilon > 0$ gegeben. Es existiert ein $n \in \mathbb{N}$, sodaß

$$|f_n(x) - f(x)| < \epsilon, \ x \in [a,b]. \tag{4.6}$$

Da $f_n \in \mathcal{R}$ existiert eine Partition P mit

$$U(P, f_n) - L(P, f_n) < \epsilon.$$

Wegen $f(x) < f_n(x) + \epsilon$ bzw. $f(x) > f_n(x) - \epsilon$ folgt

$$U(P, f) \le U(P, f_n) + \epsilon(b-a),$$

$$L(P, f) \ge L(P, f_n) - \epsilon(b-a).$$

Also erhält man

$$U(P, f) - L(P, f) < (2(b-a) + 1)\epsilon,$$

und daher ist $f \in \mathcal{R}$.

Wähle nun $N \in \mathbb{N}$, sodaß für $n \ge N$ stets (4.6) gilt. Dann ist also für alle $n \ge N$.

$$\Big|\int_a^b f\,dx - \int_a^b f_n\,dx\Big| = \Big|\int_a^b (f - f_n)\,dx\Big| \le$$

$$\leq \int_a^b |f - f_n| dx \leq \epsilon(b-a)\,.$$

❑

4.3.6 Beispiel. Sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit positivem Konvergenzradius R. Dann gilt für $[a,b] \subseteq \{x : |x| < R\}$

$$\int_a^b \left(\sum_{n=0}^{\infty} a_n x^n \right) dx = \sum_{n=0}^{\infty} a_n \int_a^b x^n \, dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} x^{n+1}|_{x=a}^{b}$$

4.4 Hauptsatz der Differential-Integralrechnung

4.4.1 Satz (Hauptsatz der Diff.-Int.Rechnung)**.** *Sei F differenzierbar auf $[a,b]$, $F'(x) =: f(x)$. Ist $f \in \mathcal{R}$ auf $[a,b]$, so gilt*

$$\int_a^b f(x)\, dx = F(b) - F(a).$$

Beweis. Sei P eine Partition von $[a,b]$. Dann existieren Zwischenstellen $t_i, x_{i-1} \leq t_i \leq x_i$, sodaß

$$F(x_i) - F(x_{i-1}) = (x_i - x_{i-1}) f(t_i).$$

Es folgt

$$F(b) - F(a) = \sum_{i=1}^{n} (F(x_i) - F(x_{i-1})) = \sum_{i=1}^{n} f(t_i) \triangle x_i.$$

Für $\mu(P) \to 0$ strebt die rechte Seite gegen $\int\limits_a^b f(x) dx$.

❑

4.4.2 Bemerkung. Um ein Integral $\int\limits_a^b f(x)\, dx$ zu berechnen genügt es also eine Funktion F zu finden von der f die Ableitung ist. Man spricht von einer *Stammfunktion* F von f. Eine solche ist nicht eindeutig bestimmt. Je zwei Stammfunktionen $F, \tilde{F}$ können sich jedoch nur um eine Konstante unterscheiden (vgl. Korollar 3.3.7). Die Gesamtheit aller Stammfunktionen bezeichnet man auch als das unbestimmte Integral von f und schreibt dafür $\int f(x)\, dx$. Die Existenz einer Stammfunktion ist nicht sowieso klar; sie wird im folgenden Satz bewiesen. Der Hauptsatz legt es nahe das Integral $\int\limits_a^b f(x) dx$ als Funktion seiner oberen Grenze zu betrachten um eine Stammfunktion zu erhalten.

4.4.3 Satz. *Sei $f \in \mathcal{R}$ auf $[a,b]$. Für $a \leq x \leq b$ definiere*

$$F(x) := \int_a^x f(t)\, dt.$$

Dann ist F stetig auf $[a,b]$. Ist f in einem Punkt x_0 stetig, so ist F bei x_0 differenzierbar und es gilt

$$F'(x_0) = f(x_0).$$

Beweis. Da $f \in \mathcal{R}$ ist, ist f beschränkt. Sei $|f(x)| \leq M, x \in [a,b]$. Für $a \leq x < y \leq b$ gilt

$$|F(y) - F(x)| = \Big|\int_x^y f(t)\,dt\Big| \leq M(y-x).$$

Insbesondere ist also F (gleichmäßig) stetig.

Sei nun f stetig bei x_0. Ist $\epsilon > 0$ gegeben, so existiert $\delta > 0$ sodaß

$$|f(t) - f(x_0)| < \epsilon \text{ für } |t - x_0| < \delta.$$

Ist $x_0 < x < x_0 + \delta$, so gilt

$$\Big|\frac{F(x)-F(x_0)}{x-x_0} - f(x_0)\Big| = \frac{1}{x-x_0}\Big|\int_{x_0}^{x} [f(t) - f(x_0)]\,dt\Big| < \epsilon.$$

Für $x_0 - \delta < x < x_0$ gilt

$$\Big|\frac{F(x)-F(x_0)}{x-x_0} - f(x_0)\Big| = \frac{1}{x_0-x}\Big|\int_{x}^{x_0} [f(t) - f(x_0)]\,dt\Big| < \epsilon.$$

Also folgt

$$\lim_{x\to x_0} \frac{F(x)-F(x_0)}{x-x_0} = f(x_0).$$

❏

Wir erhalten aus dem Hauptsatz auch noch unmittelbar zwei wichtige Rechenregeln für Integrale.

4.4.4 Korollar (Substitutionsregel)**.** *Sei $f[a,b] \to \mathbb{R}$ stetig, sei $\phi : [\alpha,\beta] \to \mathbb{R}$ differenzierbar und ϕ' stetig. Weiters sei $\phi([\alpha,\beta]) \subseteq [a,b]$, sodass also $f \circ \phi$ wohldefiniert ist. Dann gilt*

$$\int_{\phi(\alpha)}^{\phi(\beta)} f(x)\,dx = \int_{\alpha}^{\beta} f(\phi(y))\phi'(y)\,dy\,.$$

Beweis. Da f stetig ist, existiert sicher eine Stammfunktion F. Nach der Kettenregel gilt

$$[F(\phi(x))]' = F'(\phi(x))\phi'(x) = f(\phi(x))\phi'(x)\,.$$

Also ist $F(\phi(x))$ eine Stammfunktion von $f(\phi(x))\phi'(x)$. Es folgt

$$\int_\alpha^\beta f(\phi(x))\phi'(x)\,dx = F(\phi(\beta)) - F(\phi(\alpha)) = \int_{\phi(\alpha)}^{\phi(\beta)} f(t)\,dt\,.$$

❏

4.4.5 Korollar (Partielle Integration)**.** *Seien* $f, g : [a, b] \to \mathbb{R}$ *differenzierbar und seien* f', g' *stetig auf* $[a, b]$*. Dann gilt*

$$\int_a^b f(x)g'(x)\,dx = f(b)g(b) - f(a)g(a) - \int_a^b f'(x)g(x)\,dx\,.$$

Beweis. Nach der Produktregel gilt

$$[f(x)g(x)]' = f'(x)g(x) + f(x)g'(x),$$

d.h. es ist $f(x)g(x)$ eine Stammfunktion von $f'(x)g(x) + f(x)g'(x)$. Es folgt

$$\int_a^b f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx = f(b)g(b) - f(a)g(a)\,.$$

❏

Die folgenden Sätze sind oft praktisch.

4.4.6 Satz (1.Mittelwertsatz der Integralrechnung)**.** *Sei* f *stetig auf* $[a, b]$*. Dann existiert ein Punkt* $x \in [a, b]$ *mit*

$$\int_a^b f(t)\,dt = f(x)(b-a)\,.$$

Beweis. Sei $M = \sup_{t\in[a,b]} f(t), m = \inf_{t\in[a,b]} f(t)$. Dann gilt

$$m(b-a) \leq \int_a^b f\,d\alpha \leq M(b-a)\,.$$

Es existiert also eine Zahl λ, $m \leq \lambda \leq M$, mit

$$\int_a^b f(t)\,dt = \lambda(b-a)\,.$$

Da f als stetige Funktion auf der kompakten Menge $[a, b]$ sein Supremum bzw. Infimum annimmt und die Zwischenwerteigenschaft hat existiert $x \in [a, b]$ mit $f(x) = \lambda$.

❏

4.4.7 Satz (Verallgemeinerter 1.Mittelwertsatz der Integralrechnung)**.** *Sei* $f : [a, b] \to \mathbb{R}$ *stetig,* $g : [a, b] \to \mathbb{R}$ *Riemann-integrierbar,* $g(x) \geq 0$*. Dann existiert* $x \in [a, b]$ *mit*

$$\int_a^b f(t)g(t)\,dt = f(x)\int_a^b g(t)\,dt\,.$$

Beweis. Sei wieder $M = \sup_{t\in[a,b]} f(t), m = \inf_{t\in[a,b]} f(t)$, sodass $mg(x) \leq f(x)g(x) \leq Mg(x)$. Dann folgt

$$m \int_a^b g(t)\,dt \leq \int_a^b f(t)g(t)\,dt \leq M \int_a^b g(t)\,dt\,.$$

Ist $\int_a^b g(t)\,dt = 0$, so ist also auch $\int_a^b f(t)g(t)\,dt = 0$, und die Behauptung gilt für jedes $x \in [a,b]$. Sei $\int_a^b g(t)\,dt > 0$, dann setze

$$\lambda := \frac{\int_a^b f(t)g(t)\,dt}{\int_a^b g(t)\,dt}\,,$$

dann ist $m \leq \lambda \leq M$. Also existiert $x \in [a,b]$ mit $f(x) = \lambda$.

❑

4.4.8 Satz (2.Mittelwertsatz der Integralrechnung)**.** *Sei $f[a,b] \to \mathbb{R}$ monoton, differenzierbar, sei f' stetig auf $[a,b]$, und sei $g : [a,b] \to \mathbb{R}$ stetig. Dann existiert ein Punkt $x \in [a,b]$ mit*

$$\int\limits_a^b f(t)g(t)\,dt = f(a) \int_a^x g(t)\,dt + f(b) \int_x^b g(t)\,dt\,. \tag{4.7}$$

Beweis. Der Fall $f(a) = f(b)$ ist klar, denn dann ist f konstant und es gilt die Beziehung (4.7) für jedes $x \in [a,b]$. Betrachte den Fall dass $f(a) < f(b)$, also f monoton wachsend. Dann ist $f' \geq 0$. Sei $G(x) := \int_a^x g(t)\,dt$, dann ist

$$\int_a^b f(t)g(t)\,dt = f(b)G(b) - \underbrace{f(a)G(a)}_{=0} - \int_a^b f'(t)G(t)\,dt\,.$$

Nach dem Verallgemeinerten 1.Mittelwertsatz existiert $x \in (a,b)$ mit

$$\int_a^b G(t)f'(t)\,dt = G(x) \int_a^b f'(t)\,dt = G(x)[f(b) - f(a)]\,.$$

Es folgt

$$\begin{aligned}\int_a^b f(t)g(t)\,dt &= f(b)G(b) - G(x)[f(b) - f(a)] = \\ &= f(a)G(x) + f(b)[G(b) - G(x)]\,.\end{aligned}$$

❑

Als eine Anwendung beweisen wir die Taylorsche Formel (vgl. Satz 3.4.5) mit einem Restglied in Integralform.

4.4.9 Beispiel. Sei f auf $[a,b]$ n-mal stetig differenzierbar. Dann gilt

$$f(b) = f(a) + \frac{f'(a)}{1!}(b-a) + \frac{f''(a)}{2!}(b-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1} +$$

$$+\int_a^b f^{(n)}(x)\frac{(b-x)^{n-1}}{(n-1)!}\,dx.$$

Um diese Formel einzusehen starten wir mit

$$f(b)-f(a)=\int_a^b f'(x)\,dx,$$

und integrieren partiell:

$$f(b)-f(a)=-\int_a^b f'(x)(-1)\,dx=f'(a)(b-a)+\int_a^b f''(x)(b-x)dx=$$

$$=f'(a)(b-a)+f''(a)\frac{(b-a)^2}{2}+\int_a^b f'''(x)\frac{(b-x)^2}{2}\,dx=$$

$$=\cdots=\sum_{k=1}^{n-1} f^{(k)}(a)\frac{(b-a)^k}{k!}+\int_a^b f^{(n)}(x)\frac{(b-x)^{n-1}}{(n-1)!}\,dx\,.$$

Aus dieser Formel kann man mit Hilfe des Verallgemeinerten 1.Mittelwertsatzes, angewendet mit $f^{(n)}$ und $g(x)=\frac{(b-x)^{n-1}}{(n-1)!}$, auch die uns schon bekannte Formel für das Restglied erhalten. Denn dann erhalten wir

$$\int_a^b f^{(n)}(x)\frac{(b-x)^{n-1}}{(n-1)!}\,dx=f^{(n)}(x)\frac{(b-a)^n}{n!}$$

für ein geeignetes $\xi\in[a,b]$.

Schliesslich wollen wir ein paar Methoden bzw. Beispiele zur expliziten Berechnung der Stammfunktion angeben. Dazu zunächst eine

4.4.10 Bemerkung.

(*i*) Die Regel der partiellen Integration bzw. die Substitutionsregel schreiben sich für unbestimmte Integral wie folgt:

$$\int f(x)\cdot g'(x)\,dx=f(x)\cdot g(x)-\int f'(x)\cdot g(x)\,dx$$

$$\int f(\phi(x))\phi'(x)\,dx=\int f(x)\,dx$$

(*ii*) Die Stammfunktion einer gegebenen Funktion f ist (falls sie überhaupt existiert) jedenfalls nur bis auf eine Konstante bestimmt. Das Symbol $\int f(x)\,dx$ bezeichnet die Gesamtheit aller Stammfunktionen. Wenn man eine (und damit alle) Stammfunktion(en) explizit angeben will, schreibt man oft $\int f(x)\,dx=F(x)+C$, und meint mit C eine freie Konstante. Zum Beispiel schreibt man

$$\int x^2\,dx=\frac{x^3}{3}+C$$

und meint damit, dass die Menge aller Stammfunktionen von x^2 gegeben ist als $\{\frac{x^3}{3} + C : C \in \mathbb{R}\}$.

4.4.11 Beispiel. Wir wollen das Integral $\int \tan x \, dx$ berechnen:

$$\int \tan x \, dx = \int \frac{\sin x}{\cos x} \, dx \underset{\left[\substack{t=\cos x \\ dt=-\sin x dx}\right]}{=} -\int \frac{1}{t} \, dt =$$

$$= -\ln t + C = -\ln(\cos x) + C \, .$$

Diese Methode beruht darauf, dass unser Integrand zufällig von der Gestalt $f(\phi(x))\phi'(x)$ ist, und noch dazu mit einer sehr einfachen Funktion f. Daher können wir die Substitutionsregel unmittelbar anwenden und die entstehende Funktion leicht integrieren.

4.4.12 Beispiel. Wir wollen das Integral $\int \ln x \, dx$ berechnen. Dazu verwenden wir partielle Integration:

$$\int \ln x \, dx = \int \ln x \cdot 1 \, dx = \ln x \cdot x - \int \frac{1}{x} \cdot x \, dx = x \ln x - x + C$$

4.4.13 Beispiel. Wir wollen das Integral $\int x^2 \sin x \, dx$ berechnen. Wieder verwenden wir partielle Integration:

$$\int x^2 \sin x \, dx = x^2(-\cos x) - \int 2x(-\cos x) \, dx = -x^2 \cos x + 2 \int x \cos x \, dx =$$

$$= -x^2 \cos x + 2\big(x \sin x - \int \sin x \, dx\big) = -x^2 \cos x + 2x \sin x + 2 \cos x + C$$

Mit dieser Methode kann man offenbar zum Beispiel alle Integrale von der Form

$$\int P(x) e^x \, dx, \quad \int P(x) \sin x \, dx, \quad \int P(x) \cos x \, dx \, ,$$

mit einem Polynom P berechnen.

4.4.14 Beispiel (Integration rationaler Funktionen). Für eine rationale Funktion $f(x) = \frac{p(x)}{q(x)}$ mit $p, q \in \mathbb{R}[x]$, $\operatorname{grad} p < \operatorname{grad} q$, lässt sich stets eine sogenannte *Partialbruchzerlegung* durchführen. Das ist eine additive Zerlegung von f in Summanden der Gestalt

$$\frac{c}{(x-\gamma)^l}, \quad \frac{ax+b}{(x^2+\alpha x+\beta)^l}$$

wobei $c, a, b, \gamma, \alpha, \beta \in \mathbb{R}$, $l \in \mathbb{N}$, und wobei das Polynom $x^2 + \alpha x + \beta$ keine reellen Nullstellen hat, d.h. dass $\frac{\alpha^2}{4} - \beta < 0$ ist.

Für jeden dieser einzelnen Summanden kann man nun das Integral ausrechnen. Summanden des ersten Typs sind einfach:

$$\int \frac{c}{x-\gamma} \, dx = c \ln(x-\gamma) + C, \quad \int \frac{c}{(x-\gamma)^l} \, dx = -\frac{c}{l-1} \frac{1}{(x-\gamma)^{l-1}} + C, \ l > 1$$

Um Summanden des zweiten Typs mit $l = 1$ zu integrieren, schreiben wir

$$x^2 + \alpha x + \beta = \Big(x + \frac{\alpha}{2}\Big)^2 + \Big(\beta - \frac{\alpha}{2}\Big)$$

Der zweite Summand ist positiv, ist also gleich a^2 mit einer eindeutigen positiven Zahl a. Substituiert man nun $x + \frac{\alpha}{2} = t$, so folgt

$$\begin{aligned}
\int \frac{ax+b}{x^2+\alpha x+\beta}\,dx \underset{\left[\substack{t=x+\frac{\alpha}{2}\\ dt=dx}\right]}{=} & \int \frac{at+(b-\frac{a\alpha}{2})}{t^2+a^2}\,dt = \\
&= \frac{a}{2}\int \frac{2t}{t^2+a^2}\,dt + \big(b-\frac{a\alpha}{2}\big)\int \frac{1}{t^2+a^2}\,dt = \\
&= \frac{a}{2}\ln(t^2+a^2) + \big(b-\frac{a\alpha}{2}\big)\frac{1}{a}\arctan\frac{t}{a} + C = \\
&= \frac{a}{2}\ln(x^2+\alpha x+\beta) + \frac{2b-a\alpha}{\sqrt{4\beta-\alpha^2}}\arctan\frac{2x+\alpha}{\sqrt{4\beta-\alpha^2}}
\end{aligned}$$

Im Fall $l > 1$ verwendet man die gleiche Substitution.

4.4.15 Beispiel (Integration von $R(e^x)$). Hat man eine Funktion f der Gestalt $f(x) = R(e^x)$ wobei R eine rationale Funktion ist, so kann man deren Integral stets mit Hilfe der Substitution $t = e^x$ auf das Integral einer rationalen Funktion bringen. Denn es gilt, mit $t = e^x$,

$$\int R(e^x)\,dx \underset{\left[\substack{t=e^x\\ dt=e^x dx\\ dx=\frac{1}{t}dt}\right]}{=} \int R(t)\frac{1}{t}\,dt$$

4.4.16 Beispiel. Wir wollen das Integral $\int \frac{e^{2x}-1}{e^x+2}\,dx$ berechnen. Zunächst gilt

$$\int \frac{e^{2x}-1}{e^x+2}\,dx = \int \frac{(e^x)^2-1}{e^x+2}\,dx \underset{\left[\substack{t=e^x\\ dx=\frac{1}{t}dt}\right]}{=} \int \frac{t^2-1}{t+2}\cdot\frac{1}{t}\,dt$$

Nun haben wir ein Integral einer rationalen Funktion zu berechnen. Dazu schreiben wir

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

und versuchen nun den zweiten Summanden in Partialbrüche zu zerlegen:

$$\frac{2t-1}{t(t+2)} = \frac{A}{t} + \frac{B}{t+2} = \frac{(A+B)t+2A}{t(t+2)}$$

Koeffizientenvergleich führt auf $A = -\frac{1}{2}$ und $B = \frac{5}{2}$. Also haben wir

$$\frac{t^2-1}{t+2}\cdot\frac{1}{t} = 1 + \frac{1}{2t} - \frac{5}{2(t+2)}$$

und daher

$$\int \frac{t^2-1}{t+2}\cdot\frac{1}{t}\,dt = t + \frac{1}{2}\ln t - \frac{5}{2}\ln(t+2) =$$

$$= t + \frac{1}{2}\ln t - \frac{5}{2}\ln(t+2) = e^x + \frac{x}{2} - \frac{5}{2}\ln(e^x+2)$$

4.4.17 Beispiel (Integration von $R(\sin x, \cos x)$). Hat man eine Funktion f der Gestalt $f(x) = R(\sin x, \cos x)$ wobei R eine rationale Funktion ist, so kann man deren Integral stets mit Hilfe der Substitution $t = \tan \frac{x}{2}$ auf das Integral einer rationalen Funktion bringen. Denn es gilt, mit $t = \tan \frac{x}{2}$,

$$\sin x = 2 \sin \frac{x}{2} \cos \frac{x}{2} = \frac{2 \tan \frac{x}{2}}{1 + \tan^2 \frac{x}{2}} = \frac{2t}{1+t^2}$$

$$\cos x = \cos^2 \frac{x}{2} - \sin^2 \frac{x}{2} = \frac{1 - \tan^2 \frac{x}{2}}{1 + \tan^2 \frac{x}{2}} = \frac{1-t^2}{1+t^2}$$

$$\frac{dt}{dx} = \frac{1}{2}(1 + \tan^2 \frac{x}{2}) \text{ also } dx = \frac{2}{1+t^2} dt$$

4.5 Eine Näherungsformel

In den meisten Fällen wird es nicht möglich sein zu einer Funktion eine Stammfunktion explizit anzugeben, und damit das Integral $\int_a^b f\,dx$ tatsächlich auszurechnen. Dieses Phänomen tritt zum Beispiel schon bei so „einfachen" Funktionen wie $f(x) = e^{x^2}$ auf.

Ein möglicher Weg um trotzdem den Wert von $\int_a^b f\,dx$, zumindest näherungsweise, zu bestimmen, ist die gegebene Funktion f durch eine andere Funktion g zu ersetzen deren Integral man berechnen kann. Als erstes denkt man dabei für g wohl an Polynome.

Wir wissen bereits, dass man eine Funktion f so durch ein Polynom p vom Grad höchstens n interpolieren kann, dass $f(x) = p(x)$ für $n+1$ Stellen gilt. Dieses Polynom kann man mit Hilfe der Newtonschen Interpolationsformel auch explizit angeben. Oft ist es praktisch dieses Interpolationspolynom anders zu konstruieren.

4.5.1 Satz (Lagrangesche Interpolationsformel)**.** *Seien paarweise verschiedene Stützstellen $x_0, \ldots, x_n$ und dazugehörige Werte $y_0, \ldots, y_n$ gegeben. Setze*

$$L_i(x) := \prod_{\substack{k=0 \\ k \neq i}} \frac{x - x_k}{x_i - x_k}, \quad i = 0, \ldots, n\,.$$

Dann ist das Polynom

$$p(x) := \sum_{i=0}^{n} y_i L_i(x)$$

eine Lösung des gegebenen Interpolationsproblems, d.h. es gilt $p(x_i) = y_i$, $i = 0, \ldots, n$.

Beweis. Das Polynom L_i hat die Eigenschaft das

$$L_i(x_k) = \begin{cases} 1 & , \ k = i \\ 0 & , \ k \neq i \end{cases}$$

❑

Klarerweise ist das Interpolationspolynom (vom Grad $\leq n$) eindeutig bestimmt.

Interpoliert man nun eine Funktion f an den Stellen $x_0, \ldots, x_n$ durch ein Polynom P_n vom Grad $\leq n$, dann ist es natürlich wesentlich eine Information über die Abweichung von $P_n(x)$ und $f(x)$ zu haben.

4.5.2 Satz. *Die Funktion $f : [a,b] \to \mathbb{R}$ sei stetig und $(n+1)$-mal differenzierbar auf (a,b). Weiters seien $a = x_0 < x_1 < \ldots < x_n = b$ gegeben, und wir bezeichnen $\omega(x) := \prod_{i=0}^{n}(x - x_i)$. Sei P_n das Polynom vom Grad $\leq n$ mit $P(x_k) = f(x_k)$, $k = 0, \ldots, n$, und sei $R_n(x) := f(x) - P_n(x)$. Dann gilt*

$$|R_n(x)| \leq \frac{\max_{x\in(a,b)} |f^{(n+1)}(x)|}{(n+1)!} \max_{x\in(a,b)} |\omega(x)| \, .$$

Beweis. Sei $x \in [a,b]$ fest, und sei $x \neq x_i$, $i = 0, \ldots, n$. Dann betrachten wir die Hilfsfunktion

$$r_n(t) := R_n(t) - \frac{R_n(x)}{\omega(x)}\omega(t) \, .$$

Diese Funktion ist $(n+1)$-mal differenzierbar und hat die $n+2$ Nullstellen $x_0, \ldots, x_n$ und x. Daher hat $r_n^{(n+1)}$ eine Nullstelle ξ in (a,b). Nun ist

$$P_n^{(n+1)}(t) = 0 \text{ und } \omega^{(n+1)}(t) = (n+1)! \, ,$$

und daher

$$r_n^{(n+1)}(t) = f^{(n+1)}(t) - \frac{R_n(x)}{\omega(x)}(n+1)! \, .$$

Setzt man $t = \xi$, so folgt

$$R_n(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)\omega(x) \, , \tag{4.8}$$

und damit die Behauptung.

❑

4.5.3 Bemerkung. Diese Abschätzung setzt sich aus den beiden Faktoren $\max_{x\in(a,b)} |f^{(n+1)}(x)|$ und $\max_{x\in(a,b)} |\omega(x)|$. Der erste hängt nicht von der Wahl der Stützstellen x_i ab, der zweite nicht von der betrachteten Funktion f. Man wird also versuchen durch geeignete Wahl der Stützstellen den zweiten Faktor so klein wie möglich zu machen. Beispielsweise für das Intervall $[a,b] = [-1,1]$ wird dies durch

$$x_i = \cos \frac{\pi(2i+1)}{2(n+1)}, \; i = 0, \ldots, n$$

geleistet.

4.5.4 Korollar. *Die Funktion $f : [a,b] \to \mathbb{R}$ sei $(n+1)$-mal differenzierbar. Weiters seien $a = x_0 < x_1 < \ldots < x_n = b$, und P bezeichne das Polynom vom Grad $\leq n$ mit $P(x_k) = f(x_k)$, $k = 0, \ldots, n$. Dann gilt*

$$\Big| \int_a^b f(x)\, dx - \int_a^b P(x)\, dx \Big| \leq \frac{\max_{x\in(a,b)} |f^{(n+1)}(x)|}{(n+1)!} \int_a^b \prod_{i=0}^{n} |x - x_i| \, dx \, .$$

Beweis. Es gilt nach (4.8)

$$\Big|\int_a^b f(x)\,dx - \int_a^b P(x)\,dx\Big| \le \int_a^b |f(x)-P(x)|\,dx \le$$

$$\le \int_a^b \frac{\max_{x\in(a,b)} |f^{(n+1)}(x)|}{(n+1)!} \prod_{i=0}^n |x-x_i|\,dx\,.$$

❑

Die Spezialfälle $n = 1$ bzw. $n = 2$ mit äquidistanter Zerlegung sind als *Trapezformel* bzw. *Simpson-Formel* oder *Keplersche Faßregel* bekannt.

4.5.5 Korollar. *Es gilt*

(*i*) *Die Funktion* $f : [a,b] \to \mathbb{R}$ *sei* 2*-mal differenzierbar. Setze*

$$I_T := \frac{b-a}{2}\big(f(a)+f(b)\big)\,,$$

dann gilt

$$\Big|\int_a^b f(x)\,dx - I_T\Big| \le \frac{(b-a)^3}{12} \max_{x\in(a,b)} |f''(x)|\,.$$

(*ii*) *Die Funktion* $f : [a,b] \to \mathbb{R}$ *sei* 3*-mal differenzierbar. Setze*

$$I_S := \frac{b-a}{6}\Big(f(a)+4f\big(\frac{a+b}{2}\big)+f(b)\Big)\,,$$

dann gilt

$$\Big|\int_a^b f(x)\,dx - I_S\Big| \le \frac{3(b-a)^4}{64} \max_{x\in(a,b)} |f'''(x)|\,.$$

Beweis. Im Fall $n = 1$ ist das Interpolationspolynom gegeben als $P_1(x) = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$, und daher ist

$$\int_a^b P_1(x)\,dx = \big(f(a)x + \frac{f(b)-f(a)}{b-a}\frac{(x-a)^2}{2}\big)\Big|_{x=a}^b =$$

$$= f(a)(b-a) + \frac{f(b)-f(a)}{b-a}\frac{(b-a)^2}{2} = I_T\,.$$

Wir haben $|\omega(x)| = (x-a)(b-x)$, $x \in [a,b]$. Daher ist

$$\int_a^b |\omega(x)|\,dx = \int_a^b (-x^2+(a+b)x-ab)\,dx =$$

$$\big(-\frac{x^3}{3}+(a+b)\frac{x^2}{2}-abx\big)\Big|_{x=a}^b = \frac{1}{6}(b-a)^3\,.$$

Betrachte nun den Fall dass $n = 2$ und $x_1 = \frac{a+b}{2}$. Dann ist nach der Lagrangeschen Interpolationsformel das Interpolationspolynom P_2 gegeben als

$$P_2(x) = f(a)\frac{2}{(b-a)^2}(x-\frac{a+b}{2})(x-b) - f(\frac{a+b}{2})\frac{4}{(b-a)^2}(x-a)(x-b)+$$

$$+f(b)\frac{2}{(b-a)^2}(x-a)(x-\frac{a+b}{2})\,.$$

Man berechnet damit $\int_a^b P_2(x)\,dx = I_S$. Es ist $\omega(x) = (x-a)(x-\frac{a+b}{2})(x-b)$, und eine analoge Rechnung wie oben führt zur gewünschten Abschätzung.

❑

Interessant ist es zu bemerken, dass die Simpson-Formel tatsächlich eine wesentlich bessere Näherung liefert.

4.5.6 Satz. *Die Funktion* $f : [a,b] \to \mathbb{R}$ *sei* 4*-mal differenzierbar. Dann gilt*

$$\Big|\int_a^b f(x)\,dx - I_S\Big| \le \frac{(b-a)^5}{2880}\max_{x\in(a,b)}|f^{(4)}(x)|\,.$$

Beweis. Sei $P_2(x)$ das Interpolationspolynom aus der Simpson-Formel, d.h. es sei gebildet mit den Stützstellen $x_0 = a, x_1 = \frac{a+b}{2}, x_2 = b$. Sei ein weiterer, von diesen verschiedener, Punkt $x^* \in (a,b)$ gegeben, und bezeichne mit $P^*(x)$ das Interpolationspolynom vom Grad ≤ 3 welches an den Stellen x_0, x_1, x_2 und noch an x^* interpoliert. Es muss dann gelten

$$P^*(x) = P(x) + \lambda(x-x_0)(x-x_1)(x-x_2)$$

mit einer geeigneten Konstanten λ. Es folgt

$$\int_a^b P^*(x)\,dx = \int_a^b P(x)\,dx + \lambda\int_a^b (x-x_0)(x-x_1)(x-x_2)\,dx\,.$$

Da die Funktion $h(x) := (x-x_0)(x-x_1)(x-x_2)$ ungerade bezüglich $\frac{a+b}{2}$ ist, d.h. $h(\frac{a+b}{2}-x) = -h(\frac{a+b}{2}+x)$, ist das letzte Integral gleich 0, und wir erhalten $\int_a^b P^*(x)\,dx = I_S$.

Die Fehlerabschätzung aus Korollar 4.5.4 liefert nun

$$\Big|\int_a^b f(x)\,dx - I_S\Big| = \Big|\int_a^b f(x)\,dx - \int_a^b P^*(x)\,dx\Big| \le$$

$$\le \frac{\max_{x\in(a,b)}|f^{(4)}(x)|}{4!}\int_a^b \big(|x-x_0|\,|x-x_1|\,|x-x_2|\,\underbrace{|x-x^*|}_{=|x-x_1+x_1-x^*|}\big)\,dx \le$$

$$\le \frac{\max_{x\in(a,b)}|f^{(4)}(x)|}{24}\Big[\int_a^b \big(|x-x_0|\,|x-x_1|^2\,|x-x_2|\big)\,dx +$$

$$|x-x^*|\int_a^b \big(|x-x_0|\,|x-x_1|\,|x-x_2|\big)\,dx\Big]\,.$$

Da der Wert auf der linken Seite nicht von der Wahl von x^* abhängt, und diese Ungleichung für alle $x^* \ne x_0, x_1, x_2$ gilt, können wir $x^* \to x_1$ streben lassen und erhalten

$$\Big|\int_a^b f(x)\,dx - I_S\Big| \le \frac{\max_{x\in(a,b)}|f^{(4)}(x)|}{24}\int_a^b \big(|x-x_0|\,|x-x_1|^2\,|x-x_2|\big)\,dx\,.$$

Das letzte Integral berechnet man als ($x = \frac{a+b}{2} + \frac{b-a}{2}u$, $dx = \frac{b-a}{2}du$, $x = a \leftrightarrow u = -1$, $x = b \leftrightarrow u = 1$)

$$\int_a^b \big(|x - x_0|\,|x - x_1|^2\,|x - x_2|\big)\,dx =$$

$$= \Big(\frac{b-a}{2}\Big)^5 \int_{-1}^{1} (1-u)u^2(1+u)\,du = \Big(\frac{b-a}{2}\Big)^5 \frac{4}{15}\,.$$

Insgesamt erhält man die gewünschte Abschätzung.

❑

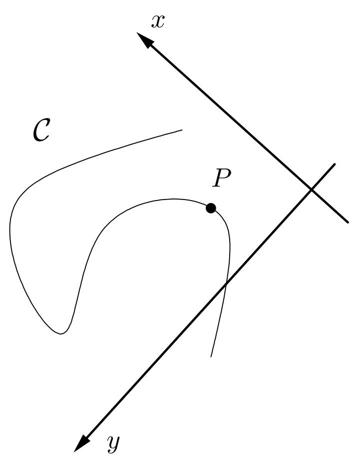
4.5.7 Bemerkung. Man sieht also, dass man Näherungsformeln für $\int_a^b f(x)\,dx$ erhält, deren Fehler sich wie eine gewisse Potenz der Intervalllänge $b-a$ verhält. Praktisch wird man also zuerst das gegebene Intervall in meherere kleine Teile zerlegen, und dann auf jeden Teil einzeln eine Näherungsformel anweden.

4.6 Differentialgeometrie

Wir betrachten als Beispiel für die Untersuchung eines geometrischen Objektes mit Hilfe der Differential- und Integralrechnung Kurven im dreidimensionalen Raum.

Um ein solches geometrisches Objekt mit analytischen Mitteln studieren zu können, muß man es in irgendeiner Weise formelmäßig erfaßt haben.

Ist eine Kurve $\mathcal{C}$ gegeben, so möchte man also die Lage der Kurvenpunkte P durch Formeln beschreiben. Dazu muß man einmal ein Koordinatensystem wählen und dann versuchen die Kurvenpunkte durch Beziehungen zwischen ihren Koordinaten zu charakterisieren.



In früheren Abschnitten haben wir uns den Graphen einer Funktion $y = f(x)$ auch als Kurve (in der Ebene) vorgestellt. Das ist eine Möglichkeit eine Kurve $\mathcal{C}$ formelmäßig festzulegen

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 : y = f(x)\}.$$

Man spricht von einer *expliziten Darstellung*. Dabei ist aber die y-Koordinate gegenüber der x-Koordinate ausgezeichnet, was geometrisch nicht Sinn macht da die Lage des Koordinatensystems ja willkürlich war.

Die Idee eine Beziehung zwischen den Koordinaten der Punkte der Kurve anzugeben führt auf eine sogenannte *implizite Darstellung*. Dabei hat man eine Funktion F, derart daß

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 : F(x, y) = 0\}.$$

Eine sehr allgemeine und für die meisten Untersuchungen gut geeignete Möglichkeit der koordinatenweisen Darstellung der Kurvenpunkte erhält man aus der folgenden Vostellung: Man fasse die Kurve auf als Bahn der Bewegung eines Punktes. Man schreibt also

$$\mathcal{C} = \{(x, y) \in \mathbb{R}^2 : (x, y) = P(t) \text{ für ein } t\}$$

wobei P eine Funktion von einem gewissen Parameterbereich in die Ebene ist. $P(t)$ ist also die Position des bewegten Punktes zum Zeitpunkt t. Man spricht in diesem Fall von einer *Parameterdarstellung* der Kurve. Wir werden im folgenden stets mit einer Kurve in Parameterdarstellung arbeiten.

Sei im folgenden $\mathcal{C}$ eine Kurve im Raum, sei ein Koordinatensystem fix gewählt und sei $\mathcal{C}$ gegeben durch die Parameterdarstellung $\mathfrak{x} = \mathfrak{x}(t)$. Basis aller analytischen Untersuchungen ist das Differenzieren. Wir setzen daher stets voraus, daß $\mathfrak{x}(t)$ hinreichend glatt ist (d.h. stetig und hinreichend oft differenzierbar).

4.6.1 Bemerkung. Spricht man von der Bewegung eines Puntes $\mathfrak{x}$, dessen Lage zum Zeitpunkt t durch die Funktion $\mathfrak{x}(t)$ gegeben ist, so hat es sich eingebürgert für die Ableitung nach dem Parameter t die Bezeichnung $\dot{\mathfrak{x}}(t)$ zu verwenden, d.h.

$$\dot{\mathfrak{x}}(t) := \frac{d\mathfrak{x}}{dt}(t).$$

4.6.2 Satz. *Sei $\mathfrak{x} = \mathfrak{x}(t)$ eine Kurve in Parameterdarstellung. Der Vektor $\dot{\mathfrak{x}}(t)$ zeigt in Richtung der Tangente an die Kurve im Punkt $\mathfrak{x}(t)$. Die Länge s des Kurvenbogens zwischen zwei Punkten $\mathfrak{x}(t)$ und $\mathfrak{x}(T)$ ist gegeben durch*

$$s = \int_t^T \sqrt{\dot{\mathfrak{x}}(t) \cdot \dot{\mathfrak{x}}(t)}\, dt. \tag{4.9}$$

Beweis. Ist $\mathfrak{x}(t)$ ein Kurvenpunkt, so stellt man sich die Tangente wieder vor als Grenzlage von Sekanten durch $\mathfrak{x}(t)$ und einen benachbarten Kurvenpunkt $\mathfrak{x}(t + \triangle t)$. Offenbar ist

$$\frac{\mathfrak{x}(t + \triangle t) - \mathfrak{x}(t)}{\triangle t}$$

ein Vektor in Richtung der Sekante, also zeigt

$$\dot{\mathfrak{x}}(t) = \lim_{\triangle t \to 0} \frac{\mathfrak{x}(t + \triangle t) - \mathfrak{x}(t)}{\triangle t}$$

in Richtung der Tangente.

Um die Formel (4.9) für die Bogenlänge zu erklären, stelle man sich das Parameterintervall in kleine Teile zerlegt und die Bewegung des Punktes stückweise geradlinig vor, d.h. man approximiert die Kurve $\mathcal{C}$ durch einen Streckenzug $\tilde{\mathcal{C}}$. Die Länge von $\tilde{\mathcal{C}}$ läßt sich nun leicht berechnen. Zerlegt man zum Beispiel das Intervall $[t, T]$ in n gleiche Teile $t, t + \triangle t, \ldots, t + n\triangle t = T$ wobei $\triangle t := \frac{T-t}{n}$, so erhält man nach dem Satz von Pythagoras für die Länge L_n des Streckenzuges

$$L_n = \sum_{i=0}^{n-1} \sqrt{[\mathfrak{x}(t + (i+1)\triangle t) - \mathfrak{x}(t + i\triangle t)]^2}$$

Hier bezeichnet $.^2$ das skalare Produkt des Sehnenvektors mit sich selbst.
Nun ist

$$L_n = \sum_{i=0}^{n-1} \sqrt{\left[\frac{\mathfrak{x}(t + (i+1)\triangle t) - \mathfrak{x}(t + i\triangle t)}{\triangle t}\right]^2} \triangle t.$$

Wendet man auf jede einzelne Komponente von $\mathfrak{x}$ ($\mathfrak{x}(t) = (x_1(t), x_2(t), x_3(t))$ den Mittelwertsatz an, so erhält man

$$\frac{\mathfrak{x}(t + (i+1)\triangle t) - \mathfrak{x}(t + i\triangle t)}{\triangle t} = \begin{pmatrix} \dot{x}_1(t + (i+\theta_1)\triangle t) \\ \dot{x}_2(t + (i+\theta_2)\triangle t) \\ \dot{x}_3(t + (i+\theta_3)\triangle t) \end{pmatrix}$$

mit gewissen $\theta_1, \theta_2, \theta_3 \in (0, 1)$. Man erhält also

$$L_n = \sum_{i=0}^{n-1} \triangle t \sqrt{\dot{x_1}^2(t + (i+\theta_1)\triangle t) + \dot{x_2}^2(t + (i+\theta_2)\triangle t) + \dot{x_3}^2(t + (i+\theta_3)\triangle t)}.$$

Da $\dot{\mathfrak{x}(t)}$ eine stetige Funktion auf $[t, T]$ ist (wir setzen ja stets $\mathfrak{x}$ als hinreichend glatt voraus), sind alle drei Komponenten auch gleichmäßig stetig. D.h. zu jedem $\epsilon > 0$ existiert $\delta > 0$ sodaß für $|t_1 - t_2| < \delta$ gilt $|\dot{\mathfrak{x}}_i(t_1) - \dot{\mathfrak{x}}_i(t_2)| < \epsilon$. Für $\triangle t$ hinreichend klein, $\triangle t < \epsilon$, erhält man mit der Dreiecksungleichung für den euklidischen Abstand im $\mathbb{R}^3$

$$\Big| \sqrt{\dot{x_1}^2(t + (i+\theta_1)\triangle t) + \dot{x_2}^2(t + (i+\theta_2)\triangle t) + \dot{x_3}^2(t + (i+\theta_3)\triangle t)} -$$

$$- \sqrt{\dot{x_1}^2(t + i\triangle t) + \dot{x_2}^2(t + i\triangle t) + \dot{x_3}^2(t + i\triangle t)} \Big| < \epsilon\sqrt{3}.$$

Also folgt

$$\left| L_n - \sum_{i=0}^{n-1} \triangle t \sqrt{\dot{x_1}^2(t + i)\triangle t) + \dot{x_2}^2(t + i\triangle t) + \dot{x_3}^2(t + i)\triangle t)} \right| <$$

$$< \epsilon\sqrt{3} \sum_{i=0}^{n-1} \triangle t = \epsilon\sqrt{3}(T - t).$$

Läßt man nun $\triangle t$ gegen 0 streben, so ist also der Grenzwert von L_n gleich jener der Summe

$$\sum_{i=0}^{n-1} \triangle t \sqrt{\dot{x_1}^2(t + i)\triangle t) + \dot{x_2}^2(t + i\triangle t) + \dot{x_3}^2(t + i)\triangle t)}.$$

Das ist eine Riemannsche Zwischensumme, und strebt daher für $\triangle t \to 0$ gegen $\int_t^T \sqrt{\dot{\mathfrak{x}}(t)^2} dt$.

❑

4.6.3 Bemerkung. Eigentlich geben wir in "Satz 4.6.2" eine Definition der Tangente und der Bogenlänge. Der "Beweis" ist also eigentlich die Aussage, daß diese Definitionen tatsächlich das wiederspiegeln was man sich anschaulich unter diesen Begriffen vorstellt.

Da wir uns nur für die Kurve $\mathcal{C}$ interessieren, nicht jedoch für die Art der Bewegung des Punktes $\mathfrak{x}(t)$, können wir o.B.d.A. verlangen, daß der Betrag der Geschwindigkeit von $\mathfrak{x}(t)$ stets gleich 1 ist, d.h. daß gilt $\sqrt{\dot{\mathfrak{x}}(t)^2} = 1$. Damit ist aber auch

$$s = \int_t^T \sqrt{\dot{\mathfrak{x}}(t)^2}\, dt = T - t,$$

d.h. der Parameter beschreibt genau die Bogenlänge.

Es scheint also sinnvoll die geometrische Größe der Bogenlänge als Parameter zu verwenden. Wir werden dies im folgenden stets tun.

4.6.4 Bemerkung. Um anzudeuten, daß man den geometrischen Parameter s (Bogenlänge) und nicht den physikalischen t (Zeit) verwendet hat es sich eingebürgert die Ableitung nach der Bogenlänge als $\mathfrak{x}'(s)$ zu bezeichnen.

Die Tangente an die Kurve im Punkt $\mathfrak{x}(s_0)$ ist gegeben durch

$$\mathfrak{y}(t) = \mathfrak{x}(s_0) + t\mathfrak{x}'(s_0).$$

Die Tangente ist jene Gerade der die Kurve am nächsten ist. Sie ist die Gerade durch "zwei benachbarte" Kurvenpunkte.

Was ist nun die Ebene welcher die Kurve am nächsten ist, d.h. die Ebene durch "drei benachbarte" Kurvenpunkte?

Dazu betrachte drei Kurvenpunkte $\mathfrak{x}(s), \mathfrak{x}(s+h), \mathfrak{x}(s+k)$. Nach der Taylorschen Formel gilt

$$\mathfrak{x}(s+h) = \mathfrak{x}(s) + h\mathfrak{x}'(s) + \frac{h^2}{2}\mathfrak{x}''(s) + \frac{h^3}{6}\mathfrak{x}'''(s+\theta_1 h),$$

$$\mathfrak{x}(s+k) = \mathfrak{x}(s) + k\mathfrak{x}'(s) + \frac{k^2}{2}\mathfrak{x}''(s) + \frac{k^3}{6}\mathfrak{x}'''(s+\theta_2 k),$$

mit gewissen $\theta_1, \theta_2 \in (0,1)$. Die gesuchte Ebene enthält also die Vektoren

$$\mathfrak{u} = \frac{1}{h}[\mathfrak{x}(s+h) - \mathfrak{x}(s)] = \mathfrak{x}'(s) + \frac{h}{2}\mathfrak{x}''(s) + \frac{h^2}{6}\mathfrak{x}'''(s+\theta_1 h),$$

$$\mathfrak{v} = \frac{1}{k}[\mathfrak{x}(s+k) - \mathfrak{x}(s)] = \mathfrak{x}'(s) + \frac{k}{2}\mathfrak{x}''(s) + \frac{k^2}{6}\mathfrak{x}'''(s+\theta_1 k),$$

$$\mathfrak{w} = \frac{2}{k-h}(\mathfrak{v} - \mathfrak{u}) = \mathfrak{x}''(s) + \frac{1}{3}\frac{k^2\mathfrak{x}'''(s+\theta_2 k) - h^2\mathfrak{x}'''(s+\theta_1 h)}{k-h}.$$

Die Grenzlage dieser Ebene für $h, k \to 0$ enthält also die Vektoren

$$\lim_{h,k\to 0} \mathfrak{u} = \mathfrak{x}'(s), \quad \lim_{h,k\to 0} \mathfrak{w} = \mathfrak{x}''(s).$$

Die Ebene aufgespannt von den Vektoren $\mathfrak{x}'(s)$ und $\mathfrak{x}''(s)$ nennt man die *Schmiegeebene* der Kurve im Punkt $\mathfrak{x}(s)$. Das macht natürlich nur dann Sinn, wenn $\mathfrak{x}'$ und $\mathfrak{x}''$ linear unabhängig sind, insbesondere also $\mathfrak{x}''(s) \neq 0$ ist, was wir im folgenden immer voraussetzen wollen. Die Bedingung $\mathfrak{x}''(s) \neq 0$ bedeutet, daß die "drei benachbarten" Kurvenpunkte nicht auf einer Geraden liegen. In diesem Fall spannen $\mathfrak{x}'(s)$ und $\mathfrak{x}''(s)$ sicher eine Ebene auf, denn wegen $\mathfrak{x}'(s)^2 = 1$ folgt

$$0 = \frac{d}{ds}(\mathfrak{x}' \cdot \mathfrak{x}') = 2(\mathfrak{x}'' \cdot \mathfrak{x}'),$$

d.h. $\mathfrak{x}'(s)$ und $\mathfrak{x}''(s)$ sind orthogonal also insbesondere linear unabhängig.

Man bezeichnet $\mathfrak{x}'(s)$ als den *Tangentialvektor* und schreibt $\mathfrak{t}(s)$, den normierten Vektor in Richtung $\mathfrak{x}''(s)$ bezeichnet man als *Hauptnormalenvektor* und schreibt

$$\mathfrak{n}(s) := \frac{\mathfrak{x}''(s)}{\sqrt{\mathfrak{x}''(s)^2}}.$$

Der Vektor $\mathfrak{b} := \mathfrak{t} \times \mathfrak{n}$ ergänzt $\mathfrak{t}, \mathfrak{n}$ zu einem positiv orientierten Dreibein. Man nennt ihn den *Binormalenvektor*.

Das von den Vektoren $\mathfrak{t}, \mathfrak{n}, \mathfrak{b}$ gebildete Dreibein nennt man das *begleitende Dreibein*. Es bewegt sich in unserer Vorstellung mit dem Punkt $\mathfrak{x}(s)$ längs der Kurve. Man kann sich aber auch vorstellen, daß Dreibein beim Kurvenpunkt auszureißen und stattdessen im Ursprung anzuheften. Dann zeichnen die Vektoren $\mathfrak{t}, \mathfrak{n}, \mathfrak{b}$ wenn der Parameter s läuft, drei Kurven auf die Einheitssphäre ($\|\mathfrak{t}\| = \|\mathfrak{n}\| = \|\mathfrak{b}\| = 1$). Man nennt diese Kurven das *sphärische Tangenbild* (bzw. *sphärische Normalenbild* oder *sphärische Binormalenbild*).

Wie sieht der Tangentialvektor an diesen Kurven aus? Aber nicht bezüglich des normalen Koordinatensystems, das wäre einfach, aber dieses Koordinatensystem hat ja mit der Kurve nichts zu tun. Man hat zu jedem Parameter ein "kanonisches" Koordinatensystem, nämlich das Dreibein $\mathfrak{t}(s), \mathfrak{n}(s), \mathfrak{b}(s)$.

4.6.5 Satz (Frenetsche Formeln). *Es gibt zwei Funktionen* $\frac{1}{\rho(s)} > 0$ *und* $\frac{1}{\tau(s)}$, *die* Krümmung *bzw.* Torsion *der Kurve* $\mathfrak{x} = \mathfrak{x}(s)$, *sodaß gilt*

$$\begin{array}{lclcccc} \mathfrak{t}'(s) & = & & & \frac{1}{\rho(s)}\mathfrak{n}(s) & & \\ \mathfrak{n}'(s) & = & -\frac{1}{\rho(s)}\mathfrak{t}(s) & & + & & \frac{1}{\tau(s)}\mathfrak{b}(s) \\ \mathfrak{b}'(s) & = & & & -\frac{1}{\tau(s)}\mathfrak{n}(s) & & \end{array} \tag{4.10}$$

Beweis. Es gibt Zahlen x_{ij} mit

$$\begin{aligned} \mathfrak{t}' &= x_{11}\mathfrak{t} + x_{12}\mathfrak{n} + x_{13}\mathfrak{b} \\ \mathfrak{n}' &= x_{21}\mathfrak{t} + x_{22}\mathfrak{n} + x_{23}\mathfrak{b} \\ \mathfrak{b}' &= x_{31}\mathfrak{t} + x_{32}\mathfrak{n} + x_{33}\mathfrak{b} \end{aligned} \tag{4.11}$$

Da $\mathfrak{t}, \mathfrak{n}, \mathfrak{b}$ ein Orthonormalsystem bilden, können wir die Koeffizienten durch Bildung von Skalarprodukten ausrechnen. Die erste Gleichung ist bereits bekannt, denn nach Definition gilt $\mathfrak{t}' = \mathfrak{x}'' = \sqrt{\mathfrak{x}'' \cdot \mathfrak{x}''}\,\mathfrak{n}$. Also ist $x_{11} = 0, x_{12} = \sqrt{\mathfrak{x}'' \cdot \mathfrak{x}''}, x_{13} = 0$. Für die zweite und dritte Gleichung erhält man

$$x_{21} = \mathfrak{n}' \cdot \mathfrak{t}, x_{22} = \mathfrak{n}' \cdot n, x_{23} = \mathfrak{n}' \cdot \mathfrak{b}, x_{31} = \mathfrak{b}' \cdot \mathfrak{t}, x_{32} = \mathfrak{b}' \cdot n, x_{33} = \mathfrak{b}' \cdot \mathfrak{b}.$$

Wegen $\|\mathfrak{n}\| = \|\mathfrak{b}\| = 1$ folgt $\mathfrak{n}' \cdot \mathfrak{n} = \mathfrak{b}' \cdot \mathfrak{b} = 0$. Wegen $\mathfrak{n} \cdot \mathfrak{t} = 0$ folgt $\mathfrak{n}' \cdot \mathfrak{t} = -\mathfrak{n} \cdot \mathfrak{t}' = -\mathfrak{t}' \cdot \mathfrak{n} = -x_{12}$. Genauso erhält man aus $\mathfrak{b} \cdot \mathfrak{b} = 0$ und daher $x_{23} = -x_{32}$. Setze

$x_{12} =: \frac{1}{\rho(s)}, x_{23} = \frac{1}{\tau(s)}$.

❑

Die erste Zeile der Frenetschen Formeln gilt da wir als Parameter die Bogenlänge verwenden. Die anderen folgen da das begleitende Dreibein ein Orthonormalsystem bildet, denn daher ist die Koeffizientenmatrix schiefsymmetrisch.

Die Frenetschen Formeln für die Krümmung und Torsion bzw. für $\rho(s)$ und $\tau(s)$ heißen die *natürlichen Gleichungen* der Kurve. Denn sie enthalten nur noch geometrische Größen der Kurve (nämlich die Krümmung, Torsion und Bogenlänge).

Man kann zeigen, dass zu jeder Wahl von $\frac{1}{\rho(s)}, \frac{1}{\tau(s)}$ eine (bis auf Bewegungen eindeutige) Kurve mit dieser vorgegebenen Krümmung und Torsion existiert.

4.6.6 Bemerkung.

(*i*) Man bezeichnet die Krümmung mit $\frac{1}{\rho}$, denn ρ bedeutet den sogenannten Krümmungsradius, das ist der Radius jenes Kreises in der von $\mathfrak{t}$ und $\mathfrak{n}$ aufgespannten Ebene welcher die Kurve am besten annähert. Um dies einzusehen, nimmt man wieder drei Kurvenpunkte, legt einen Kreis durch, und läßt sie in einen einzigen Punkt zusammenwandern.

(*ii*) Das Bogenelement des sphärischen Tangentenbildes berechnet sich als

$$\sqrt{\mathfrak{t}'(s) \cdot \mathfrak{t}'(s)} = \sqrt{\mathfrak{x}''(s) \cdot \mathfrak{x}''(s)} = \frac{1}{\rho(s)},$$

jenes des sphärischen Binormalenbildes als

$$\sqrt{\mathfrak{b}'(s) \cdot \mathfrak{b}'(s)} = \sqrt{\frac{1}{\tau^2}\mathfrak{n}(s)} = \frac{1}{|\tau|}.$$

Krümmung bzw. Torsion beschreibt also das Verhältnis des Bogenelementes des sphärischen Tangenten- bzw. Binormalenbildes zum Bogenelement der Kurve, d.h. wie schnell sich die Tangente bzw. Schmiegeebene verdreht.

Literaturverzeichnis

[EL] K.Endl,W.Luh: *Analysis I-III,* Aula Verlag, Wiesbaden 1986.

[F] G.M.Fichtenholz: *Differential- und Integralrechnung I-III,* Deutscher Verlag der Wissenschaften, Berlin 1964.

[GO] B.Gelbaum,J.Olmsted: *Counterexamples in Analysis,* Holden-Day Inc., SanFrancisco 1964.

[GL] H.Grauert, I.Lieb: *Differential- und Integralrechnung I-III,* Springer Verlag, Heidelberg 1967.

[EL] H.Heuser: *Lehrbuch der Analysis 1,2,* Teubner Verlag, Stuttgart 1989.

[EL] S.Lang: *A first course in calculus,* Springer Verlag, Heidelberg 1986.

[R] W.Rudin: *Principles of Mathematical Analysis,* McGraw-Hill, New York 1953.

[R] W.Walter: *Ordinary Differential Equations,* Springer, New York 1998.

Index

Überdeckung
offene, 8
Teil-, 8
1.Mittelwertsatz der Differentialrechnung, 60

Abelsches Kriterium, 36
Ableitung
einer Funktion, 58
höhere, 58
im Punkt x, 53
allgemeine Potenz, 46

Banachraum, 77
begleitende Dreibein, 115
Binormalenvektor, 115
Bogenlänge, 112

Cosinus, 44

Darstellung
explizite, 112
implizite, 112
Parameter-, 112
differenzierbar
n-mal, 58
im Punkt x, 53
stetig, 58
Dirichlet Kriterium, 36
Dreiecksungleichung, 76
Dreiecksungleichung für Integrale, 97

Eulersche Identität, 44
Eulersche Zahl, 47
Exponent, 46
Extremum
absolutes, 59
lokales, 59

Fixpunkt, 78
Formel von de Moivre, 44
Frenetsche Formeln, 115
Funktion
beschränkte, 33
gerade, 74
monoton fallend, 30
monoton wachsende, 29
stetige, 21
streng monoton wachsende, 30
ungerade, 74
unstetige, 28
Funktionen
trigonometrische, 44

Gaußklammer, 23
gleichgradig stetig, 39
gleichmäßig beschränkt, 39
Grenzwert
einer Funktion, 23
linksseitiger, 28
rechtsseitiger, 28

Häufungspunkt, 3
Hauptnormalenvektor, 115
Hauptsatz der Differential-Integralrechnung, 100
hebbare Unstetigkeit, 29

Integral, 88
integrierbar
Riemann-, 88
Intervall, 6
Isometrie, 6

Keplersche Faßregel, 109
Kettenregel, 56
kompakt, 8
Kontraktion, 78
konvergent
gleichmäßig, 34
punktweise, 32
Konvergenzradius, 37
Krümmung, 115
Kugel, 1

Lagrangeschen Interpolationsformel, 107
Leibnizsche Formel, 58
Limes inferior, 17
Limes superior, 17
Lipschitz-Bedingung, 78
Logarithmus
  natürlicher, 46

Maximum
  absolutes, 59
  lokales, 59
Menge
  abgeschlossene, 3
  dichte, 5
  offene, 2
  zusammenhängende, 6
Metrik
  Supremums-, 33
Minimum
  absolutes, 59
  lokales, 59
Mittelwertsatz, 60
  1.der Integralrechnung, 102
  2.der Differentialrechnung, 61
  2.der Integralrechnung, 103
  verallgemeinerter, 61
  Verallgemeinerter 1.der Integralrechnung, 102
monoton fallend, 30
monoton wachsend, 29

natürlichen Gleichungen, 116
Newton-Verfahren, 79
Newtonsche Interpolationsformel, 68
Norm, 76

Parameterdarstellung, 112
Partialbruchzerlegung, 105
Partielle Integration, 102
Partition, 86
periodisch, 45
Pi, 41
Potenzreihen, 36
Produktregel, 56
punktweise beschränkt, 39

Quotientenregel, 56

Raum
  normierter, 76
Regel von de L'Hospital, 61
Reihe
  Potenzreihen, 36
Riemann Integral, 88
Riemannsche Obersumme, 87
Riemannsche Untersumme, 87
Riemannsche Zwischensumme, 93

Satz
  1.Mittelwertsatz der Differentialrechnung, 60
  2.Mittelwertsatz der Differentialrechnung, 61
  Abelscher Grenzwertsatz, 38
  Fixpunktsatz von Banach, 78
  Frenetsche Formeln, 115
  Fundamentalsatz der Algebra, 47
  Hauptsatz der Differential-Integralrechnung, 100
  Kettenregel, 56
  Mittelwertsatz, 60
  Regel von de L'Hospital, 61
  Taylorscher Lehrsatz, 71
  Verallgemeinerter Mittelwertsatz, 61
  von Arzela-Ascoli, 39
  von Bolzano-Weierstraß, 13
  von Heine-Borel, 10
  von Rolle, 60
  Zwischenwertsatz, 26
Schmiegeebene, 115
Simpson-Formel, 109
Sinus, 44
sphärisches
  Binormalenbild, 115
  Normalenbild, 115
  Tangenbild, 115
Sprungstelle, 29
stückweise stetig, 92
Stammfunktion, 100
stetig, 21
  an der Stelle x, 21
  gleichmäßig, 27
streng monoton fallend, 30
Substitutionsregel, 101
sukzessiven Approximation, 79
Summensätze, 45
Supremumsnorm, 77
  gewichtete, 78

Tangente, 112
Tangentenverfahren, 79
Tangentialvektor, 115
Taylorreihe, 72
Teilüberdeckung, 8
Torsion, 115
Trapezformel, 109

Unstetigkeit
1.Art, 28
2.Art, 28
hebbare, 29
Sprungstelle, 29

Verfeinerung, 86

Weierstraß Kriterium, 36

Zwischenwertsatz, 26