1 Topologie in metrischen Räumen

1.1 Metrische und normierte Räume

Das Ziel dieses Abschnitts ist es, die Erkenntnisse über Folgen (und Reihen) reeller Zahlen aus dem vergangenen Semester zu verallgemeinern. Insbesondere wollen wir Aussagen über die Konvergenz von Folgen im $\mathbb{R}^n$ formulieren und beweisen können. Wir wählen jedoch den allgemeineren Ansatz des metrischen Raumes, einer abstrakten mathematischen Struktur auf der ein Abstandsbegriff definiert ist. Die Theorie wird dadurch nicht komplizierter (nur abstrakter) aber wir können später mit den Werkzeugen die wir uns erarbeiten werden auch zum Beispiel die Konvergenz von Funktionenfolgen untersuchen.

Definition 1.1 (Metrischer Raum). Ein metrischer Raum (X, d) ist ein Paar bestehend aus einer Menge X zusammen mit einer Abbildung $d : X \times X \to \mathbb{R}$, die als Metrik bezeichnet wird und die folgenden Bedingungen für alle $x, y, z \in X$ erfüllt:

$$\begin{aligned} d(x,y) &\geq 0 \text{ (Positivität)} \\ d(x,y) = 0 &\Longleftrightarrow x = y \text{ (Definitheit)} \\ d(x,y) &= d(y,x) \text{ (Symmetrie)} \\ d(x,y) &\leq d(x,z) + d(z,y) \text{ (Dreiecksungleichung).} \end{aligned}$$

Die Elemente eines metrischen Raumes werden oft als Punkte bezeichnet. Die Bezeichnungen „Raum“ und „Punkt“ rühren daher, dass es sich hier um Objekte der Geometrie (in einem sehr weit gefassten Sinn des Wortes) handelt. Ihnen kommt keine tiefere Bedeutung zu. Ein einzelner Punkt eines abstrakten metrischen Raumes besitzt keine besonderen Eigenschaften außer der Element eben dieses Raumes zu sein. Interessante Eigenschaften treten erst durch die Struktur (die Metrik mit ihren Eigenschaften) auf, die verschiedene Punkte des Raumes zueinander in Beziehung setzt.

In vielen für uns wichtigen Beispielen besitzt die Menge X außer der metrischen noch eine damit kompatible Vektorraumstruktur:

Definition 1.2 (Normierter Raum). Ein normierter Raum $(V, \|\cdot\|)$ ist ein Paar bestehend aus einem Vektorraum V über dem Körper $\mathbb{K}$ der reellen oder komplexen Zahlen und einer Abbildung $\|\cdot\| : V \to \mathbb{R}$, die als Norm bezeichnet wird und die folgenden Bedingungen für alle $x, y \in V$ erfüllt:

$$\begin{aligned} \|x\| &\geq 0 \text{ (Positivität)} \\ \|x\| = 0 &\Longleftrightarrow x = 0 \text{ (Definitheit)} \\ \|\lambda x\| &= |\lambda| \, \|x\| \text{ (absolute Homogenität)} \\ \|x+y\| &\leq \|x\| + \|y\| \text{ (Dreiecksungleichung).} \end{aligned}$$

Jeder normierte Raum erhält die Struktur eines metrischen Raumes indem wir definieren $d(x, y) := \|x - y\|$. Man überprüft leicht, dass die so definierte Abbildung eine Metrik ist, sie wird als die von der Norm induzierte Metrik bezeichnet. Das für uns bei Weitem wichtigste Beispiel:

Beispiel 1.3 ($\mathbb{R}^n$ mit euklidischer Norm). Definiere die euklidische Norm

$$\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$$
$$\|(x_1, \dots, x_n)\| := \sqrt{x_1^2 + \dots + x_n^2}.$$

Dann ist $(\mathbb{R}^n, \|\cdot\|)$ ein normierter Raum.

Im Fall des komplexen Vektorraums $\mathbb{C}^n$ müssen wir die Definition der Norm leicht verändern:

$$\|(x_1, \dots, x_n)\| := \sqrt{\overline{x_1}x_1 + \dots + \overline{x_n}x_n}.$$

Beispiel 1.4. (i) Die Abbildung

$$(x_1, \dots, x_n) \mapsto \max\{|x_1|, \dots, |x_n|\}$$

definiert ebenfalls eine Norm auf dem $\mathbb{R}^n$. Bereits auf dem $\mathbb{R}^n$ sind also verschiedene Normen möglich.

(ii) Sei M eine beliebige Menge. Dann wird durch

$$d(x, y) := \begin{cases} 0 & \text{falls } x = y \\ 1 & \text{sonst} \end{cases}$$

eine Metrik definiert, die als diskrete Metrik bezeichnet wird.

(iii) Seien (X, d_X) und (Y, d_Y) metrische Räume. Dann ist $(X \times Y, d_{X\times Y})$ mit der Metrik

$$d_{X\times Y}((x_1, y_1), (x_2, y_2)) := d_X(x_1, x_2) + d_Y(y_1, y_2) \tag{1.1}$$

ein metrischer Raum.

Analog ist $\|(v, w)\|_{V\times W} := \|v\|_V + \|w\|_W$ eine Norm auf dem direkten Produkt $V \times W$.

Bemerkung 1.5.

- Sei (X, d) ein metrischer Raum und $Y \subset X$ dann ist $(Y, d|_{Y\times Y})$ ebenfalls ein metrischer Raum.
- Analog sind Untervektorräume (nicht beliebige Teilmengen) normierter Vektorräume ebenfalls normierte Vektorräume.
- Zu einer vorgegebenen Menge gibt es keine „natürliche“ Metrik. Man kann vielmehr auf ein und der selben Menge unterschiedliche Metriken (im Fall von Vektorräumen auch unterschiedliche Normen) mit ganz unterschiedlichen Eigenschaften betrachten. Im konkreten Fall wählt man sich dann die Metrik, die dem Problem am besten angemessen ist.
- Die „Minkowski-Metrik“ die in der Relativitätstheorie eine zentrale Rolle spielt ist keine Metrik im hier eingeführten Sinne.

Definition 1.6 (Kugeln und Beschränktheit). Sei (X, d) ein metrischer Raum, $x \in X$ und $r > 0$. Die Menge

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

heißt offene Kugel mit Radius r um x (manchmal auch r-Kugel um x) und die Menge

$$K_r(x) := \{y \in X |\ d(y, x) \leq r\}$$

heißt abgeschlossene Kugel mit Radius r um x.

Eine Teilmenge M von X heißt beschränkt, wenn sie in irgendeiner Kugel enthalten ist.

Die Bezeichnungen „offen“ und „abgeschlossen“ werden sich später erschließen.

1.2 Folgen, Reihen und Grenzwerte

Mittels der soeben eingeführten Begriffe können wir nun Konvergenz von Folgen definieren.

Definition 1.7 (Folgen und Konvergenz). Sei (X, d) ein metrischer Raum. Eine Folge in X ist eine Abbildung von $\mathbb{N}$ nach X. Wir schreiben dafür $(x_n)_{n\in\mathbb{N}}$ oder noch kürzer (x_n).

Die Folge $(x_n)_{n\in\mathbb{N}}$ heißt konvergent gegen $x \in X$, wenn für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass

$$d(x_n, x) < \epsilon \text{ für alle } n > N.$$

Der Punkt x heißt in diesem Fall der Grenzwert der Folge (x_n).

Eine Folge die nicht konvergent ist heißt divergent.

Bemerkung 1.8. Auf den reellen Zahlen stimmt die euklidische Metrik mit der Betragsfunktion überein. Man überzeuge sich davon, dass in diesem Fall die neue Definition (Konvergenz von Folgen im metrischen Raum) mit der alten Definition (Konvergenz von reellen Zahlenfolgen) übereinstimmt. Die obige Definition ist also eine echte Verallgemeinerung der Konvergenz von reellen Zahlenfolgen.

Satz 1.9. *Sei (x_n) eine Folge in einem metrischen Raum (X, d) und $x \in X$. Dann sind äquivalent:*

(i) Die Folge (x_n) konvergiert gegen x.

(ii) Für jedes $\epsilon > 0$ enthält $B_\epsilon(x)$ alle bis auf endlich viele Folgenglieder.

(iii) Die Folge (reeller Zahlen) $(d(x_n, x))$ ist eine Nullfolge.

Beweis. Die Aussage (ii) ist lediglich eine Umformulierung der Definition der Konvergenz. Wir zeigen $(i) \iff (iii)$. Es gilt $d(x_n, x) = |d(x_n, x) - 0|$. Damit ist $d(x_n, x) < \epsilon$, genau dann wenn $|d(x_n, x) - 0| < \epsilon$. Die Folge (x_n) konvergiert gegen x, genau dann wenn $d(x_n, x)$ gegen 0 konvergiert. □

Satz 1.10. *Sei $(x^{(n)}) = (x_1^{(n)}, \dots, x_d^{(n)})$ eine Folge in $\mathbb{R}^d$ und $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. Dann konvergiert $(x^{(n)})$ gegen x, genau dann wenn für jedes $i \in \{1, \dots, d\}$ die Folge $(x_i^{(n)})$ gegen x_i konvergiert.*

Beweis. Man überzeugt sich leicht, dass $|y_i| \leq \|y\|$ für alle $y \in \mathbb{R}^n$ und alle $i \in \{1, \cdots, d\}$ gilt. Wenn also $(x^{(n)})$ gegen x konvergiert, dann konvergiert nach Satz 1.9 $\left\|x^{(n)} - x\right\|$ gegen 0 und damit auch $\left|x_i^{(n)} - x_i\right| \leq \left\|x^{(n)} - x\right\|$.

Es gelte andererseits $x_i^{(n)} \to x_i$ für alle $i \in \{1, \dots, d\}$. Sei $\epsilon > 0$ Dann existiert $N \in \mathbb{N}$, so dass $\left|x_i^{(n)} - x_i\right| \leq \frac{\epsilon}{\sqrt{d}}$ für alle $n > N$ und alle $i \in \{1, \dots d\}$. Es gilt folglich für $n > N$

$$\left\|x^{(n)} - x\right\| = \sqrt{\left(x_1^{(n)} - x_1\right)^2 + \cdots + \left(x_d^{(n)} - x_d\right)^2} \leq \sqrt{\frac{\epsilon^2}{d} + \cdots + \frac{\epsilon^2}{d}} = \epsilon. \qquad \square$$

Satz 1.11. *(i) Jede Folge besitzt höchstens einen Grenzwert.*

(ii) Jede konvergente Folge ist beschränkt.

Beweis. Angenommen x und y seien zwei verschiedene Grenzwerte der Folge (x_n). Aus der Definitheit der Norm folgt dann $d(x, y) > 0$. Da (x_n) sowohl gegen x als auch gegen y konvergiert existiert N, so dass $d(x_n, x) < d(x, y)/2$ und $d(x_n, y) < d(x, y)/2$ für alle $n > N$. Damit gilt für jedes $n > N$, dass $d(x_n, x) + d(x_n, y) < d(x, y)$ was jedoch im Widerspruch zur Dreiecksungleichung steht.

Zur Beschränktheit: Wenn (x_n) eine Folge ist, die gegen x konvergiert, so existiert per definitionem ein $N \in \mathbb{N}$, so dass $d(x_n, x) < 1$ für alle $n > N$. Dann gilt $d(x_n, x) \leq \max\{d(x_1, x), \ldots, d(x_N, x), 1\}$ für jedes $n \in \mathbb{N}$. □

Definition 1.12 (Teilfolge und Häufungspunkt). Sei (X, d) ein metrischer Raum, (x_n) eine Folge in X und $n : \mathbb{N} \to \mathbb{N}$, $k \mapsto n_k$ eine streng monoton wachsende Folge natürlicher Zahlen. Dann heißt $(x_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge von (x_n).

Ein Punkt $x \in X$ heißt Häufungspunkt von (x_n), wenn es eine Teilfolge (x_{n_k}) gibt, die gegen x konvergiert.

Satz 1.13. *Ersatzlos gestrichen.*

Satz 1.14. *Sei $(V, \|\cdot\|)$ ein normierter Raum über dem Vektorraum $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$), $(x_n), (y_n)$ Folgen in V mit Grenzwerten x und y und (λ_n) eine Folge in $\mathbb{K}$ mit Grenzwert λ. Dann konvergiert die Folge $(x_n + y_n)$ gegen $x + y$ und die Folge $\lambda_n x_n$ gegen λx.*

Beweis. Übung □

Definition 1.15 (Cauchy-Folge und Vollständigkeit). Eine Folge (x_n) im metrischen Raum (X, d) heißt Cauchy-Folge (alternativ besitzt die Cauchy-Eigenschaft oder erfüllt die Cauchy-Bedingung), wenn für jedes $\epsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass

$$d(x_n, x_m) < \epsilon \text{ für alle } n, m > N.$$

Ein metrischer Raum (X, d) heißt vollständig, wenn jede Cauchy-Folge konvergent ist.

Satz 1.16. *Jede konvergente Folge ist eine Cauchy-Folge.*

Beweis. Sei (x_n) eine konvergente Folge im metrischen Raum (X, d). Sei $\epsilon > 0$. Da (x_n) konvergent ist, existiert ein Grenzpunkt $x \in X$ und ein $N \in \mathbb{N}$, so dass $d(x_n, x) < \epsilon/2$ für alle $n > N$. Dann gilt für alle $n, m > N$, dass

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Die Folge (x_n) ist eine Cauchy-Folge. □

Bemerkung 1.17.

- Um an Hand der Definition der Konvergenz zu überprüfen ob eine Folge konvergent ist, muss man ihren Grenzpunkt bereits kennen. Um zu überprüfen ob eine Folge eine Cauchy-Folge ist muss kein Grenzpunkt bekannt sein. Darin liegt die Bedeutung der Vollständigkeit, denn sie erlaubt es auf die Existenz eines Grenzpunktes zu schließen ohne diesen zu kennen.

- Ein vollständiger normierter Raum heißt Banachraum.

Satz 1.18. *Der $\mathbb{R}^d$ versehen mit der euklidischen Norm ist vollständig (also ein Banachraum).*

Beweis. Wir betrachten zunächst den Fall $d = 1$. Sei (x_n) eine Cauchy-Folge reeller Zahlen. Zunächst zeigen wir, dass (x_n) beschränkt ist. Wähle dazu mittels der Cauchy-Eigenschaft $N \in \mathbb{N}$, so dass $|x_n - x_m| < 1$ für alle $n, m \geq N$. Dann gilt für $m \geq N$

$$|x_m| \leq |x_m - x_N| + |x_N| < 1 + |x_N|$$

und damit ist

$$|x_m| \leq \max\{|x_1|, \cdots, |x_{N-1}|, |x_N| + 1\}$$

für alle m.

Sei nun x ein Häufungspunkt von x_n (existiert nach dem Satz von Bolzano-Weierstraß aus dem vergangenen Semester da die Folge x_n beschränkt ist). Wir zeigen $x_n \to x$. Sei dazu $\epsilon > 0$ und (x_{n_k}) eine Teilfolge die gegen x konvergiert. Wähle wiederum mittels der Cauchy-Eigenschaft ein $N \in \mathbb{N}$, so dass $|x_n - x_m| < \frac{\epsilon}{2}$ für alle $n, m > N$. Wähle nun $k \in \mathbb{N}$, so dass $n_k > N$ und $|x_{n_k} - x| < \frac{\epsilon}{2}$. Dann gilt für jedes $n > n_k$

$$|x_n - x| < |x_n - x_{n_k}| + |x_{n_k} - x| < \epsilon,$$

also konvergiert (x_n) gegen x.

Betrachte schließlich den Fall $d > 1$. Für beliebiges $y = (y_1, \ldots, y_d) \in \mathbb{R}^d$ gilt $|y_i| \leq \|y\|$ für jedes $i \in \{1, \ldots, d\}$. Daraus folgt, dass für eine Cauchy-Folge $(x^{(n)})$ in $\mathbb{R}^d$ jede der Koordinatenfolgen $(x_i^{(n)})$ eine Cauchy-Folge in $\mathbb{R}$ ist. Aus der Vollständigkeit der reellen Zahlen folgt damit, dass die Koordinatenfolgen konvergent sind und damit ist nach Satz 1.10 die Folge $x^{(n)}$ konvergent. □

Bemerkung 1.19. Im Beweis haben wir die folgenden Aussagen gezeigt (nur für reelle Folgen, sie gelten aber allgemein): Cauchy-Folgen sind beschränkt. Cauchy-Folgen mit einem Häufungspunkt sind konvergent.

Definition 1.20 (Reihen und (absolute) Konvergenz). Sei $(V, \|\cdot\|)$ ein normierter Vektorraum und (x_n) eine Folge in V. Bezeichne mit

$$s_n := \sum_{k=1}^{n} x_k$$

die n-te Partialsumme. Dann heißt die Folge (s_n) die zu (x_n) gehörige Reihe. Ist die Folge (s_n) konvergent, so heißt

$$\sum_{k=1}^{\infty} x_k := \lim_{n\to\infty} s_n$$

der Grenzwert der Reihe (auch Wert oder Summe der Reihe). Häufig verwendet man (formal nicht ganz korrekt) die Notation $\sum_{k=1}^{\infty} x_n$ für die Reihe selbst (egal ob diese konvergiert oder divergiert).

Die Reihe $\sum_{k=1}^{\infty} x_n$ heißt absolut konvergent, wenn $\sum_{k=1}^{\infty} \|x_n\|$ konvergiert.

Satz 1.21. *In Banachräumen sind absolut konvergente Reihen konvergent.*

Beweis. Sei $(V, \|\cdot\|)$ ein Banachraum und die Reihe $\sum_{n=1}^{\infty} x_n$ in V absolut konvergent. Es reicht zu zeigen, dass die Partialsummenfolge (s_n) eine Cauchy-Folge ist. Sei dazu $\epsilon > 0$. Bezeichne mit

$$t_n := \sum_{k=1}^{n} \|x_k\|$$

die Partialsummenfolge der Reihe $\sum_{n=1}^{\infty} \|x_n\|$. Diese ist monoton wachsend und konvergiert nach Voraussetzung. Insbesondere ist sie also Cauchy-Folge, das heißt es existiert $N \in \mathbb{N}$, so dass $|t_n - t_m| < \epsilon$ für alle $n, m > N$.

Dann gilt für $n > m > N$

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^{n} x_k \right\| \leq \sum_{k=m+1}^{n} \|x_k\| = t_n - t_m = |t_n - t_m| < \epsilon.$$

Die Partialsummenfolge (s_n) ist also eine Cauchy-Folge und damit konvergent. □

Theorem 1.22 (Umordnungssatz). *Sei $(V, \|\cdot\|)$ ein Banachraum. Sei $\sum_{n=1}^{\infty} x_n$ eine absolut konvergente Reihe und $\sigma : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Dann ist die umgeordnete Reihe $\sum_{n=1}^{\infty} x_{\sigma(n)}$ ebenfalls konvergent mit Grenzwert $\sum_{n=1}^{\infty} x_n$.*

Beweis. Sei $\epsilon > 0$. Sei (s_n) die zu (x_n) und (t_n) die zu $(\|x_n\|)$ gehörige Partialsummenfolge. Nach Voraussetzung ist (t_n) konvergent, also insbesondere Cauchy-Folge, das heißt es existiert N, so dass $\|t_n - t_m\| < \frac{\epsilon}{2}$ für $n, m \geq N$. Nach Satz 1.21 konvergiert (s_n), es existiert also s und ein (eventuell größeres) N, so dass $\|s_n - s\| < \frac{\epsilon}{2}$ für $n \geq N$. Setze

$A := \sigma^{-1}\{1, \ldots, N\}$ und $M := \max A$. Dann gilt für jedes $m > M$

$$\begin{aligned}
\left\| \sum_{k=1}^{m} x_{\sigma(k)} - s \right\| &= \left\| \sum_{k=1, k\in A}^{m} x_{\sigma(k)} + \sum_{k=1, k\notin A}^{m} x_{\sigma(k)} - s \right\| \\
&\leq \left\| \sum_{l=1}^{N} x_l - s \right\| + \sum_{k=1, k\notin A}^{m} \left\| x_{\sigma(k)} \right\| \\
&\leq \|s_N - s\| + \sum_{l=N+1}^{L} \|x_l\| \\
&< \frac{\epsilon}{2} + t_L - t_N < \epsilon.
\end{aligned}$$

Wobei $L \geq \max \sigma(\{1, \ldots m\})$ sein soll. Damit ist die Konvergenz gezeigt. □

Bemerkung 1.23. Da wir obigen Satz insbesondere auf die Reihe $\sum_{n=1}^{\infty} \left\| x_{\sigma(n)} \right\|$ anwenden können, ist die umgeordnete Reihe auch absolut konvergent.

1.3 Offene und abgeschlossene Mengen

Definition 1.24 (Innere Punkte und Randpunkte). Sei (X, d) ein metrischer Raum, $A \subset X$ und $x \in X$.

(i) Der Punkt x heißt innerer Punkt von A, wenn es ein $\epsilon > 0$ gibt, so dass $B_\epsilon(x) \subset A$,

(ii) Der Punkt x heißt Randpunkt von A, wenn für jedes $\epsilon > 0$ die Kugel $B_\epsilon(x)$ mindestens je einen Punkt aus A sowie aus $X \backslash A$ enthält.

(iii) Die Menge aller inneren Punkte von A heißt das Innere von A und wird mit $\mathring{A}$ bezeichnet.

(iv) Die Menge aller Randpunkte von A heißt Rand von A und wird mit ∂A bezeichnet.

(v) Die Menge $A \cup \partial A$ heißt Abschluss von A und wird mit $\overline{A}$ bezeichnet

Bemerkung 1.25.

- Aus der Definition ist ersichtlich, dass ein innerer Punkt von A stets selbst in A liegt es gilt also $\mathring{A} \subset A$.
- Randpunkte von A können in A liegen, müssen es jedoch nicht.
- Jeder Punkt in A ist entweder ein innerer Punkt oder ein Randpunkt von A (niemals beides). Es gilt also $A \subset \mathring{A} \cup \partial A$ und $\mathring{A} \cap \partial A = \emptyset$.
- Man überzeugt sich leicht, dass der Rand von A und der Rand von $X \backslash A$ übereinstimmen.

Satz 1.26. *Seit (X, d) ein metrischer Raum, $A \subset X$ und $x \in X$ innerer Punkt oder Randpunkt von A. Dann gibt es eine Folge (x_n) in A die gegen x konvergiert.*

Beweis. Falls $x \in A$ (also insbesondere falls x innerer Punkt von A ist), so liegt die konstante Folge $(x)_{n \in \mathbb{N}}$ in A (und ist natürlich konvergent gegen x). Sei also $x \notin A$ (und damit Randpunkt von A). Für jedes $n \in \mathbb{N}$ wähle einen Punkt $x_n \in B_{\frac{1}{n}}(x) \cap A$. Diese Menge ist nicht leer, da x Randpunkt von A ist. Wir zeigen nun $x_n \to x$. Sei dazu $\epsilon > 0$. Dann existiert $N \in \mathbb{N}$, so dass $\frac{1}{N} < \epsilon$. Sei $n > N$ dann gilt nach der Definition der Folge (x_n), dass $x_n \in B_{\frac{1}{n}}(x) \subset B_{\frac{1}{N}}(x)$ also

$$d(x_n, x) < \frac{1}{N} < \epsilon.$$

Damit konvergiert die Folge gegen x. □

Definition 1.27 (Offene und abgeschlossene Mengen). Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt

(i) offen wenn jeder ihrer Punkte innerer Punkt von A ist,

(ii) abgeschlossen wenn die Menge A jeden ihrer Randpunkte enthält.

Definition 1.28 (Umgebung). Sei (X, d) ein metrischer Raum und $x \in X$. Eine Menge $U \subset X$ heißt Umgebung von x genau dann wenn ein $\epsilon > 0$ existiert, so dass $B_\epsilon(x) \subset U$ ist.

Bemerkung 1.29. Aus der Definition ist ersichtlich, dass U genau dann Umgebung von x ist, wenn x innerer Punkt von U ist. Insbesondere sind offene Teilmengen Umgebungen für jedes ihrer Elemente.

Satz 1.30. *Eine Teilmenge A eines metrischen Raumes (X, d) ist offen genau dann, wenn ihr Komplement $X \backslash A$ abgeschlossen ist.*

Beweis. Per definitionem ist die Menge A offen, genau dann wenn sie nur aus inneren Punkten von A besteht. Da jeder Punkt in A entweder innerer Punkt oder Randpunkt von A ist, ist das der Fall genau dann wenn A keinen Randpunkt von A enthält. Da der Rand von A und der Rand von $X \backslash A$ übereinstimmen, gilt das genau dann wenn A keinen Randpunkt von $X \backslash A$ enthält also genau dann wenn $X \backslash A$ alle Randpunkte von $X \backslash A$ enthält. Das ist per definitionem der Fall genau dann wenn $X \backslash A$ abgeschlossen ist. □

Beispiel 1.31. Sei (X, d) ein metrischer Raum, $x \in X$ und $r > 0$. Die offene Kugel $B_r(x) = \{y \in X |\ d(x, y) < r\}$ ist eine offene, die abgeschlossene Kugel $\{y \in X |\ d(x, y) \leq r\}$ eine abgeschlossene Menge.

Beweis. Übung □

Bemerkung 1.32. Ein Spezialfall des vorhergehenden Beispiels sind Intervalle der Form (a, b) – offene Intervalle – und $[a, b]$ – abgeschlossene Intervalle – für $a < b$ Elemente der reellen Zahlen.

Satz 1.33. *Sei (X, d) ein metrischer Raum und $A \subset X$. Die Menge A ist abgeschlossen genau dann, wenn für jede konvergente Folge (x_n) in A auch ihr Grenzwert in A liegt.*

Beweis. Sei zunächst A abgeschlossen, $(x_n) \subset A$ eine Folge und x ihr Grenzwert. Für beliebiges $\epsilon > 0$ existiert $n \in \mathbb{N}$, so dass $x_n \in B_\epsilon(x)$. Das heißt, in jeder ϵ-Kugel liegt mindestens ein Punkt von A, x ist also innerer Punkt von A oder Randpunkt von A. Innere Punkte von A liegen stets in A, Randpunkte von A liegen ebenfalls in A da A abgeschlossen ist.

Erfülle nun A die Eigenschaft, dass der Grenzwert jeder konvergenten Folge in A wieder in A liegt. Sei $x \in \partial A$. Dann existiert nach Satz 1.26 eine Folge (x_n) in A die gegen x konvergiert. Nach Voraussetzung ist dann $x \in A$. Damit haben wir $\partial A \subset A$ gezeigt, A ist also abgeschlossen. □

Satz 1.34. *Sei (X, d) ein metrischer Raum und τ die Menge aller offenen Teilmengen von X. Es gelten:*

(i) die leere Menge $\emptyset$ und der ganze Raum X sind in τ,

(ii) die Vereinigung beliebig vieler Mengen aus τ ist wiederum in τ,

(iii) der Schnitt endlich vieler Mengen aus τ ist wiederum in τ.

Beweis. Übung □

Bemerkung 1.35.

- Das System von Teilmengen τ in obigem Satz bezeichnet man als die von der Metrik d induzierte Topologie. (Topologieen sind Strukturen – noch allgemeiner als metrische Räume – die es erlauben über Begriffe wie Konvergenz und Stetigkeit zu reden)
- Der obige Satz impliziert für abgeschlossene Mengen mittels Satz 1.30 und der De Morganschen Regeln:
  (i) die leere Menge $\emptyset$ und der ganze Raum X sind abgeschlossen,
  (ii) der Schnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen,
  (iii) die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.
- Unendliche Schnitte offener Mengen müssen nicht mehr offen, unendliche Vereinigungen abgeschlossener Mengen nicht mehr abgeschlossen sein wie die folgenden Beispiele zeigen

$$\bigcap_{n\in\mathbb{N}} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$$
$$\bigcup_{n\in\mathbb{N}} [1/n, 2] = (0, 2]\,.$$

Satz 1.36. *Sei (X, d) ein metrischer Raum und $A \subset X$. Die Menge $\mathring{A}$ ist offen, die Mengen ∂A und $\overline{A}$ sind abgeschlossen.*

Beweis. Sei zunächst $x \in \mathring{A}$. Wir wissen, dass x innerer Punkt von A ist und wir wollen zeigen, dass es innerer Punkt von $\mathring{A}$ ist. Nach Voraussetzung existiert $\epsilon > 0$, so dass $B_\epsilon(x) \subset A$. Sei nun $y \in B_\epsilon(x)$. Nach Beispiel 1.31 ist $B_\epsilon(x)$ offen, das heißt es existiert $\delta > 0$, so dass $B_\delta(y) \subset B_\epsilon(x) \subset A$ ist. Damit ist y innerer Punkt von A und wir haben gezeigt, dass $B_\epsilon(x) \subset \mathring{A}$. Dann ist aber x innerer Punkt von $\mathring{A}$ und $\mathring{A}$ ist offen.

Setze nun $B := X \backslash A$. Dann gilt (Bemerkung 1.25) $\partial A = \partial B$. Damit erhält man

$$X = A \cup B = \mathring{A} \cup \partial A \cup \mathring{B} \cup \partial B = \mathring{A} \cup \partial A \cup \mathring{B} \tag{1.2}$$

Da A und B disjunkt sind, gilt das auch für $\mathring{A} \subset A$ und $\mathring{B} \subset B$. Außerdem ist $\partial A = \partial B$ disjunkt zu $\mathring{A}$ und $\mathring{B}$ (nochmals Bemerkung 1.25). Wir haben also die disjunkte Zerlegung $X = \mathring{A} \cup \partial A \cup \mathring{B}$. Damit sind die Mengen $\mathring{A} \cup \partial A = X \backslash \mathring{B}$ und $\partial A = X \backslash (\mathring{A} \cup \mathring{B})$ abgeschlossen. □

Satz 1.37. *Sei (X, d) ein metrischer Raum und $A \subset X$. Es gelten*

$$\bigcup_{O \text{ offen},\ O \subset A} O = \mathring{A} \text{ und}$$

$$\bigcap_{C \text{ abgeschlossen},\ A \subset C} C = \overline{A}$$

das heißt $\mathring{A}$ ist die größte offene Teilmenge von A, $\overline{A}$ ist die kleinste abgeschlossene Teilmenge von X die A enthält.

Beweis. Da $\mathring{A}$ offen und in A enthalten ist gilt in jedem Fall

$$\bigcup_{O \subset A \text{ offen}} O \supset \mathring{A}.$$

Sei nun $O \subset A$ offen und $x \in O$. Dann existiert $\epsilon > 0$, so dass $B_\epsilon(x) \subset O \subset A$ also ist x innerer Punkt von A. Da das für jedes $x \in O$ gilt haben wir gezeigt, dass $O \subset \mathring{A}$ ist. Da das für jedes offene $O \subset A$ gilt ist auch

$$\bigcup_{O \subset A \text{ offen}} O \subset \mathring{A}.$$

Die zweite Aussage erhält man nun durch Komplementbildung. □

Satz 1.38. *Sei (X, d) ein vollständiger metrischer Raum und $A \subset X$ abgeschlossen. Dann ist (A, d) ebenfalls vollständig.*

Beweis. Übung □

Bemerkung 1.39. Es ist wichtig sich bewusst zu machen, dass topologische Eigenschaften – Konvergenz einer Folge (x_n), Offenheit/Abgeschlossenheit einer Menge A (später Stetigkeit von Funktionen, Kompaktheit von Mengen ...) – nicht nur von der Folge (x_n) oder Menge A abhängen, sondern auch vom umgebenden metrischen Raum (X, d).

So ist die Folge $\left(\frac{1}{n}\right)$ im Raum $(\mathbb{R}, |\cdot|)$ konvergent, in $((0,1), |\cdot|)$ jedoch nicht (der Grenzwert liegt nicht im Raum).

In ähnlicher Weise ist $(0,1)$ im Raum $((0,1), |\cdot|)$ abgeschlossen (der ganze Raum ist stets abgeschlossen), nicht jedoch im Raum $(\mathbb{R}, |\cdot|)$.

1.4 Grenzwerte von Abbildungen

Definition 1.40 (Häufungspunkt, isolierter Punkt). Sei (X, d) ein metrischer Raum und $x \in X$. Der Punkt x heißt

(i) isolierter Punkt, wenn es ein $\epsilon > 0$ gibt, so dass $B_\epsilon(x) = \{x\}$,

(ii) Häufungspunkt, wenn er kein isolierter Punkt ist.

Definition 1.41 (Punktierte Umgebung, punktierte Kugel). Sei (X, d_X) ein metrischer Raum und $x \in X$ ein Häufungspunkt. Eine Menge $\dot{U}$ heißt punktierte Umgebung von x, wenn $\dot{U} \cup \{x\}$ eine Umgebung von x ist. Sei $r > 0$ dann ist die punktierte Kugel definiert als $\dot{B}_r(x) = B_r(x) \setminus \{x\}$.

Bemerkung 1.42. Durch die Voraussetzung, dass x ein Häufungspunkt von X ist, enhält jede ϵ-Kugel um x noch mindestens einen weiteren Punkt. Die punktierten Kugeln um x, und damit die punktierten Umgebungen um x, die jeweils eine punktierte Kugel enthalten müssen, sind also nicht leer.

Definition 1.43 (Grenzwert von Abbildungen). Seien (X, d_X), (Y, d_Y) metrische Räume, $x \in X$ ein Häufungspunkt und $\dot{U}$ eine punktierte Umgebung von x. Sei $f : \dot{U} \to Y$ eine Abbildung. Dann konvergiert f in x gegen $y \in Y$ (und y heißt Grenzwert von f im Punkt x), wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass $f(x') \in B_\epsilon(y)$ für alle $x' \in \dot{B}_\delta(x)$. Wir schreiben $y = \lim_{x' \to x} f(x')$.

Beispiel 1.44. Definiere $f, g : \mathbb{R} \setminus \{0\} \to \mathbb{R}$

$$f(x) := x^2$$
$$g(x) := \frac{1}{x}.$$

Dann ist f in 0 konvergent (Wogegen?), g jedoch nicht (Warum?).

Satz 1.45. *Seien (X, d_X), (Y, d_Y) metrische Räume, $x \in X$ ein Häufungspunkt, $\dot{U}$ eine punktierte Umgebung von x und $f : \dot{U} \to Y$ eine Abbildung. Die Abbildung f konvergiert in x gegen $y \in Y$, genau dann wenn für jede Folge (x_n) in $\dot{U}$ mit $x_n \to x$ gilt, dass $(f(x_n))$ gegen y konvergiert.*

Beweis. Es gelte zunächst $\lim_{x' \to x} f(x') = y$ und (x_n) sei eine gegen x konvergente Folge. Sei $\epsilon > 0$ dann existiert nach Voraussetzung $\delta > 0$, so dass $f(x') \in B_\epsilon(y)$ für alle $x' \in \dot{B}_\delta(x)$. Da $x_n \to x$ existiert $N \in \mathbb{N}$, so dass $x_m \in \dot{B}_\delta(x)$ für alle $m > N$. Dann gilt aber $f(x_m) \in B_\epsilon(y)$ für alle $m > N$ und wir haben gezeigt, dass $(f(x_n))$ gegen y konvergiert.

Es gelte nun dass $f(x_n) \to y$, wann immer die Folge (x_n) aus $\dot{U}$ gegen x konvergiert. Angenommen f konvergiert in x nicht gegen y, das heißt es existiert ein $\epsilon > 0$ und für jedes $n \in \mathbb{N}$ ein $x_n \in B_{\frac{1}{n}}(x)$ mit $f(x_n) \notin B_\epsilon(y)$. Die Folge x_n konvergiert gegen x da $d_X(x_n, x) < \frac{1}{n} \to 0$. Die Folge $f(x_n)$ konvergiert nicht gegen y, da $d_Y(f(x_n), y) \geq \epsilon$. Da das im Widerspruch zur Voraussetzung steht, muss f in x gegen y konvergieren. $\square$

Korollar 1.46. *Seien (X, d) ein metrischer Raum, $x \in X$ ein Häufungspunkt, $\dot{U}$ eine punktierte Umgebung von x und $f, g : \dot{U} \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) Funktionen. Es gelte $\lim_{x' \to x} f(x') = y_1$ und $\lim_{x' \to x} g(x') = y_2$. Dann konvergiert $f + g$ in x gegen $y_1 + y_2$ und $f \cdot g$ gegen $y_1 y_2$. Ist $y_1 \neq 0$, dann ist $1/f$ auf einer punktierten Umgebung von x definiert und $1/f$ konvergiert gegen $1/y_1$.*

Beweis. Sei (x_n) eine beliebige Folge in X mit $x_n \to x$. Dann gilt nach Voraussetzung $f(x_n) \to y_1$ und $g(x_n) \to y_2$. Daraus folgt mittels der Konvergenzregeln für Zahlenfolgen $f(x_n) + g(x_n) \to y_1 + y_2$ und $f(x_n)g(x_n) \to y_1 y_2$. Damit gilt nach Satz 1.45, dass die Funktionen $\lim_{x' \to x} f + g = y_1 + y_2$, sowie $\lim_{x' \to x} f \cdot g = y_1 y_2$.

Sei nun zusätzlich $y_1 \neq 0$, das heißt $|y_1| > 0$. Da f in x gegen y_1 konvergiert, existiert $\delta > 0$, so dass $f(x') \in B_{\frac{|y_1|}{2}}(y_1)$ für alle $x' \in \dot{B}_\delta(x)$. Für solche x' gilt also insbesondere (umgekehrte Dreiecksungleichung) $|f(x')| > |y_1| - |f(x') - y_1| > \frac{y_1}{2}$. Damit ist die Funktion $1/f$ auf der punktierten Umgebung $\dot{B}_\delta(x)$ von x definiert. Es gilt jetzt wieder für jede gegen x konvergente Folge (x_n), dass $f(x_n) \to y_1$ und damit $\frac{1}{f(x_n)} \to \frac{1}{y_1}$. Wir verwenden nochmals Satz 1.45 und erhalten $\lim_{x' \to x} \frac{1}{f(x')} = \frac{1}{y_1}$. $\square$

1.5 Stetigkeit

Der Begriff der Stetigkeit formalisiert die Eigenschaft einer Funktion f, dass sich die Funktionswerte $f(x)$ nicht zu stark ändern, wenn sich das Argument x nicht zu stark ändert. Je nach dem was genau man unter „zu stark ändern“ versteht ergeben sich unterschiedliche Stetigkeitsbegriffe die in unterschiedlichen Situationen nützlich sein können und von denen wir einige im Folgenden untersuchen wollen.

Viele funktionale Zusammenhänge die in der Natur vorkommen (Position eines Objektes in Abhängigkeit von der Zeit, Lufttemperatur in Abhängigkeit vom Ort, ...) sind stetig.

Definition 1.47 (Stetigkeit). Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Abbildung $f : X \to Y$ heißt stetig im Punkt $x_0 \in X$, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass $f(x) \in B_\epsilon(f(x_0))$ für alle $x \in B_\delta(x_0)$. Wir sagen f ist stetig (ohne angabe eines Punktes), wenn f in jedem Punkt von X stetig ist.

Beispiel 1.48. (i) Sei (X, d) ein metrischer Raum. Die Identitätsabbildung $\mathrm{id} : X \to X$, $x \mapsto x$ ist stetig in jedem Punkt.

(ii) Die Funktion $\chi_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$ definiert durch

$$\chi_{\mathbb{Q}}(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & \text{sonst} \end{cases}$$

ist in keinem Punkt stetig.

Bemerkung 1.49. (i) Leicht umformuliert bedeutet die Stetigkeit von f in x, dass für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass $f(B_\delta(x)) \subset B_\epsilon(f(x))$.

(ii) Stetigkeit (im Punkt x_0) ist eine sogenannte lokale Eigenschaft, das heißt ob f in x_0 stetig ist oder nicht kann man entscheiden, wenn man f nur in einer beliebig kleinen Umgebung von x_0 kennt. Formal lautet die Aussage: Seien f, g Funktionen von (X, d_X) nach (Y, d_Y), U eine Umgebung von x_0 und $f|_U = g|_U$ Dann ist f in x_0 stetig, genau dann wenn g in x_0 stetig ist.

Um das zu zeigen, sei f stetig und $\epsilon > 0$. Dann existiert $\delta > 0$, so dass $f(B_\delta(x_0)) \subset B_\epsilon(f(x_0))$. Da U Umgebung von x ist, existiert $0 < \rho \leq \delta$, so dass $B_\rho(x_0) \subset U$ ist. Dann gilt

$$g(B_\rho(x_0)) = f(B_\rho(x_0)) \subset f(B_\delta(x_0)) \subset B_\epsilon(f(x_0)) = B_\epsilon(g(x_0)).$$

Damit ist g stetig in x_0.

Durch Vertauschen von f und g erhält man „genau dann wenn“.

Satz 1.50. *Seien (X, d_X), (Y, d_Y) und (Z, d_Z) metrische Räume, $x_0 \in X$, $g : X \to Y$ und $f : Y \to Z$ Abbildungen. Wenn g stetig in x_0 ist und f stetig in $g(x_0)$, dann ist $f \circ g$ stetig in x_0.*

Beweis. Sei $\epsilon > 0$. Nach Voraussetzung existiert $\delta > 0$, so dass $f(y) \in B_\epsilon(f(g(x_0)))$ für alle $y \in B_\delta(g(x_0))$. Außerdem existiert $\rho > 0$, so dass $g(x) \in B_\delta(g(x_0))$ für alle $x \in B_\rho(x_0)$. Dann gilt für jedes $x \in B_\rho(x_0)$, dass $g(x) \in B_\delta(g(x_0))$ und damit auch $f \circ g(x) = f(g(x)) \in B_\epsilon(f \circ g(x_0))$. □

Definition 1.51 (Lipschitz-Stetigkeit, Kontraktion, Isometrie). Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Abbildung $f : X \to Y$ heißt Lipschitz-stetig, wenn es ein $L > 0$ gibt, so dass

$$d_Y(f(x), f(y)) \leq L \cdot d_X(x, y)$$

für alle $x, y \in X$ gilt. Die Konstante L heißt dann Lipschitz-Konstante für die Funktion f. Die Abbildung f heißt Kontraktion falls $d_Y(f(x), f(y)) \leq d_X(x, y)$ und Isometrie, falls $d_Y(f(x), f(y)) = d_X(x, y)$.

Beispiel 1.52. Betrachte die Funktionen $f, g : [0, 1] \to \mathbb{R}$

$$\begin{aligned} f(x) &:= x^2 \\ g(x) &:= \sqrt{x}. \end{aligned}$$

Die Funktion f ist Lipschitz-stetig da

$$|f(x) - f(y)| = \left|x^2 - y^2\right| = |x - y|\,|x + y| \leq 2\,|x - y|\,.$$

Die Funktion g ist nicht Lipschitz-stetig, da

$$\frac{|g(x) - g(0)|}{|x - 0|} = \frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}} \to \infty$$

für x gegen 0. Wie wir noch sehen werden, ist g stetig.

Satz 1.53. *Jede Lipschitz-stetige Funktion ist stetig.*

Beweis. Seien $(X, d_X), (Y, d_Y)$ metrische Räume, $x \in X$ und $f : X \to Y$ eine Lipschitz-stetige Funktion mit Lipschitz-Konstante L. Sei $\epsilon > 0$. Dann gilt für jedes $x' \in B_{\frac{\epsilon}{L}}(x)$

$$d_Y(f(x'), f(x)) \leq L \cdot d_X(x', x) < \epsilon.$$

Damit ist die Stetigkeit gezeigt. □

Beispiel 1.54. (i) Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Die Abbildung $p_i : \mathbb{K}^n \to \mathbb{K}$

$$x = (x_1, \ldots, x_n) \mapsto x_i$$

ist eine Kontraktion und damit stetig.

(ii) Sei (X, d) ein metrischer Raum und $(X \times X, d_{X\times X})$ das Produkt von X mit sich selbst und der Produktmetrik (Beispiel 1.4 (iii)). Wir zeigen, dass die Metrik, also die Abbildung

$$\begin{aligned} d : X \times X &\to \mathbb{R} \\ (x, y) &\mapsto d(x, y) \end{aligned}$$

eine Kontraktion von $(X \times X, d_{X\times X})$ nach $\mathbb{R}$ ist (insbesondere also stetig).

Für $x_1, x_2, y_1, y_2 \in X$ gilt nach der Dreiecksungleichung

$$d(x_1, y_1) \leq d(x_1, x_2) + d(x_2, y_1) \leq d(x_1, x_2) + d(y_2, y_1) + d(x_2, y_2).$$

Damit erhalten wir

$$|d(x_1, y_1) - d(x_2, y_2)| \leq d(x_1, x_2) + d(y_1, y_2) = d_{X\times X}((x_1, x_2), (y_1, y_2)).$$

(iii) Analog gilt mit der umgekehrten Dreiecksungleichung

$$|\|x\| - \|y\|| \leq \|x - y\|$$

und $\|\cdot\|$ ist eine Kontraktion von $(V, \|\cdot\|)$ nach $\mathbb{R}$ also insbesondere stetig.

Satz 1.55. *Seien (X, d_X) und (Y, d_Y) metrische Räume und $f : X \to Y$. Die Abbildung f ist stetig in $x \in X$ genau dann, wenn entweder x isolierter Punkt von X ist, oder $\lim_{x'\to x} f(x') = f(x)$ gilt.*

Beweis. Sei zunächst f stetig. Wenn x isolierter Punkt ist, ist nichts zu zeigen. Sei also x Häufungspunkt von X. Sei $\epsilon > 0$, dann existiert wegen der Stetigkeit $\delta > 0$, so dass $f(x') \in B_\epsilon(f(x))$ für alle $x' \in B_\delta(x)$ also insbesondere für alle $x' \in \dot{B}_\delta(x)$. Damit gilt $\lim_{x'\to x} f(x') = f(x)$.

Es sei nun x isolierter Punkt von X, das heißt es existiert $\delta > 0$, so dass $B_\delta(x) = \{x\}$. Dann ist $f(B_\delta(x)) = \{f(x)\} \subset B_\epsilon(f(x))$ für jedes $\epsilon > 0$ und damit f stetig in x.

Sei schließlich x Häufungspunkt von X und es gelte $\lim_{x'\to x} f(x') = f(x)$. Die Grenzwertbedingung bedeutet gerade, dass für jedes $\epsilon > 0$ ein $\delta > 0$ existiert, so dass $f(\dot{B}_\delta(x)) \subset B_\epsilon(f(x))$. Dann gilt aber auch $f(B_\delta(x)) \subset B_\epsilon(f(x))$ also ist f stetig in x. □

Mit dem obigen Satz können wir nun Aussagen über Grenzwerte von Funktionen in entsprechende Aussagen über stetige Funktionen übersetzen. So zum Beispiel die folgenden Korollare deren Beweise dem Leser zur eigenständigen Übung überlassen werden. Dazu macht man sich zunächst bewusst, dass eine Folge (x_n) genau dann gegen einen isolierten Punkt x konvergiert, wenn sie ab einem bestimmten Index konstant gleich x ist.

Korollar 1.56. *Seien (X, d_X) und (Y, d_Y) metrische Räume und $f : X \to Y$ eine Abbildung. Dann ist f stetig in $x \in X$, genau dann wenn für jede Folge (x_n) in X, die gegen x konvergiert auch $(f(x_n))$ gegen $f(x)$ konvergiert.*

Korollar 1.57. *Seien $f_1, \dots, f_n : \mathbb{R}^m \to \mathbb{R}$ Funktionen. Dann ist die Abbildung $f : \mathbb{R}^m \to \mathbb{R}^n$, $x \mapsto (f_1(x), \dots, f_n(x))$ stetig im Punkt x_0, genau dann wenn f_i für jedes $i \in \{1, \dots, n\}$ stetig im Punkt x_0 ist.*

Korollar 1.58. *Seien (X, d) ein metrischer Raum und $f, g : X \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) Funktionen die stetig in $x \in X$ sind, dann sind die Funktionen $f + g$ und $f \cdot g$ stetig in x. Ist $f(x) \neq 0$, dann ist $1/f$ auf einer Umgebung von x definiert und in x stetig.*

Beispiel 1.59. Für $\eta = (\eta_1, \dots, \eta_n) \in \mathbb{N}_0^n$ und $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) definieren wir

$$x^\eta := x_1^{\eta_1} x_2^{\eta_2} \cdots x_n^{\eta_n}.$$

Dabei bezeichnen wir η als Multiindex. Wir setzen außerdem $|\eta| := \eta_1 + \cdots + \eta_n$. Seien $c_\eta \in \mathbb{K}$ Konstanten für jedes η mit $|\eta| \leq N$, dann heißt die Funktion

$$\begin{aligned} f : \mathbb{K}^n &\to \mathbb{K} \\ x &\mapsto \sum_{\eta, |\eta| \leq N} c_\eta x^\eta \end{aligned}$$

Polynom in n Variablen (über dem Körper $\mathbb{K}$). Solche Polynome sind stetig.

Bemerkung 1.60. Im Kontext von Polynomen (und später Potenzreihen) interpretieren wir den Ausdruck 0^0, der auftreten kann wenn $\eta_i = x_i = 0$ ist, als 1.

Der folgende Satz ist eine Charakterisierung stetiger Funktionen, die insbesondere in Beweisen oft hilfreich ist.

Satz 1.61. *Seien (X, d_X) und (Y, d_Y) metrische Räume und $f : X \to Y$ eine Abbildung.*

(i) Die Abbildung f ist stetig in $x \in X$ genau dann wenn $f^{-1}(U)$ eine Umgebung von x ist für jede Umgebung U von $f(x)$.

(ii) Die Abbildung f ist stetig (in jedem Punkt), genau dann wenn $f^{-1}(O)$ offen ist für jede offene Menge $O \subset Y$.

(iii) Die Abbildung f ist stetig (in jedem Punkt), genau dann wenn $f^{-1}(C)$ abgeschlossen ist für jede abgeschlossene Menge $C \subset Y$.

Beweis. Übung □

Beispiel 1.62. Betrachte die Funktion

$$\begin{aligned} f : \mathbb{R}^n &\to \mathbb{R} \\ x &\mapsto \|x\| . \end{aligned}$$

Wir wissen bereits, dass diese Funktion stetig ist. Sei $r > 0$. Dann folgt mit dem vorhergehenden Satz, dass die Sphäre mit Radius r

$$f^{-1}(\{r\}) = \{x \in \mathbb{R}^n |\ \|x\| = r\}$$

und die abgeschlossene Kugel mit Radius r

$$f^{-1}([0,r]) = \{x \in \mathbb{R}^n |\ \|x\| \leq r\} = K_r(0)$$

abgeschlossen sind, und die offene Kugel

$$f^{-1}((-1,r)) = \{x \in \mathbb{R}^n |\ \|x\| < r\} = B_r(0)$$

offen ist.

Theorem 1.63 (Zwischenwertsatz). *Seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a,b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) \leq f(b)$. Dann existiert zu jedem $y \in [f(a), f(b)]$ ein $x \in [a,b]$, so dass $f(x) = y$.*

Beweis. Sei $y \in [f(a), f(b)]$. Setze

$$x := \sup \left\{x' \in [a,b] \middle|\ f(x') \leq y\right\}$$

und wähle eine Folge (x_n) in $\{x' \in [a,b] |\ f(x') \leq y\}$ die gegen x konvergiert. Da f stetig ist konvergiert $(f(x_n))$ gegen $f(x)$. Wir zeigen $f(x) = y$. Zunächst gilt nach der Definition von x, dass $f(x_n) \leq y$ für alle $n \in \mathbb{N}$ und damit gilt für den Grenzwert ebenfalls $f(x) \leq y$. Falls $x = b$, so gilt $f(x) = f(b) \leq y$ und $y \in [f(a), f(b)]$ also $f(x) = f(b) = y$. Falls $x < b$, so konvergiert die Folge $x + \frac{1}{n}$ gegen x. Nach der Definition von x ist $f(x + \frac{1}{n}) > y$ (für hinreichend große n). Wir verwenden nochmals die Stetigkeit von f und erhalten $f(x) = \lim_{n\to\infty} f(x + \frac{1}{n}) \geq y$. Damit ist $f(x) = y$. □

Satz 1.64. *Sei $A \subset \mathbb{R}^n$ eine beschränkte und abgeschlossene Teilmenge und $f : A \to \mathbb{R}^m$ eine stetige Abbildung. Dann ist das Bild von f, $f(A)$, beschränkt und abgeschlossen.*

Beweis. Wir wollen das Folgenkriterium für Abgeschlossenheit (Satz 1.33) verwenden. Sei dazu (y_n) eine Folge in $f(A)$, die in $\mathbb{R}^m$ gegen y konvergiert. Wir müssen zeigen, dass der Grenzwert y ebenfalls in $f(A)$ liegt. Da $(y_n) \subset f(A)$ existieren x_n, so dass $f(x_n) = y_n$ für jedes $n \in \mathbb{N}$. Nach dem Satz von Bolzano-Weierstraß existiert eine konvergente Teilfolge (x_{n_k}) von (x_n). Da die Teilfolge (x_{n_k}) in A liegt und A abgeschlossen ist, liegt auch ihr Grenzwert x in A. Wegen der Stetigkeit von f konvergiert die Folge $(f(x_{n_k}))$ gegen $f(x)$. Andererseits ist $(f(x_{n_k}))$ Teilfolge der konvergenten Folge $f(x_n)$ und konvergiert daher gegen deren Grenzwert y. Da Grenzwerte (im metrischen Raum) stets eindeutig bestimmt sind, muss $f(x) = y$ gelten und wir haben gezeigt, dass $y \in f(A)$.

Es bleibt noch die Beschränktheit von $f(A)$ zu zeigen. Angenommen $f(A)$ wäre unbeschränkt, das heißt es ist in keiner noch so großen Kugel um den Koordinatenursprung enthalten. Dann existiert für jedes $n \in \mathbb{N}$ ein $x_n \in A$, so dass $\|f(x_n)\| > n$.

Wir verwenden nochmals Bolzano-Weierstraß um aus (x_n) eine konvergente Teilfolge (x_{n_k}) auszuwählen. Wie oben liegt deren Grenzwert in A und wegen der Stetigkeit von f gilt $\lim_{k\to\infty} f(x_{n_k}) = f(x)$. Für jedes $N \in \mathbb{N}$ und jedes $n \in \mathbb{N}$ mit $n > N$ gilt dann $\|f(x_n)\| > n > N$. Damit gilt wegen der Stetigkeit der Norm

$$\|f(x)\| = \lim_{k\to\infty} \|f(x_{n_k})\| > N.$$

Da $f(x) \in \mathbb{R}^m$ ist, kann diese Ungleichung unmöglich für alle $N \in \mathbb{N}$ gelten, $f(A)$ muss also beschränkt sein. □

Bemerkung 1.65. (i) Beschränkte und abgeschlossene Mengen im $\mathbb{R}^n$ heißen kompakt. Wir werden uns später eingehender mit Kompaktheit in allgemeineren metrischen Räumen beschäftigen und um bis dahin keine Verwirrung zu stiften, werden wir den Begriff zunächst noch meiden.

(ii) Stetige Abbildungen bilden im Allgemeinen nicht beschränkte Mengen auf beschränkte Mengen ab und auch nicht abgeschlossene auf abgeschlossene.

Korollar 1.66. *Sei $A \subset \mathbb{R}^n$ eine beschränkte und abgeschlossene Teilmenge und $f : A \to \mathbb{R}$ stetig, dann nimmt f auf A sein Maximum an, das heißt es existiert $x \in A$, so dass $f(y) \leq f(x)$ für alle $y \in A$.*

Beweis. Wir zeigen, dass eine beschränkte, abgeschlossene Teilmenge von $K \subset \mathbb{R}$ ein Maximum hat. Sei dazu $M = \sup K$ und wähle eine Folge $(x_n) \subset K$, so dass $x_n \to M$. Da K abgeschlossen ist liegt der Grenzwert M dieser Folge in K.

Wenden wir dass auf die Menge $f(A)$ an, die nach dem vorhergehenden Satz abgeschlossen und beschränkt ist, so folgt die Aussage des Korollars. □

Satz 1.67. *Seien $A \subset \mathbb{R}^n$ und $B \subset \mathbb{R}^m$ beschränkte, abgeschlossene Teilmengen und $f : A \to B$ eine stetige, bijektive Funktion. Dann ist die Umkehrfunktion $f^{-1} : B \to A$ ebenfalls stetig.*

Beweis. Da f bijektiv ist, besitzt es eine Umkehrabbildung. Setze nun $g = f^{-1}$. Die Abbildung g ist bijektiv da f bijektiv ist. Sei $C \subset A$ abgeschlossen (da A selbst abgeschlossen ist, ist C abgeschlossen in $\mathbb{R}^n$, genau dann wenn es in A abgeschlossen ist). Aus der Beschränktheit von A folgt, dass C ebenfalls beschränkt ist. Wegen der Bijektivität von g gilt

$$\begin{aligned} f(C) &= \{f(x) \in B |\ x \in C\} \\ &= \{f(g(y)) \in B |\ g(y) \in C\} \\ &= \{y \in B |\ g(y) \in C\} = g^{-1}(C). \end{aligned}$$

Nach Satz 1.64 ist $f(C) = g^{-1}(C)$ abgeschlossen (und beschränkt). Da C eine beliebige abgeschlossene Teilmenge von A war, folgt damit aus Satz 1.61 die Stetigkeit von $g = f^{-1}$. □

Beispiel 1.68. Sei $n \in \mathbb{N}$. Dann ist die Funktion

$$\begin{aligned} f : [0, \infty) &\to [0, \infty) \\ x &\mapsto \sqrt[n]{x} \end{aligned}$$

stetig.

Beweis. Da $[0, \infty)$ nicht beschränkt ist, können wir Satz 1.67 nicht unmittelbar anwenden. Wir nutzen hier aus, dass Stetigkeit eine lokale Eigenschaft ist (Bemerkung 1.49). Sei also $x_0 \in [0, \infty)$.

Definiere

$$\begin{aligned} \tilde{f} : [0, x_0 + 1] &\to \left[0, \sqrt[n]{x_0 + 1}\right] \\ x &\mapsto \sqrt[n]{x}. \end{aligned}$$

Dann ist $\tilde{f}$ die Umkehrfunktion von

$$\begin{aligned} g : \left[0, \sqrt[n]{x_0 + 1}\right] &\to [0, x_0 + 1] \\ y &\mapsto y^n. \end{aligned}$$

Die Funktion g ist stetig und streng monoton (also insbesondere bijektiv). Damit folgt nach Satz 1.67 die Stetigkeit von $\tilde{f}$, insbesondere im Punkt x_0. Da f und $\tilde{f}$ in einer Umgebung von x_0 übereinstimmen, ist also auch f in x_0 stetig. Da $x_0 \in [0, \infty)$ beliebig war, ist f überall stetig. $\square$

1.6 Konvergenz von Funktionenfolgen

Definition 1.69 (Punktweise Konvergenz). Sei M eine Menge, $f_n : M \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) Funktionen für jedes $n \in \mathbb{N}$ und $f : M \to \mathbb{K}$ eine weitere Funktion. Die Folge (f_n) heißt punktweise konvergent gegen f, wenn gilt

$$\lim_{n\to\infty} f_n(x) = f(x) \text{ für alle } x \in M.$$

Punktweise Konvergenz von Funktionenfolgen ist oft vergleichsweise einfach zu zeigen, aber oft auch nicht sehr nützlich, da viele wichtige Eigenschaften von Funktionen bei punktweiser Konvergenz nicht erhalten bleiben. Die punktweise Konvergenz läßt sich im Allgemeinen – konkreter falls M überabzählbar ist – auch nicht durch eine Metrik charakterisieren.

Beispiel 1.70. Seien $f_n : [0,1] \to \mathbb{R}$ die folgenden Funktionen

$$f_n(x) := \begin{cases} 1 - nx & x \in [0, 1/n] \\ 0 & \text{sonst.} \end{cases}$$

Dann ist f_n stetig für jedes $n \in \mathbb{N}$ und die Folge (f_n) konvergiert punktweise gegen $f : [0,1] \to \mathbb{R}$

$$f(x) := \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases}$$

welches nicht mehr stetig ist.

Wir können nun den Formalismus des normierten Raumes nutzen, um einen anderen Konvergenzbegriff für Funktionenfolgen zu definieren, der viele gute Eigenschaften von Funktionen erhält.

Beispiel 1.71. Sei M eine Menge und $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Betrachte den Raum der beschränkten Funktionen über M, das heißt

$$\mathbf{B}(M) = \{f : M \to \mathbb{K} |\ \exists K \in \mathbb{R} \text{ so dass } |f(x)| \leq K \text{ für alle } x \in M\}. \tag{1.3}$$

Mit der punktweisen Addition und skalaren Multiplikation, das heißt für $f, g \in \mathbf{B}(M)$, $\lambda \in \mathbb{K}$ und für alle $x \in M$

$$\begin{aligned} (f+g)(x) &:= f(x) + g(x) \\ (\lambda f)(x) &:= \lambda f(x) \end{aligned}$$

erhält $\mathbf{B}(M)$ die Struktur eines Vektorraumes.

Dann ist

$$\begin{aligned} \|\cdot\|_\infty : \mathbf{B}(M) &\to \mathbb{R} \\ f &\mapsto \sup_{x\in M} |f(x)| \end{aligned}$$

eine Norm auf $\mathbf{B}(M)$.

Beweis. Wir ziegen, dass $\|\cdot\|_\infty$ eine Norm ist. Die Positivität und absolute Homogenität sind einfach nachzuprüfen. Um die Definitheit zu zeigen sei $f \in \mathbf{B}(M)$ so, dass $\|f\|_\infty = 0$ gilt. Dann ist $|f(x)| = 0$, also $f(x) = 0$ für alle $x \in M$ und damit ist f die Nullfunktion.

Es bleibt noch die Dreiecksungleichung zu zeigen. Zunächst stellen wir fest, dass nach der Definition der Norm $|f(x)| \leq \|f\|_\infty$ für jedes $x \in M$ gilt. Dann folgt für $f, g \in \mathbf{B}(M)$ mit der Dreiecksungleichung für den Betrag

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty .$$

Da das für alle $x \in M$ gilt, ist dann auch

$$\|f + g\|_\infty = \sup_{x \in M} |f(x) + g(x)| \leq \|f\|_\infty + \|g\|_\infty . \qquad \square$$

Definition 1.72 (Gleichmäßige Konvergenz). Eine Folge von beschränkten Funktionen (f_n), $f_n : M \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$), die bezüglich der Norm $\|\cdot\|_\infty$ gegen $f : M \to \mathbb{K}$ konvergiert heißt gleichmäßig konvergent. Wenn von dem Funktionenraum $\mathbf{B}(M)$ (und später $\mathbf{C}(M)$) die Rede ist werden wir die Norm $\|\cdot\|_\infty$ oft nicht explizit erwähnen.

Bemerkung 1.73. In formaler Notation bedeutet die punktweise Konvergenz einer Funktionenfolge f_n gegen f, dass

$$\forall \epsilon > 0 \quad \forall x \in M \quad \exists N \in \mathbb{N} \quad \forall n > N \quad |f_n(x) - f(x)| < \epsilon$$

erfüllt ist, bei gleichmäßiger Konvergenz hingegen gilt

$$\forall \epsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall x \in M \quad \forall n > N \quad |f_n(x) - f(x)| < \epsilon.$$

Die Definitionen unterscheiden sich also nur darin, ob N von x abhängen darf oder nicht. Insbesondere ist jede gleichmäßig konvergente Folge auch punktweise konvergent.

Satz 1.74. *Der Funktionenraum $\mathbf{B}(M)$ ist vollständig, also ein Banachraum.*

Beweis. Sei $(f_n) \subset \mathbf{B}(M)$ eine Cauchy-Folge. Wähle $x \in M$ fest. Dann ist die Folge $(f_n(x))$ ebenfalls eine Cauchy-Folge, da

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty$$

gilt. Da $\mathbb{K}$ (also $\mathbb{R}$ oder $\mathbb{C}$) vollständig sind, konvergiert also die Folge $(f_n(x))$. Wir setzen jetzt

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

Damit haben wir einen Grenzwert f für die Folge (f_n) gefunden, allerdings nur bezüglich der punktweisen Konvergenz.

Da jede Cauchy-Folge beschränkt ist (Bemerkung 1.19) existiert $K \in \mathbb{R}$, so dass $\|f_n\|_\infty \leq K$ für alle $n \in \mathbb{N}$. Damit folgt für jede $x \in M$

$$|f(x)| = \lim_{n\to\infty} |f_n(x)| \leq \limsup_{n\to\infty} \|f_n\|_\infty \leq K,$$

unsere Grenzfunktion f ist also beschränkt.

Sei nun $\epsilon > 0$. Wir wählen mit der Cauchy-Eigenschaft $N \in \mathbb{N}$ so, dass $\|f_n - f_m\|_\infty < \epsilon$ für alle $n, m \geq N$. Die Wahl von N ist nicht von x abhängig. Dann gilt für jedes $x \in M$ und jedes $n, m > N$

$$|f_m(x) - f_n(x)| \leq \|f_m - f_n\|_\infty \leq \epsilon.$$

Da die rechte Seite nicht mehr von m abhängt, können wir den Grenzwert $m \to \infty$ bilden und erhalten, da die Betragsfunktion stetig ist

$$|f(x) - f_n(x)| \leq \epsilon.$$

Da die rechte Seite auch von x unabhängig ist können wir das Supremum bilden und erhalten

$$\sup_{x\in M} |f(x) - f_n(x)| = \|f - f_n\|_\infty \leq \epsilon.$$

Damit haben wir die Konvergenz von (f_n) gegen f bezüglich der Norm gezeigt. □

Definition 1.75. Sei (X, d) ein metrischer Raum und $\mathbf{B}(X)$ der Raum der beschränkten Funktionen auf X. Der Teilraum der stetigen, beschränkten Funktionen wird mit $\mathbf{C}_b(X)$ bezeichnet.

Bemerkung 1.76. Falls X eine beschränkte, abgeschlossene Teilmenge von $\mathbb{R}^n$ ist, dann ist nach Satz 1.64 jede stetige Funktion beschränkt. In diesem Fall enthält der Raum $\mathbf{C}_b(X)$ alle stetigen Funktionen auf X und wir schreiben $\mathbf{C}(X)$.

Im Gegensatz zur punktweisen Konvergenz erhält die gleichmäßige Konvergenz viele gute Eigenschaften von Funktionen.

Satz 1.77. *Sei (X, d) ein metrischer Raum. Der Unterraum $\mathbf{C}_b(X)$ ist abgeschlossen in $\mathbf{B}(X)$. Insbesondere ist $\mathbf{C}_b(X)$ vollständig.*

Beweis. Wir wollen das Folgenkriterium für Abgeschlossenheit Satz 1.33 verwenden. Sei also $(f_n) \subset \mathbf{C}_b(X)$ eine Folge stetiger Funktionen, die (gleichmäßig) gegen $f \in \mathbf{B}(X)$ konvergiert. Sei $x \in X$ und $\epsilon > 0$. Da (f_n) gleichmäßig gegen f konvergiert, existiert $N \in \mathbb{N}$ so, dass $\|f - f_n\|_\infty < \frac{\epsilon}{3}$ für jedes $n > N$. Wähle jetzt $n > N$ fest. Da f_n stetig ist, existiert $\delta > 0$ so, dass $|f_n(x) - f_n(y)| < \frac{\epsilon}{3}$ falls $d(x, y) < \delta$. Dann gilt für jedes $y \in B_\delta(x)$

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &\leq 2\,\|f - f_n\|_\infty + |f_n(x) - f_n(y)| < \epsilon. \end{aligned}$$

Damit ist f stetig in x und da $x \in X$ beliebig war, ist f stetig.

Da die Folge (f_n) beliebig war, ist damit auch die Abgeschlossenheit von $\mathbf{C}_b(X)$ gezeigt. Die Vollständigkeit dieses Raumes folgt dann aus Satz 1.38. □

Definition 1.78. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$, $x_0 \in \mathbb{K}$ und $(a_n)_{n\in\mathbb{N}_0} \subset \mathbb{K}$. Dann heißt $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine (formale) Potenzreihe um den Punkt x_0 über dem Körper $\mathbb{K}$.

Die Zahl

$$\sup\left\{ |x-x_0| \,\middle|\, \sum_{n=0}^{\infty} a_n(x-x_0)^n \text{ konvergiert} \right\}$$

in $[0,\infty) \cup \{\infty\}$ heißt Konvergenzradius der Potenzreihe.

Satz 1.79. *Sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe über dem Körper $\mathbb{K}$ und $r \in [0,\infty) \cup \{\infty\}$ ihr Konvergenzradius. Für $|x-x_0| < r$ konvergiert die Reihe absolut, für $|x-x_0| > r$ divergiert sie.*

Es gilt die Formel

$$\frac{1}{r} = \limsup_{n\to\infty} \sqrt[n]{|a_n|}. \tag{1.4}$$

Wobei wir $\frac{1}{\infty} = 0$ und $\frac{1}{0} = \infty$ setzen.

Beweis. Um die Notation einfach zu halten, führen wir den Beweis nur für $x_0 = 0$. Dass die Reihe für $|x| > r$ divergiert folgt aus der Definition des Konvergenzradius. Wir zeigen zunächst die absolute Konvergenz der Reihe falls $|x| < r$. Falls $r = 0$ ist, so ist nichts zu zeigen. Sei also $r > 0$ und $|x| < r$. Nach der Definition des Konvergenzradius existert y, so dass $|x| < |y| \leq r$ und $\sum_{n=0}^{\infty} a_n y^n$ konvergiert. Insbesondere muss also $(a_n y^n)$ eine Nullfolge sein und so gilt $|a_n|\,|y|^n \leq C$ für ein $C \in [0,\infty)$ und alle $n \in \mathbb{N}_0$. Damit gilt

$$|a_n|\,|y|^n \left(\frac{|x|}{|y|}\right)^n \leq C \left(\frac{|x|}{|y|}\right)^n$$

die Reihe

$$\sum_{n=0}^{\infty} |a_n|\,|y|^n \left(\frac{|x|}{|y|}\right)^n$$

wird also durch eine konvergente geometrische Reihe majorisiert und ist somit selbst konvergent. Damit ist die Reihe

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} a_n y^n \left(\frac{x}{y}\right)^n \tag{1.5}$$

absolut konvergent.

Die Formel für den Konvergenzradius ergibt sich aus dem Wurzelkriterium. Setze

$$b := \limsup_{n\to\infty} \sqrt[n]{|a_n|}.$$

Sei zunächst $b \in (0,\infty)$. Dann gilt

$$\limsup_{n\to\infty} \sqrt[n]{|a_n x^n|} = |x| \limsup_{n\to\infty} \sqrt[n]{|a_n|} = |x|\, b.$$

Damit ist die Reihe konvergent falls $|x| < \frac{1}{b}$ – also gilt $r \geq \frac{1}{b}$ – und divergent falls $|x| > \frac{1}{b}$ also ist $r \leq \frac{1}{b}$. Schließlich sieht man leicht, dass die Potenzreihe lediglich in $x = 0$ konvergiert falls $b = \infty$ und für jedes $x \in \mathbb{K}$ konvergiert falls $b = 0$ ist. $\square$

Bemerkung 1.80. (i) Aus dem obigen Beweis ist ersichtlich, dass

$$r = \sup \left\{ a \in [0, \infty) \middle| \sum_{n=0}^{\infty} |a_n| \, a^n \text{ konvergiert} \right\}.$$

(ii) Falls der folgende Grenzwert existiert (also insbesondere $a_n \neq 0$ ist ab einem bestimmten Index n), so gilt auch

$$r = \lim_{n \to \infty} \frac{|a_n|}{|a_{n+1}|}.$$

Beispiel 1.81. Die Potenzreihe

$$\sum_{n=0}^{\infty} x^n$$

ist konvergent auf $(-1, 1)$ (beziehungsweise auf $B_1(0) \subset \mathbb{C}$) und konvergiert dort gegen die Funktion $f(x) = \frac{1}{1-x}$.

Man beachte dabei, dass der Grenzwert der Reihe nur auf dem Konvergenzkreis mit f übereinstimmt (außerhalb dieses Kreises existiert der Grenzwert nicht) obwohl der maximale Definitionsbereich von f darüber hinaus geht.

Satz 1.82. *Sei $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ eine Potenzreihe über dem Körper $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$) mit Konvergenzradius $r > 0$. Sei $0 < a < r$. Dann konvergiert die Potenzreihe auf $K_a(x_0)$ gleichmäßig.*

Insbesondere ist die Funktion

$$\begin{aligned} f : B_r(x_0) &\to \mathbb{K} \\ x &\mapsto \sum_{n=0}^{\infty} a_n (x - x_0)^n \end{aligned}$$

stetig.

Beweis. Wir betrachten wiederum nur den Fall $x_0 = 0$. Sei $0 < a < r$. Wir wissen bereits, dass die Reihe $\sum_{n=0}^{\infty} a_n x^n$ auf $K_a(0)$ (sogar auf $B_r(0)$) punktweise konvergiert, wir wollen jedoch die Konvergenz bezüglich der Supremumnorm $\|\cdot\|_\infty$ zeigen. Setze

$$\begin{aligned} f_n : K_a(0) &\to \mathbb{K} \\ x &\mapsto x^n \end{aligned}$$

für alle $n \in \mathbb{N}_0$. Dann gilt

$$\|f_n\|_\infty = \sup_{x \in K_a(0)} |x^n| = a^n$$

und wir erhalten

$$\sum_{n=0}^{\infty} \|a_n f_n\|_\infty = \sum_{n=0}^{\infty} |a_n| \, a^n < \infty \tag{1.6}$$

da $a < r$. Die Reihe $\sum_{n=0}^{\infty} a_n f_n$ ist also absolut konvergent in $\mathbf{B}(K_a(0))$. Da dieser Raum vollständig ist, folgt daraus die Konvergenz der Reihe in $\mathbf{B}(K_a(0))$ (Satz 1.21) also die gleichmäßige Konvergenz.

Da die gleichmäßige Konvergenz die punktweise Konvergenz impliziert (Bemerkung 1.73), muss der gleichmäßige Grenzwert mit dem punktweisen übereinstimmen also mit der Funktion $x \mapsto \sum_{n=0}^{\infty} a_n x^n$. Da die Partialsummen $\sum_{n=0}^{N} a_n x^n$ für jedes N Polynome und somit stetig sind, ist nach Satz 1.77 auch der Grenzwert stetig (auf $K_a(0)$).

Wir wollen nun noch die Stetigkeit auf $B_r(0)$ zeigen. Wir setzen

$$\begin{aligned} s : B_r(0) &\to \mathbb{K} \\ x &\mapsto \sum_{n=0}^{\infty} a_n x^n \end{aligned}$$

Sei nun $y \in B_r(0)$ und wähle a so, dass $|y| < a < r$. Dann ist nach dem soeben gezeigten $s|_{K_a(0)}$ stetig, also insbesondere stetig in $y \in K_a(0)$. Da $K_a(0)$ eine Umgebung von y enthält, stimmen s und $s|_{K_a(0)}$ auf dieser Umgebung überein und da Stetigkeit eine lokale Eigenschaft ist, folgt daraus die Stetigkeit von s in y. Da das für jedes $y \in B_r(0)$ gilt, ist s stetig. □

Beispiel 1.83. (i) Wir definieren die Exponentialfunktion

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Der Konvergenzradius ist

$$r = \lim_{n\to\infty} \frac{\frac{1}{n!}}{\frac{1}{(n+1)!}} = \lim_{n\to\infty} n+1 = \infty,$$

die Reihe also auf ganz $\mathbb{R}$ (oder ganz $\mathbb{C}$) konvergent. Nach obigem Satz ist die Exponentialfunktion stetig.

(ii) Ebenfalls definieren wir für $x \in \mathbb{R}$

$$\begin{aligned} \sin(x) &:= \Im e^{\mathrm{i}x} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \\ \cos(x) &:= \Re e^{\mathrm{i}x} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}. \end{aligned}$$

Beide Reihen konvergieren auf ganz $\mathbb{R}$ und sind daher stetige Funktionen.

Bemerkung 1.84. Die Ergebnisse dieses Abschnitts können – inklusive der Beweise – ohne Weiteres auf $\mathbb{R}^n$-wertige Funktionen oder sogar auf Funktionen mit Werten in einem beliebigen Banachraum verallgemeinert werden.

2 Differenzierbarkeit

2.1 Differenzierbarkeit von Funktionen einer reellen Variablen

Wie bisher sei $\mathbb{K}$ einer der Zahlkörper $\mathbb{R}$ oder $\mathbb{C}$.

Definition 2.1. Sei $U \subset \mathbb{R}$ offen, $f : U \to \mathbb{K}$ eine Funktion und $x \in U$. Die Funktion f heißt differenzierbar in x, wenn der Grenzwert

$$f'(x) := \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

existiert. Dann heißt $f'(x)$ die Ableitung der Funktion f an der Stelle x.

Die Funktion f heißt differenzierbar, falls sie in jedem Punkt des Definitionsbereiches U differenzierbar ist.

Bemerkung 2.2. (i) Genauso können wir Differenzierbarkeit und Ableitung für Kurven im $\mathbb{R}^n$ (oder sogar in einem beliebigen Banachraum V), also für Abbildungen $f : U \to \mathbb{R}^n$ ($f : U \to V$) definieren. Für $\mathbb{R}^n$-wertige Abbildungen entspricht die Differenzierbarkeit genau der Differenzierbarkeit der einzelnen Komponentenfunktionen.

(ii) Ist eine Funktion in x differenzierbar, so gilt

$$\lim_{h \to 0} f(x+h) - f(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} h = 0$$

also ist f in x stetig.

(iii) Differenzierbarkeit ist eine lokale Eigenschaft, hängt also nur vom Verhalten der Funktion in einer beliebig kleinen Umgebung ab.

(iv) Die Funktion f ist differenzierbar in x, genau dann wenn

$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x}$$

existiert.

Satz 2.3. *Seien $U \subset \mathbb{R}$ offen, $x \in U$ und $f, g : U \to \mathbb{K}$ in x differenzierbare Funktionen. Dann sind auch $f + g$ und $f \cdot g$ differenzierbar in x und es gilt*

$$\begin{aligned} (f+g)'(x) &= f'(x) + g'(x), \\ (f \cdot g)'(x) &= f'(x)g(x) + f(x)g'(x). \end{aligned}$$

Ist $f(x) \neq 0$, dann ist $\frac{1}{f}$ auf einer (offenen) Umgebung von x definiert und es gilt

$$\left(\frac{1}{f}\right)'(x) = -\frac{f'(x)}{f(x)^2}.$$

Beweis. Übung $\square$

Satz 2.4. *Seien $U, V \subset \mathbb{R}$ offen, $g : U \to V$ differenzierbar in x und $f : V \to \mathbb{K}$ differenzierbar in $g(x)$, dann ist $f \circ g$ differenzierbar in x und es gilt*

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

Beweis. Erfolgt weiter unten allgemeiner für Funktionen mehrerer Variablen. □

Definition 2.5. Sei (X, d) ein metrischer Raum, $f : X \to \mathbb{R}$ eine Funktion und $x \in X$. Die Funktion f hat in x ein lokales Maximum, wenn es $\epsilon > 0$ gibt, so dass

$$f(y) \leq f(x) \text{ für alle } y \in B_\epsilon(x).$$

Die Funktion f hat in x ein lokales Minimum, wenn $-f$ in x ein lokales Maximum hat. Ein lokales Extremum ist entweder ein lokales Maximum oder ein lokales Minimum.

Satz 2.6. *Sei $U \subset \mathbb{R}$ offen, $x \in U$ und $f : U \to \mathbb{R}$ differenzierbar in x. Falls f in x ein lokales Extremum hat, dann ist $f'(x) = 0$.*

Beweis. Die Funktion f habe oBdA. in x ein lokales Maximum (andernfalls betrachte $-f$). Dann gibt es $\epsilon > 0$, so dass $f(y) \leq f(x)$ für alle $y \in B_\epsilon(x)$. Dann gilt

$$f(x + h) - f(x) \leq 0$$

für alle $h \in B_\epsilon(0)$. Damit gilt für $h_n = \frac{1}{n}$

$$\lim_{n\to\infty} \frac{f(x + h_n) - f(x)}{h_n} \leq 0 \text{ und}$$
$$\lim_{n\to\infty} \frac{f(x - h_n) - f(x)}{-h_n} \geq 0.$$

Da f in x differenzierbar ist, müssen beide Grenzwerte mit

$$\lim_{h\to 0} \frac{f(x + h) - f(x)}{h} = f'(x)$$

übereinstimmen, es muss also $f'(x) = 0$ gelten. □

Bemerkung 2.7. Das Verschwinden der Ableitung ist für die Existenz eines lokalen Extremums nur eine notwendige, keine hinreichende Bedingung.

Theorem 2.8 (Mittelwertsatz der Differentialrechnung). *Seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a, b] \to \mathbb{R}$ eine stetige Funktion die auf (a, b) differenzierbar ist. Dann existiert $\xi \in (a, b)$, so dass*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi).$$

Beweis. Wir betrachten zunächst den Fall $f(a) = f(b)$. Dann ist zu zeigen, dass $\xi \in (a,b)$ existiert, so dass $f'(\xi) = 0$. Da $[a,b]$ beschränkt und abgeschlossen ist, nimmt die Funktion f dort sowohl ihr Maximum als auch ihr Minimum an (Korollar 1.66). Falls sowohl das Maximum, als auch das Minimum in a und b angenommen werden, dann ist f konstant und $f'(\xi) = 0$ für jedes $\xi \in (a,b)$. Andernfalls bezeichne ξ eine der Extremstellen in (a,b). Da ein globales Extremum insbesondere ein lokales Extremum ist, gilt nach Satz 2.6 $f'(\xi) = 0$.

Wir betrachten nun allgemeine f und definieren

$$\begin{aligned} g : [a,b] &\to \mathbb{R} \\ x &\mapsto f(x) - \frac{f(b)-f(a)}{b-a}(x-a). \end{aligned}$$

Dann ist g ebenfalls stetig auf $[a,b]$ und differenzierbar auf (a,b) und es gilt $g(a) = f(a) = g(b)$. Dann existiert nach dem ersten Teil des Beweises $\xi \in (a,b)$ mit

$$0 = g'(\xi) = f'(\xi) - \frac{f(b)-f(a)}{b-a}. \qquad \square$$

Korollar 2.9. *Seien $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$, $a < b$, $f : (a,b) \to \mathbb{R}$ differenzierbar und es gelte $f'(x) = 0$ für alle $x \in (a,b)$. Dann ist f konstant.*

Beweis. Seien $x, y \in (a,b)$ und $y > x$. Dann gibt es nach dem Mittelwertsatz $\xi \in (x,y)$, so dass

$$0 = f'(\xi) = \frac{f(y)-f(x)}{y-x},$$

also gilt $f(x) = f(y)$. Insbesondere gilt also für festes $x \in (a,b)$, dass $f(x) = f(y)$ für alle $y \in (a,b)$, also ist f konstant. $\square$

Korollar 2.10. *Seien $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$, $a < b$, $f : (a,b) \to \mathbb{R}$ differenzierbar. Die Funktion f ist monoton wachsend, genau dann wenn $f'(x) \geq 0$ für alle $x \in (a,b)$. Wenn $f'(x) > 0$ ist, so ist f streng monoton wachsend.*

Beweis. Es gelte zunächst $f'(x) \geq 0$ ($f'(x) > 0$) für alle $x \in (a,b)$. Seien wiederum $x, y \in (a,b)$, $y > x$. Nach dem Mittelwertsatz gibt es $\xi \in (x,y)$, so dass

$$f'(\xi) = \frac{f(y)-f(x)}{y-x}.$$

Damit ist $f(y) \geq f(x)$ ($f(y) > f(x)$). Da x und y beliebig waren, ist f (streng) monoton wachsend.

Sei andererseits f monoton wachsend, dann gilt für jedes x

$$f'(x) = \lim_{h \to 0} \frac{f(x+h)-f(x)}{h} = \lim_{n \to \infty} \frac{f\left(x + \frac{1}{n}\right) - f(x)}{\frac{1}{n}} \geq 0. \qquad \square$$

Bemerkung 2.11. (i) In Korollar 2.9 ist entscheidend, dass die Funktion auf einem Intervall definiert ist. Die Ableitung der Funktion

$$\begin{aligned} f : \mathbb{R}\setminus\{0\} &\to \mathbb{R} \\ x &\mapsto \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases} \end{aligned}$$

verschwindet zwar auf ihrem Definitionsbereich, aber die Funktion ist offensichtlich nicht konstant.

(ii) Die Umkehrung der zweiten Aussage in Korollar 2.10 gilt nicht. Die Funktion $x \mapsto x^3$ ist streng monoton wachsend aber es gilt $f'(0) = 0$.

Korollar 2.12 (Erweiterter Mittelwertsatz). *Seien $a, b \in \mathbb{R}$, $a < b$, $f, g : [a, b] \to \mathbb{R}$ stetige Funktionen die auf (a, b) differenzierbar sind. Dann existiert $\xi \in (a, b)$, so dass*

$$f'(\xi)(g(b) - g(a)) = g'(\xi)(f(b) - f(a)).$$

Beweis. Falls $g(b) = g(a)$ ist, so existiert nach dem Mittelwertsatz ξ mit $g'(\xi) = 0$ und damit ist die Aussage gezeigt.

Sei also $g(b) \neq g(a)$. Wir wenden den Mittelwertsatz auf die Hilfsfunktion

$$h(x) := f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a))$$

an. Es gilt $h(b) = f(a) = h(a)$, also existiert $\xi \in (a, b)$, so dass

$$0 = h'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(\xi).$$

Multiplikation mit $g(b) - g(a)$ ergibt die gewünschte Aussage. □

Satz 2.13 (Regel von de l'Hospital). *Sei $U \subset \mathbb{R}$ offen, $x \in U$ und $f, g : U \to \mathbb{R}$ differenzierbare Funktionen. Falls $f(x) = g(x) = 0$, $g' \neq 0$ in einer punktierten Umgebung von x und*

$$\lim_{y \to x} \frac{f'(y)}{g'(y)}$$

existiert, dann gilt

$$\lim_{y \to x} \frac{f(y)}{g(y)} = \lim_{y \to x} \frac{f'(y)}{g'(y)}.$$

Beweis. Wir zeigen zunächst, dass x eine isolierte Nullstelle von g ist. Angenommen x ist Häufungspunkt von Nullstellen von g, das heißt es existieren Nullstellen $x_n \neq x$ für alle $n \in \mathbb{N}$, so dass $x_n \to x$. Wir können oBdA. annehmen, dass (x_n) streng monoton ist (sonst wähle eine Teilfolge aus). Sei (x_n) streng monoton wachsend. Nach dem Mittelwertsatz

existiert für jedes $n \in \mathbb{N}$ ein $\xi_n \in (x_n, x_{n+1})$ mit $g'(\xi_n) = 0$. Da $\xi_n \to x$ konvergieren, steht das im Widerspruch zur Voraussetzung $g' \neq 0$ auf einer punktierten Umgebung von x. Für den Fall streng monoton fallender Folge (x_n) liefert ein analoges Argument ebenfalls einen Widerspruch.

Sei nun $\dot{V}$ eine Umgebung von x, so dass $g \neq 0$ und $g' \neq 0$ auf $\dot{V}$. Wähle eine beliebige Folge $x_n \in V$, so dass $x_n \to x$. Für jedes $n \in \mathbb{N}$ existiert nach dem erweiterten Mittelwertsatz $\xi_n \in (\min\{x, x_n\}, \max\{x, x_n\})$, so dass

$$\frac{f'(\xi_n)}{g'(\xi_n)} = \frac{f(x) - f(x_n)}{g(x) - g(x_n)} = \frac{f(x_n)}{g(x_n)}.$$

Es gilt $\lim_{n\to\infty} \xi_n \to x$. Nach Voraussetzung existiert

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

Da die Folge (x_n) beliebig war, folgt mit dem Folgenkriterium für Konvergenz von Funktionen die Existenz von

$$\lim_{y\to x} \frac{f(y)}{g(y)} = \lim_{y\to x} \frac{f'(y)}{g'(y)}. \qquad \square$$

Beispiel 2.14. Wir wollen mittels der Regel von l'Hospital den Grenzwert

$$\lim_{x\to 0} \frac{1 - \cos x}{x^2}$$

bestimmen. Es sind also $f(x) = 1 - \cos x$, $f'(x) = \sin x$ und $f''(x) = \cos x$ sowie $g(x) = x^2$, $g'(x) = 2x$ und $g''(x) = 2$. Da $f'(0) = g'(0) = 0$ und wegen der Stetigkeit des cos

$$\lim_{x\to 0} \frac{f''(x)}{g''(x)} = \frac{1}{2}$$

ist, existiert nach l'Hospital auch

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \frac{1}{2}.$$

Da auch $f(x) = g(x) = 0$ ist und der obige Grenzwert existiert, gilt dann auch

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = \frac{1}{2}.$$

Definition 2.15. Sei $U \subset \mathbb{R}$ offen und $f : U \to \mathbb{R}$ eine Funktion. Die Funktion heißt $n+1$ mal differenzierbar wenn die n-te Ableitung $f^{(n)}$ differenzierbar ist und die $n+1$-te Ableitung ist die Funktion $f^{(n)\prime}$. Die Funktion heißt beliebig differenzierbar oder glatt falls sie n mal differenzierbar ist für jedes $n \in \mathbb{N}$.

Die Funktion heißt n mal stetig differenzierbar, falls $f^{(n)}$ stetig ist.

Definition 2.16. Sei $U \subset \mathbb{R}$ offen und $f : U \to \mathbb{R}$ eine in x differenzierbare Funktion. Dann heißt

$$T_n(y) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k$$

das n-te Taylor-Polynom um den Entwicklungspunkt x.

Theorem 2.17 (Satz von Taylor). *Sei $(a,b) \subset \mathbb{R}$ ein offenes Intervall und $f : (a,b) \to \mathbb{R}$ eine $n+1$ mal differenzierbare Funktion. Seien $x, y \in (a,b)$ zwei verschiedene Punkte und I das abgeschlossene, $\mathring{I}$ das offene von x und y aufgespannte Intervall. Sei $h : I \to \mathbb{R}$ eine stetige auf $\mathring{I}$ differenzierbare Funktion, so dass $h' \neq 0$ auf $\mathring{I}$. Dann existiert $\xi \in \mathring{I}$, so dass*

$$f(y) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k + \frac{f^{(n+1)}(\xi)}{n!}(y-\xi)^n \frac{h(y)-h(x)}{h'(\xi)}. \tag{2.1}$$

Beweis. Wir wenden den verallgemeinerten Mittelwertsatz auf

$$\begin{aligned} g : (a,b) &\to \mathbb{R} \\ t &\mapsto \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(y-t)^k \end{aligned}$$

und h an, es existiert also $\xi \in \mathring{I}$, so dass

$$\frac{g'(\xi)}{h'(\xi)}(h(y)-h(x)) = g(y) - g(x) = f(y) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k.$$

Wir müssen nun noch g' ausrechnen:

$$\begin{aligned} g'(t) &= \sum_{k=0}^{n} \left(\frac{f^{(k+1)}(t)}{k!}(y-t)^k - \frac{f^{(k)}(t)}{k!}k(y-t)^{k-1} \right) \\ &= \frac{f^{(n+1)}(t)}{n!}(y-t)^n \end{aligned}$$

wobei die letzte Umformung aus der Tatsache folgt, dass es sich um eine Teleskopsumme handelt. Einsetzen von ξ liefert dann die gesuchte Gleichung. □

Bemerkung 2.18. (i) Je nach Wahl der Funktion h erhält man nun verschiedene Formen des Restgliedes

$$R_{n+1}(y,x) := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(y-x)^k.$$

(ii) Für $h(t) = (y-t)^{n+1}$ also $h'(t) = -(n+1)(y-t)^n$ ergibt sich die Lagrange-Form des Restgliedes:

$$R_{n+1}(y,x) = \frac{f^{(n+1)}(\xi)}{n!}(y-\xi)^n \frac{(y-x)^{n+1}}{(n+1)(y-\xi)^n} = \frac{f^{(n+1)}(\xi)}{(n+1)!}(y-x)^{n+1}$$

Korollar 2.19. *Sei $(a,b) \subset \mathbb{R}$ ein intervall und $f : (a,b) \to \mathbb{R}$ mindestens n mal stetig differenzierbar. Sei $x \in (a,b)$ und es gelte $0 = f'(x) = f''(x) = \cdots = f^{(n-1)}(x)$ und $f^{(n)}(x) \neq 0$. Ist n ungerade so liegt bei x kein lokales Extremum vor, ist n gerade, so hat f bei x ein lokales Minimum falls $f^{(n)}(x) > 0$ und ein lokales Maximum falls $f^{(n)}(x) < 0$.*

Beweis. Wir setzten oBdA. voraus, dass $f^{(n)}(x) > 0$ ist (sonst betrachte $-f$). Wir verwenden die Taylor-Formel mit Lagrange-Restglied. Nach Voraussetzung verschwinden in der Entwicklung (2.1) die Terme 1 bis $n-1$. Da nach Voraussetzung $f^{(n)}(x)$ stetig ist, existiert $\epsilon > 0$, so dass $f^{(n)}(x) > 0$ auf ganz $(x-\epsilon, x+\epsilon)$. Dann existiert nach der Taylor-Formel für jedes $y \in (x-\epsilon, x+\epsilon)$ ein $\xi_y \in (x-\epsilon, x+\epsilon)$, so dass

$$f(y) = f(x) + \frac{f^{(n)}(\xi_y)}{n!}(y-x)^n.$$

Falls n ungerade ist, so ist $f(x-\delta) < f(x) < f(x+\delta)$ für jedes $\delta > 0$. Es liegt also kein lokales Extremum vor. Ist n gerade, so ist der zweite Term stets nichtnegativ, es gilt also $f(y) \geq f(x)$ für alle $y \in (x-\epsilon, x+\epsilon)$, an x liegt also ein lokales Minimum vor. □

Satz 2.20. *Sei $U \subset \mathbb{R}$ offen und $x \in U$. Eine Funktion $f : U \to \mathbb{K}$ ist differenzierbar in x, genau dann wenn $a \in \mathbb{K}$ und $\varphi : U \to \mathbb{K}$ existieren, so dass*

$$f(y) = f(x) + a(y-x) + \varphi(y)\,|y-x| \tag{2.2}$$

für alle $y \in U$ und $\lim_{y\to x} \varphi(y) = 0$.

Beweis. Seien zunächst $a \in \mathbb{K}$ und $\varphi : U \to \mathbb{K}$ so, dass $\lim_{y\to x} \varphi(y) = 0$ und Gleichung (2.2) gilt. Dann gilt

$$\frac{f(y)-f(x)}{y-x} = a + \varphi(y)\frac{|y-x|}{y-x}$$

und die rechte Seite dieser Gleichung konvergiert für $y \to x$ nach Voraussetzung gegen a. Damit ist f in x differenzierbar und $f'(x) = a$.

Sei andererseits f in x differenzierbar und definiere

$$\varphi(y) := \begin{cases} \left(\frac{f(y)-f(x)}{y-x} - f'(x)\right)\frac{y-x}{|y-x|} & y \neq x \\ 0 & y = x \end{cases}.$$

Dann gilt Gleichung (2.2) (für $a = f'(x)$) und $\lim_{y\to x} \varphi(y) = 0$. □

2.2 Differenzierbarkeit von Funktionen mehrerer Veränderlicher

Wiederholung lineare Abbildungen und Matrizen

Im Folgenden seien V, W stets Vektorräume über einem Körper $\mathbb{K}$.

Definition 2.21. Eine Abbildung $\varphi : V \to W$ heißt linear, wenn

$$\varphi(v + \lambda\tilde{v}) = \varphi(v) + \lambda\varphi(\tilde{v})$$

für alle $v, \tilde{v} \in V$ und alle $\lambda \in \mathbb{K}$.

Bemerkung 2.22. (i) Die Menge $\mathcal{L}(V|W)$ aller linearen Abbildungen zwischen festen Räumen V und W ist selbst ein Vektorraum über $\mathbb{K}$.

(ii) Seien $e_1, \ldots, e_k$ und $f_1, \ldots, f_l$ Basen für V beziehungsweise W, dann ist durch

$$\varphi(e_j) = \sum_{i=1}^{l} a_{ij} f_i \tag{2.3}$$

für alle $j \in \{1, \ldots, k\}$ eindeutig eine Matrix $A_\varphi = (a_{ij})$ definiert, die sogenannte Darstellungs- oder Abbildungsmatrix der linearen Abbildung φ bezüglich der gegebenen Basen.

Umgekehrt wird durch jede $l \times k$-Matrix $A = (a_{ij})$ mittels der Gleichungen (2.3) eindeutig eine lineare Abbildung φ_A definiert.

(iii) Für fest gewählte Basen von V und W hängt die Darstellungsmatrix, also die Abbildung

$$\begin{aligned} \mathcal{L}(V|W) &\to \mathbb{R}^{l\times k} \\ \varphi &\mapsto A_\varphi \end{aligned}$$

linear von φ ab, es gilt also für lineare Abbildungen $\varphi, \psi : V \to W$ und $\lambda \in \mathbb{K}$

$$A_{\varphi+\lambda\psi} = A_\varphi + \lambda A_\psi.$$

(iv) Seien $v = v_1 e_1 + \cdots v_k e_k \in V$ und $w = \varphi(v) = w_1 f_1 + \cdots w_l f_l$, dann gilt

$$\begin{pmatrix} w_1 \\ \vdots \\ w_l \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{l1} & \cdots & a_{lk} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_k \end{pmatrix}$$

die Anwendung von φ auf v entspricht also der Matrizenmultiplikation der Matrixdarstellungen von φ und v.

(v) Falls U ein weiterer Vektorraum mit fest gewählter Basis ist und $\varphi : V \to W$ und $\psi : W \to U$ lineare Abbildungen sind, so gilt für die zugehörigen Darstellungsmatrizen

$$A_{\psi\circ\varphi} = A_\psi A_\varphi,$$

die Hintereinanderausführung der Abbildungen entspricht also der Matrixmultiplikation der Darstellungsmatrizen.

(vi) Falls $V = \mathbb{R}^k$ und $W = \mathbb{R}^l$ gilt, so gibt es auf jedem der Räume die Standardbasis. Wird nichts anderes gesagt, so beziehen sich alle Darstellungsmatrizen auf diese Standardbasen und wir werden die lineare Abbildung φ mit ihrer Abbildungsmatrix A_φ identifizieren.

(vii) Auf $\mathbb{R}^n$ existiert ebenfalls das Standardskalarprodukt $\langle \cdot | \cdot \rangle$. Für

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \qquad \tilde{v} = \begin{pmatrix} \tilde{v}_1 \\ \vdots \\ \tilde{v}_n \end{pmatrix}$$

erhalten wir

$$\langle v | \tilde{v} \rangle = \sum_{i=1}^{n} v_i \tilde{v}_i = (v_1, \dots, v_n) \begin{pmatrix} \tilde{v}_1 \\ \vdots \\ \tilde{v}_n \end{pmatrix}.$$

Definition 2.23. Sei $U \subset \mathbb{R}^n$ offen, $x = (x_1, \cdots, x_n) \in U$ und $f : U \to \mathbb{R}^m$. Die Abbildung f heißt partiell differenzierbar in x bezüglich x_i, wenn die Funktionen $t \mapsto f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots x_n)$ differenzierbar ist, das heißt wenn der Grenzwert

$$\partial_i f(x) := \lim_{h\to 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots x_n) - f(x_1, \dots, x_n)}{h}$$

existiert. Der Grenzwert $\partial_i f(x) \in \mathbb{R}^m$ heißt die partielle Ableitung von f an der Stelle x bezüglich x_i und wird auch als $\frac{\partial f}{\partial x_i}(x)$ geschrieben.

Die Abbildung f heißt partiell differenzierbar in x, wenn sie partiell differenzierbar bezüglich x_i für jedes $i \in \{1, \dots n\}$ ist.

Die Existenz nur der partiellen Ableitungen ist in vielen Fällen eine zu schwache Bedingung wie das folgende Beispiel zeigt.

Beispiel 2.24. Die Funktion

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto \begin{cases} 1 & x = 0 \text{ oder } y = 0 \\ 0 & \text{sonst} \end{cases} \end{aligned}$$

ist im Punkt $(0, 0)$ partiell differenzierbar, aber nicht stetig.

Über lediglich partiell differenzierbare Funktionen können wir kaum Aussagen treffen, es bedarf also eines stärkeren Differenzierbarkeitsbegriffes. Dazu bietet sich die Verallgemeinerung der in Satz 2.20 eingeführten Bedingung an.

Definition 2.25. Sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und $f : U \to \mathbb{R}^m$. Die Abbildung f heißt in x differenzierbar oder total differenzierbar, falls eine lineare Abbildung $A \in \mathcal{L}(\mathbb{R}^n | \mathbb{R}^m)$ und $\varphi : U \to \mathbb{R}^m$ existieren, so dass

$$f(y) = f(x) + A(y - x) + \varphi(y) \, \|y - x\| \tag{2.4}$$

für alle $y \in U$ und $\lim_{y \to x} \varphi(y) = 0$.

Genauso können wir Differenzierbarkeit von Abbildungen zwischen beliebigen normierten Vektorräumen V und W definieren.

Satz 2.26. *Sei $U \subset \mathbb{R}^n$ offen, $x \in U$ und $f : U \to \mathbb{R}^m$. Seien $f_1, \ldots, f_m : U \to \mathbb{R}$ die Komponentenfunktionen von f, das heißt $f(x) = (f_1(x), \ldots, f_m(x))$ für alle $x \in X$. Wenn die Abbildung f total differenzierbar ist, so ist sie partiell differenzierbar und die Darstellungsmatrix der Abbildung A aus Definition 2.25 ist gegeben durch $A_{ji} = \partial_i f_j(x)$.*

Beweis. Seien A und φ wie in Definition 2.25. Seien $e_1, \ldots, e_n$ und $e'_1, \ldots, e'_m$ die Standardbasen von $\mathbb{R}^n$ beziehungsweise $\mathbb{R}^m$. Wir setzen $y = x + he_i$ in (2.4) ein und erhalten

$$\begin{aligned} f(x + he_i) &= f(x) + hA(e_i) + \varphi(x + he_i) \, |h| \\ \frac{f(x + he_i) - f(x)}{h} &= A(e_i) + \varphi(x + he_i) \frac{|h|}{h}. \end{aligned}$$

Für $h \to 0$ konvergiert diese Gleichung gegen

$$\sum_{j=1}^{m} \partial_i f_j(x) e'_j = \partial_i f(x) = A(e_i) = \sum_{j=1}^{m} A_{ji} e'_j,$$

also existiert die partielle Ableitung von f bezüglich x_i. Aus der Eindeutigkeit der Basisdarstellung erhält man durch Koeffizientenvergleich die Gleichung $\partial_i f_j(x) = A_{ji}$. □

Damit ist insbesondere gezeigt, dass A in (2.4) eindeutig bestimmt ist, wir können also von dieser Matrix (oder vielmehr von der linearen Abbildung die sie darstellt) als der totalen Ableitung von f sprechen.

Definition 2.27. Die Matrix $(\partial_i f_j(x))_{ji}$ heißt Jacobi-Matrix (auch Funktionalmatrix oder Ableitungsmatrix) von f an der Stelle x. Wir werden sowohl diese Matrix, als auch die lineare Abbildung die sie repräsentiert, mit $\mathrm{D}f(x)$ bezeichnen auch dann, wenn die Funktion nicht total differenzierbar ist.

Für reellwertige f, also $f : \mathbb{R}^n \to \mathbb{R}$, hat die Jakobi-Matrix lediglich eine Spalte und wir bezeichnen sie auch als Gradient von f und schreiben $\mathrm{grad} f := \mathrm{D}f$.

Beispiel 2.28. (i) Seien V und W (endlichdimensionale) Vektorräume und $A : V \to W$ eine lineare Abbildung. Dann ist A total differenzierbar und es gilt $\mathrm{D}A(x) = A$ für alle $x \in V$. Um das zu sehen schreiben wir

$$A(y) = A(x) + A(y - x)$$

womit (2.4) für $\varphi \equiv 0$ erfüllt ist.

(ii) Betrachte die Abbildung

$$\begin{aligned} f : (0, \infty) \times (0, 2\pi) &\to \mathbb{R}^2 \\ (r, \varphi) &\mapsto (r \cos \varphi, r \sin \varphi). \end{aligned}$$

Dann ist f partiell differenzierbar und es gilt

$$Df(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}.$$

Bemerkung 2.29. (i) Eine Abbildung f ist also total differenzierbar in x genau dann, wenn die Jacobi-Matrix $\mathrm{D}f(x)$ existiert und

$$\lim_{y \to x} \frac{f(x) - f(y) - \mathrm{D}f(x)(y - x)}{\|y - x\|} = 0.$$

(ii) Aus dieser Formulierung ist ersichtlich, dass eine Abbildung $f = (f_1, \dots f_m)$, mit $f_1, \dots, f_m : U \to \mathbb{R}$ total differenzierbar ist, genau dann wenn f_i total differenzierbar ist für jedes $i \in \{1, \dots m\}$.

(iii) Wie im eindimensionalen Fall folgt aus der totalen Differenzierbarkeit insbesondere die Stetigkeit von f.

(iv) Sowohl partielle, als auch totale Differenzierbarkeit sind lokale Eigenschaften.

(v) Um Verwirrung zu vermeiden: Wenn f auf einer offenen Menge U differenzierbar ist, dann ist die Abbildung $h \mapsto \mathrm{D}f(x)h$ linear für jedes $x \in U$. Die Abbildung $x \mapsto \mathrm{D}f(x)$ ist in den meisten Fällen nicht linear.

Definition 2.30. Seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ normierte Vektorräume. Dann definiert

$$\|\varphi\|_{\mathcal{L}(V|W)} := \sup \{ \|\varphi(v)\|_W | \ v \in V, \|v\|_V = 1\}$$

eine Norm auf $\mathcal{L}(V|W)$.

Bemerkung 2.31. Der Beweis der Normeigenschaften verläuft analog zu Aufgabe 1 vom Übungsblatt 0,5. Die obige Norm ist auf $\mathcal{L}(V|W)$ für uns die Standardnorm, wir werden

also häufig nur $\|\varphi\|$ schreiben wenn durch das Argument klar ist von welcher Norm die Rede ist. Es gilt die Ungleichung

$$\|\varphi(x)\| \leq \|\varphi\| \, \|x\| \, .$$

Damit folgt insbesondere die Lipschitz-Stetigkeit (und damit Stetigkeit) linearer Abbildungen:

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

Satz 2.32. *Seien $U \subset \mathbb{R}^n$ und $V \subset \mathbb{R}^m$ offen, $x \in U$, $g : U \to V$ total differenzierbar in x und $f : V \to \mathbb{R}^k$ total differenzierbar in $g(x)$. Dann ist $f \circ g$ total differenzierbar in x und es gilt*

$$\mathrm{D}(f \circ g)(x) = \mathrm{D}f(g(x))\mathrm{D}g(x).$$

Beweis. Differenzierbarkeit von g im Punkt x bedeutet, dass $\varphi_g : U \to \mathbb{R}^m$ existiert, so dass $\lim_{y \to x} \varphi_g(y) = 0$ und

$$g(y) - g(x) = \mathrm{D}g(x)(y - x) + \varphi_g(y) \, \|y - x\| \, .$$

Insbesondere gilt also

$$\|g(y) - g(x)\| \leq \|\mathrm{D}g(x)(y - x)\| + \|y - x\| \, \|\varphi_g(y)\| \leq (\|\mathrm{D}g(x)\| + \|\varphi_g(y)\|) \, \|y - x\| \, .$$

Da $\varphi_g(y)$ für $y \to x$ gegen 0 konvergiert, ist es insbesondere beschränkt in einer Umgebung von x. Es gibt also ein $K \in \mathbb{R}$, so dass in einer punktierten Umgebung von x

$$\frac{\|g(y) - g(x)\|}{\|y - x\|} \leq K \tag{2.5}$$

gilt.

Differenzierbarkeit von f im Punkt $g(x)$ bedeutet, dass es ein $\varphi_f : V \to \mathbb{R}^k$ gibt, so dass $\lim_{z \to g(x)} \varphi_f(z) = 0$ und

$$f(z) = f(g(x)) + \mathrm{D}f(g(x))(z - g(x)) + \varphi_f(z) \, \|z - g(x)\| \, .$$

Wir setzen $z = g(y)$ und benutzen die obige Formel für $g(y) - g(x)$

$$\begin{aligned} f(g(y)) &= f(g(x)) + \mathrm{D}f(g(x))(g(y) - g(x)) + \varphi_f(g(y)) \, \|g(y) - g(x)\| \\ (f \circ g)(y) &= (f \circ g)(x) + \mathrm{D}f(g(x))\mathrm{D}g(x)(y - x) \\ &\quad + \left(\varphi_f(g(y)) \frac{\|g(y) - g(x)\|}{\|y - x\|} + \mathrm{D}f(g(x))\varphi_g(y) \right) \|y - x\| \, . \end{aligned}$$

Um den Beweis abzuschließen muss gezeigt werden, dass für $y \to x$ der Ausdruck in Klammern gegen 0 konvergiert. Da g in x differenzierbar ist, ist es insbesondere stetig und so gilt $\lim_{y \to x} g(y) = g(x)$. Dann gilt $\lim_{y \to x} \varphi_f(g(y)) = 0$ und wegen (2.5) konvergiert der erste Summand in der Klammer gegen 0. Nach Voraussetzung konvergiert $\varphi_g(y)$ genen 0 und da $D(f(g(x)))$ linear und somit stetig ist konvergiert auch der zweite Summmand in der Klammer gegen 0. □

Satz 2.33. *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$. Wenn für jedes $x \in U$ und jedes $i \in \{1, \dots, n\}$ die partielle Ableitung $\partial_i f(x)$ existiert und stetig ist, dann ist f für jedes x total differenzierbar.*

Lemma 2.34. *Sei $x \in U$, $r > 0$ und $f : B_r(x) \to \mathbb{R}$ partiell differenzierbar. Sei $h = (h_1, \dots, h_n) \in B_r(0)$. Dann existieren $\xi_1, \dots, \xi_n \in K_{\|h\|}(x)$, so dass*

$$f(x+h) - f(x) = \partial_1 f(\xi_1)h_1 + \cdots + \partial_n f(\xi_n)h_n.$$

Beweis. Sei $e_1, \dots, e_n$ die Standardbasis des $\mathbb{R}^n$ und definiere $x_0 = x$, $x_i = x_{i-1} + h_i e_i$. Dann ist $x_n = x + h$. Weiter gilt

$$\|x_i - x\| = \|h_1 e_1 + \cdots + h_i e_i\| = \sqrt{|h_1|^2 + \cdots + |h_i|^2} \leq \sqrt{|h_1|^2 + \cdots + |h_n|^2} = \|h\|$$

die Punkte x_i sind also alle in $K_{\|h\|}(x) \subset B_r(x)$ enthalten.

Für ein festes $i \in \{1, \dots, n\}$ verfahren wir wie folgt, falls $h_i \neq 0$ ist. Im Folgenden betrachten wir den Fall $h_i > 0$, der Fall $h_i < 0$ kann jedoch analog behandelt werden. Wähle $\epsilon > 0$, so dass $B_\epsilon(x_{i-1}), B_\epsilon(x_i) \subset B_r(x)$. Dann gilt für die Funktion

$$\begin{aligned} g : (-\epsilon, h_i + \epsilon) &\to \mathbb{R} \\ t &\mapsto f(x_{i-1} + te_i), \end{aligned}$$

dass

$$\lim_{s \to 0} \frac{g(t+s) - g(t)}{s} = \lim_{s \to 0} \frac{f(x_{i-1} + (t+s)e_i) - f(x_{i-1} + te_i)}{s} = \partial_i f(x_{i-1} + te_i).$$

Damit ist g differenzierbar und es gilt $g'(t) = \partial_i f(x_{i-1} + te_i)$. Insbesondere ist g also stetig auf $[0, h_i]$. Jetzt können wir den Mittelwertsatz anwenden um $\theta \in (0, h_i)$ zu erhalten, so dass

$$\partial_i f(x_{i-1} + \theta e_i)h_i = g'(\theta)h_i = g(h_i) - g(0) = f(x_i) - f(x_{i-1}).$$

Für $\xi_i := x_{i-1} + \theta e_i$ erhalten wir also

$$\partial_i f(\xi_i)h_i = f(x_i) - f(x_{i-1}).$$

Da die Kugel $K_{\|h\|}(x)$ konvex ist und x_i und x_{i-1} enthält, ist auch $\xi_i \in K_{\|h\|}(x)$.

Schließlich erhalten wird

$$f(x+h) - f(x) = \sum_{i=1}^n f(x_i) - f(x_{i-1}) = \sum_{i=1}^n \partial_i f(\xi_i)h_i.$$

Dabei ist zu beachten, dass falls $h_i = 0$ ist, sowohl $f(x_i) - f(x_{i-1})$ als auch $f(\xi_i)h_i$ verschwinden, unabhängig vom Wert von ξ_i. Wir können für diese Fälle also willkürlich z.B. $\xi_i = x_i$ setzen. $\square$

Beweis von Satz 2.33. Da totale Differenzierbarkeit komponentenweise festgestellt werden kann, reicht es den Fall $m = 1$ zu betrachten. Wähle $x \in U$ fest und sei $\epsilon > 0$, so dass $B_\epsilon(x) \subset U$. Nach dem vorhergehenden Lemma existieren für jedes $h = (h_1, \dots, h_n) \in B_\epsilon(0)$ Punkte $\xi_1(h), \dots, \xi_n(h) \in K_{\|h\|}(x)$, so dass

$$f(x+h) - f(x) = \partial_1 f(\xi_1(h))h_1 + \cdots + \partial_n f(\xi_n(h))h_n.$$

Wir setzen

$$a(h) := (\partial_1 f(\xi_1(h)) - \partial_1 f(x), \cdots, \partial_n f(\xi_n(h)) - \partial_n f(x))$$

und erhalten

$$\begin{aligned} &f(x+h) - f(x) - \mathrm{D}f(x)h \\ &\qquad = (\partial_1 f(\xi_1(h)) - \partial_1 f(x))h_1 + \cdots + (\partial_n f(\xi_n(h)) - \partial_n f(x))h_n = \langle a(h) | h \rangle, \end{aligned}$$

wobei $\langle \cdot | \cdot \rangle$ das Standardskalarprodukt im $\mathbb{R}^n$ ist.

Dann folgt mit der Cauchy-Schwarz-Ungleichung

$$|f(x+h) - f(x) - \mathrm{D}f(x)h| = |\langle a(h) | h \rangle| \leq \|a(h)\| \, \|h\| \, .$$

Da $\|\xi_i(h) - x\| \leq \|h\|$ gilt $\lim_{h \to 0} \xi_i(h) = x$ für jedes $i \in \{1, \dots, n\}$. Nach Voraussetzung ist $\partial_i f$ stetig, also gilt $\lim_{h \to 0} \partial_i f(\xi_i(h)) = \partial_i f(x)$. Also konvergiert $a(h)$ gegen 0 und es gilt

$$\lim_{h \to 0} \frac{|f(x+h) - f(x) - \mathrm{D}f(x)h|}{\|h\|} \leq \lim_{h \to 0} \|a(h)\| = 0.$$

Damit ist f in x total differenzierbar und da $x \in U$ beliebig war ist f total differenzierbar auf ganz U. □

Beispiel 2.35. Seien $N \in \mathbb{N}$ und $c_\eta \in \mathbb{K}$ Koeffizienten für jeden Multiindex $\eta \in \mathbb{N}_0^n$ mit $|\eta| \leq \mathbb{N}$. Dann ist das Polynom

$$\begin{aligned} p : \mathbb{R}^n &\to \mathbb{K} \\ x &\mapsto \sum_{\eta, |\eta| < N} c_\eta x^\eta \end{aligned}$$

total differenzierbar.

Dazu sehen wir uns die partielle Ableitung eines Monoms an:

$$\partial_i x^\eta = \partial_i x_1^{\eta_1} \cdots x_{i-1}^{\eta_{i-1}} x_i^{\eta_i} x_{i+1}^{\eta_{i+1}} \cdots x_n^{\eta_n} = \eta_i x_1^{\eta_1} \cdots x_{i-1}^{\eta_{i-1}} x_i^{\eta_i - 1} x_{i+1}^{\eta_{i+1}} \cdots x_n^{\eta_n}$$

und stellen fest, dass diese ein Polynom ist. Damit ist jede partielle Ableitung eines Polynoms wiederum ein Polynom also insbesondere stetig. Dann folgt mit Satz 2.33, dass p total differenzierbar ist.

Korollar 2.36. *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$. Dann ist f stetig differenzierbar, das heißt die Abbildung $x \mapsto \mathrm{D}f(x)$ ist stetig, genau dann wenn es partiell differenzierbar ist und alle partiellen Ableitungen stetig sind.*

Beweis. Das folgt sofort, da Stetigkeit der Abbildung $x \mapsto \mathrm{D}f(x)$ äquivalent zur Stetigkeit der einzelnen Komponentenfunktionen, also der partiellen Ableitungen ist. □

Satz 2.37 (Mittelwertsatz für reellwertige Funktionen). *Sei $U \subset \mathbb{R}^n$, $x, y \in U$ und $S_{xy} := \{x + t(y - x) |\ t \in [0, 1]\} \subset U$. Sei $f : U \to \mathbb{R}$ differenzierbar auf S_{xy} und stetig in x und y, dann existiert $\xi \in S_{xy}$, so dass*

$$f(y) - f(x) = \mathrm{D}f(\xi)(y - x).$$

Beweis. Wir definieren die Hilfsfunktion

$$\begin{aligned} g : [0, 1] &\to \mathbb{R} \\ t &\mapsto f(x + t(y - x)). \end{aligned}$$

Nach der Kettenregel und der Voraussetzung ist g differenzierbar auf $(0, 1)$ und stetig auf $[0, 1]$, also können wir den (eindimensionalen) Mittelwertsatz anwenden. Wir erhalten also $\theta \in (0, 1)$, so dass $g(1) - g(0) = g'(\theta)$. Daraus folgt mit der Kettenregel

$$f(y) - f(x) = g(1) - g(0) = g'(\theta) = \mathrm{D}f(x + \theta(y - x))(y - x)$$

also für $\xi = x + \theta(y - x)$ die gesuchte Aussage. □

Satz 2.38 (Allgemeiner Mittelwertsatz). *Sei $U \subset \mathbb{R}^n$, $x, y \in U$ und*

$$S_{xy} := \{x + t(y - x) |\ t \in [0, 1]\} \subset U.$$

Ist $f : U \to \mathbb{R}^m$ differenzierbar dann gilt

$$\|f(y) - f(x)\| \leq \sup_{t \in [0,1]} \|\mathrm{D}f(x + t(y - x))\| \, \|y - x\| \, .$$

Ohne Beweis.

Korollar 2.39. *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ differenzierbar. Falls $\mathrm{D}f \equiv 0$ auf U gilt, so ist f lokal konstant, das heißt für jedes $x \in U$ existiert eine Umgebung V, so dass f auf V konstant ist.*

Beweis. Sei $x \in U$ und wähle $\epsilon > 0$ so, dass $B_\epsilon(x) \subset U$. Dann gilt für alle $y \in B_\epsilon(x)$ nach dem allgemeinen Mittelwertsatz

$$\|f(y) - f(x)\| \leq \sup_{t \in [0,1]} \|\mathrm{D}f(x + t(y - x))\| \, \|y - x\| = 0$$

also $f(y) = f(x)$. Damit ist f auf $B_\epsilon(x)$ konstant. □

2.3 Multilineare Abbildungen und höhere Ableitungen

Definition 2.40. Seien $V_1, \ldots, V_n, W$ Vektorräume. Die Abbildung $\varphi : V_1 \times \cdots \times V_n \to W$ heißt multilinear falls sie linear in jedem ihrer Argumente ist, das heißt die Abbildung von $V_i \to V$

$$v \mapsto \varphi(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n)$$

ist linear für jedes $i \in \{1, \ldots, n\}$ und alle $v_k \in V_k$ für $k \in \{1, \ldots, n\}$ und $k \neq i$.

Wir bezeichnen den Raum der multilinearen Abbildungen von $\varphi : V_1 \times \cdots \times V_n \to W$ mit $\mathcal{L}(V_1, \ldots, V_n | W)$. Falls $V_1 = V_2 = \cdots = V_n = V$ ist, so schreiben wir $\mathcal{L}^n(V|W)$.

Beispiel 2.41. (i) Multilineare Abbildungen sind Verallgemeinerungen von linearen Abbildungen.

(ii) Die Determinante ist eine multilineare Abbildung von $\mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$.

(iii) Ist V ein reeller Vektorraum, dann ist jedes Skalarprodukt $\langle \cdot | \cdot \rangle$ eine multilineare Abbildung von $V \times V \to \mathbb{R}$.

(iv) Für Vektorräume V und W ist die Abbildung

$$\begin{aligned} \mathcal{L}(V|W) \times V &\to W \\ (A, v) &\mapsto A(v) \end{aligned}$$

eine multilineare Abbildung.

Bemerkung 2.42. (i) Multilineare Abbildungen der Form $V \times V \times \cdots \times V \to \mathbb{R}$, sogenannte Multilinearformen, sind (für endlichdimensionales V) eine Möglichkeit (kovariante) Tensoren über dem Vektorraum V einzuführen.

(ii) Genau wie für lineare Abbildungen, ist auch eine multilineare Abbildungen φ eindeutig durch ihre Werte bezüglich festgeleter Basen von $V_1, \ldots, V_n$ und W bestimmt. Wir betrachten den Fall $V_1 = V_2 = \cdots = V_n = V$. Es seien Basen $e_1, \ldots, e_k$ von V und $f_1, \ldots, f_l$ von W gegeben. Für vorgegebenes $\varphi : V^n \to W$ definieren die Gleichungen

$$\varphi(e_{i_1}, e_{i_2}, \cdots, e_{i_n}) = \sum_{j=1}^{l} a_{ji_1i_2\cdots i_n} f_j$$

für $i_1, \ldots, i_n \in \{1, \ldots, k\}$ die Komponenten $a_{ji_1i_2\cdots i_n}$ der multilinearen Abbildung φ. Für vorgegebene Komponenten $a_{ji_1i_2\cdots i_n}$, $i_1, \ldots, i_n \in \{1, \ldots, k\}$ und $j \in \{1, \ldots, l\}$ definieren obige Gleichungen eindeutig eine multilineare Abbildung φ.

(iii) Insbesondere ist für $V = W = \mathbb{R}$ jede multilineare Abbildung durch genau eine Zahl bestimmt.

Satz 2.43. *Seien V, W Vektorräume, dann ist die Abbildung $\Phi : \mathcal{L}(V|\mathcal{L}(V|W)) \to \mathcal{L}(V, V|W)$, definiert durch*

$$[\Phi(\varphi)]\,(v_1, v_2) := [\varphi(v_1)]\,(v_2)$$

ein Isomorphismus von Vektorräumen.

Beweis. Zunächst überprüfen wir die Linearität. Seien dazu $\varphi, \tilde{\varphi} \in \mathcal{L}(V|\mathcal{L}(V|W))$ und $\lambda \in \mathbb{R}$. Dann gilt für $v_1, v_2 \in V$

$$\begin{aligned}\Phi(\varphi + \lambda\tilde{\varphi})(v_1, v_2) &= (\varphi + \lambda\tilde{\varphi})(v_1)(v_2) = \varphi(v_1)(v_2) + \lambda\tilde{\varphi}(v_1)(v_2)\\ &= \Phi(\varphi)(v_1, v_2) + \lambda\Phi(\varphi)(v_1, v_2),\end{aligned}$$

also $\Phi(\varphi + \lambda\tilde{\varphi}) = \Phi(\varphi) + \lambda\Phi(\tilde{\varphi})$.

Wir müssen nun noch Injektivität und Surjektivität zeigen. Sei $\Phi(\varphi) = 0$, das heißt es gilt $0 = \Phi(\varphi)(v_1, v_2) = \varphi(v_1)(v_2)$ für alle $v_1, v_2 \in V$. Dann ist aber $\varphi(v_1) = 0$ für alle $v_1 \in V$ und es folgt auch $\varphi = 0$. Die Abbildung Φ ist also injektiv.

Sei nun $\theta \in \mathcal{L}(V, V|W)$. Wir setzen $\varphi(v_1)(v_2) := \theta(v_1, v_2)$. Da θ im zweiten Argument linear ist, ist $\varphi(v_1) \in \mathcal{L}(V|W)$ für jedes $v_1 \in V$. Da θ auch im ersten Argument linear ist, ist die Abbildung $v_1 \mapsto \varphi(v_1)$ ebenfalls linear, also $\varphi \in \mathcal{L}(V|\mathcal{L}(V|W))$. Es gilt nach Definition von φ

$$\Phi(\varphi)(v_1, v_2) = \varphi(v_1)(v_2) = \theta(v_1, v_2)$$

für alle $v_1, v_2 \in V$ und somit ist $\Phi(\varphi) = \theta$. Da θ beliebig war, ist damit die Surjektivität von Φ gezeigt. □

Analog zu obigem Resultat erhalten wir:

Korollar 2.44. *Seien V, W Vektorräume. Dann ist $\mathcal{L}^n(V|\mathcal{L}^m(V|W))$ isomorph zu $\mathcal{L}^{n+m}(V|W)$.*
Setze $W_0 = W$ und $W_{n+1} = \mathcal{L}(V|W_n)$. Dann ist W_n isomorph zu $\mathcal{L}^n(V|W)$ für $n \in \mathbb{N}$.

Definition 2.45. Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$. Die Abbildung f heißt $k+1$ mal differenzierbar falls $\mathrm{D}^k f : U \to \mathcal{L}^k(R^n|R^m)$ differenzierbar ist und $\mathrm{D}^{k+1}f(x) = \mathrm{D}\mathrm{D}^k(f)(x)$. Die Abbildung heißt beliebig differenzierbar falls sie k mal differenzierbar ist für jedes $k \in \mathbb{N}$. Die Abbildung heißt k mal stetig differenzierbar falls $\mathrm{D}^k f$ stetig ist. Mit $\mathbf{C}^k(U, \mathbb{R}^n)$ bezeichnen wir den Vektorraum der k mal stetig differenzierbaren Abbildungen von U nach $\mathbb{R}^n$.

Bemerkung 2.46. (i) In obiger Definition haben wir die Isomorphismen aus Korollar 2.44 benutzt und wir werden sie im weiteren Verlauf oft nutzen ohne jedes mal explizit darauf hinzuweisen.

(ii) Aus Satz 2.33 erhalten wir induktiv, dass eine Abbildung f k mal stetig differenzierbar ist, genau dann wenn alle partiellen Ableitungen vom Grad k existieren und stetig sind.

Lemma 2.47. *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}^m$ mindestens $k+1$ mal differenzierbar und seien $v_1, \dots, v_k \in \mathbb{R}^n$. Dann ist die Abbildung*

$$\begin{aligned} g : U &\to \mathbb{R}^m \\ x &\mapsto \mathrm{D}^k f(x)(v_1, \dots, v_k) \end{aligned}$$

differenzierbar und es gilt $\mathrm{D}g(x) = \mathrm{D}^{k+1} f(x)(v_1, \dots, v_k, \cdot)$.

Beweis. Definiere die Abbildung

$$\begin{aligned} I : \mathcal{L}^k(\mathbb{R}^n | \mathbb{R}^m) &\to \mathbb{R}^m \\ A &\mapsto A(v_1, \dots, v_k). \end{aligned}$$

Dann gilt $g = I \circ \mathrm{D}^k f$. Die Abbildung $\mathrm{D}^k f$ ist nach Voraussetzung differenzierbar, die Abbildung I ist linear und damit insbesondere differenzierbar und es gilt $\mathrm{D}I(A) = I$ für alle $A \in \mathcal{L}^k(\mathbb{R}^n | \mathbb{R}^m)$ (Beispiel 2.28). Dann ist nach der Kettenregel auch g differenzierbar und es gilt

$$\mathrm{D}g(x)(v) = (\mathrm{D}I)(\mathrm{D}^k f(x))\mathrm{D}\mathrm{D}^k f(x)(v) = I\mathrm{D}^{k+1} f(x)(v) = \mathrm{D}^{k+1} f(x)(v_1, \dots, v_n, v). \quad \square$$

Satz 2.48. *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ mindestens k mal differenzierbar und $f = (f_1, \dots, f_m)$. Sei $e_1, \dots, e_n$ die Standardbasis von $\mathbb{R}^n$. Dann gilt für die Komponenten von $\mathrm{D}^k f(x)$*

$$\left(\mathrm{D}^k f(x)\right)_{ji_1 \dots i_k} = \partial_{i_k} \cdots \partial_{i_1} f_j(x)$$

für alle $j \in \{1, \dots, m\}$ und $i_1, \dots, i_k \in \{1, \dots, n\}$.

Beweis. Für $k = 1$ ist die Aussage bereits bekannt. Wir führen nun eine Induktion über k aus. Angenommen wir wissen bereits, dass die Aussage für $1 \leq l < k$ stimmt, das heißt es gilt

$$\left(\mathrm{D}^l f(x)\right)(e_{i_1}, \cdots, e_{i_l}) = \sum_{j=1}^m (\mathrm{D}^l f(x))_{ji_1 \cdots i_k} = \sum_{j=1}^m \partial_{i_l} \cdots \partial_{i_1} f_j(x).$$

Dann gilt nach dem vorhergehenden Lemma

$$\begin{aligned} \mathrm{D}^{l+1} f(x)(e_{i_1}, \cdots, e_{i_l}, e_{i_{l+1}}) &= \mathrm{D}\left(\sum_{j=1}^m \partial_{i_l} \cdots \partial_{i_1} f_j(x)\right)(e_{i_{l+1}}) \\ &= \sum_{j=1}^m \mathrm{D}\left(\partial_{i_l} \cdots \partial_{i_1} f_j\right)(x)(e_{i_{l+1}}) \\ &= \sum_{j=1}^m \partial_{i_{l+1}} \partial_{i_l} \cdots \partial_{i_1} f_j(x). \end{aligned} \quad \square$$

Satz 2.49 (Satz von Schwarz). *Sei $U \subset \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ zweimal stetig differenzierbar, dann gilt für alle $x \in U$*

$$\mathrm{D}^2 f(x)(u, v) = \mathrm{D}^2 f(x)(v, u).$$

Insbesondere gilt für die partiellen Ableitungen

$$\partial_{i_1} \partial_{i_2} f(x) = \partial_{i_2} \partial_{i_1} f(x)$$

für alle $i_1, i_2 \in \{1, \ldots, n\}$.

Beweis. Es genügt den Fall $m = 1$ zu betrachten, da alle Ableitungen komponentenweise berechnet werden können.

Sei $x \in U$ fest und $\epsilon > 0$ so, dass $B_\epsilon(x) \subset U$. Seien $u, v \in \mathbb{R}^n$ und $0 < t < c = \frac{\epsilon}{2 \max\{\|u\|, \|v\|\}}$. Wir wenden den (eindimensionalen) Mittelwertsatz auf die Hilfsfunktion

$$\begin{aligned} g_1 : [0, t] &\to \mathbb{R} \\ s &\mapsto f(x + tv + su) - f(x + su) \end{aligned}$$

an und finden $\xi \in (0, t)$ so, dass

$$g_1(t) - g_1(0) = g_1'(\xi) t = (\mathrm{D} f(x + tv + \xi u)(u) - \mathrm{D} f(x + \xi u)(u)) t.$$

Dabei haben wir für die letzte Umformung die Kettenregel verwendet. Wir wenden nun erneut den Mittelwertsatz an, diesmal auf die Hilfsfunktion

$$\begin{aligned} g_2 : [0, t] &\to \mathbb{R} \\ s &\mapsto \mathrm{D} f(x + sv + \xi u)(u) \end{aligned}$$

und erhalten $\eta \in (0, t)$ so, dass

$$g_2(t) - g_2(0) = g_2'(\eta) t = \mathrm{D}^2 f(x + \eta v + \xi u)(u, v) t.$$

Für die letzte Umformung haben wir wiederum die Kettenregel sowie Lemma 2.47 benutzt.

Zusammen haben wir für jedes $t \in (0, c)$ Zahlen $\xi, \eta \in (0, t)$ gefunden, so dass

$$\begin{aligned} &f(x + tv + tu) - f(x + tv) - f(x + tu) + f(x) \\ &\qquad = g_1(t) - g_1(0) = (\mathrm{D} f(x + tv + \xi u)(u) - \mathrm{D} f(x + \xi u)(u)) t \\ &\qquad = (g_2(t) - g_2(0)) t = \mathrm{D}^2 f(x + \eta v + \xi u)(u, v) t^2. \end{aligned}$$

Da die zweite Ableitung als stetig vorausgesetzt wurde und ξ und η für $t \to 0$ gegen 0 konvergieren, erhalten wir

$$\lim_{t \to 0^+} \frac{f(x + tv + tu) - f(x + tv) - f(x + tu) + f(x)}{t^2} = \mathrm{D}^2 f(x)(u, v). \tag{2.6}$$

Diese Gleichung gilt für jedes $u, v \in \mathbb{R}^n$. Die linke Seite der Gleichung ist symmetrisch unter Vertauschung von u und v, also muss die rechte Seite es auch sein und es gilt $\mathrm{D}^2 f(x)(u, v) = \mathrm{D}^2 f(x)(v, u)$.

Setzen wir für u und v Elemente der Standardbasis ein, also $u = e_{n_1}$ und $v = e_{n_2}$, so folgt aus Satz 2.48 die Vertauschbarkeit der partiellen Ableitungen. □

Bemerkung 2.50. (i) Mittels Induktion folgt, dass $\mathrm{D}^k f(x)(u_1, \ldots, u_k)$ seinen Wert unter beliebigen Vertauschungen der u_i nicht ändert vorausgesetzt $\mathrm{D}^k f$ ist stetig. Man sagt, die Multilinearform $\mathrm{D}^k f(x)$ ist symmetrisch. Die Reihenfolge der Anwendung von höchstens k partiellen Ableitungen ist also unerheblich falls $f \in \mathbf{C}^k(\mathbb{R}^n, \mathbb{R}^m)$ ist.

(ii) Ohne Stetigkeit von $\mathrm{D}^2 f$ können partielle Ableitungen von der Reihenfolge abhängen.

(iii) Gleichung (2.6) stellt die zweite Ableitung mittels eines einfachen Grenzwertes dar. Eine solche Darstellung kann zum Beispiel für die Diskretisierung von Differenzialgleichungen hilfreich sein. Es wurde jedoch zweimal stetig differenzierbares f vorausgesetzt. Es gibt Funktionen, für die der Grenzwert auf der linken Seite existiert, die aber nicht zweimal differenzierbar sind.

Satz 2.51 (Taylorformel). *Sei $U \subset \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine $l+1$ mal differenzierbare Funktion und sei $h \in \mathbb{R}^n$, so dass $x+th \in U$ für jedes $t \in [0,1]$. Dann gilt für ein $\theta \in (0,1)$*

$$f(x+h) = \sum_{k=0}^{l} \frac{1}{k!} \mathrm{D}^k f(x)(h, \ldots, h) + \frac{1}{(l+1)!} \mathrm{D}^{l+1} f(x+\theta h)(h, \ldots, h).$$

Beweisidee. Wende die eindimensionale Taylorformel mit Lagrange-Restglied auf die Hilfsfunktion

$$\begin{aligned} g : [0,1] &\to \mathbb{R} \\ t &\mapsto f(x+th) \end{aligned}$$

an. □

Bemerkung 2.52. (i) Stellen wir h bezüglich der Standardbasis dar $h = \sum_{i=1}^n h_i e_i$, dann können wir Koordinatendarstellungen der einzelnen Terme bestimmen wie wir hier exemplarisch für den Term zweiten Grades tun wollen:

$$\mathrm{D}^2 f(x)(h,h) = \sum_{i,j=1}^{n} h_i h_j \mathrm{D}^2 f(x)(e_i, e_j) = \sum_{i,j=1}^{n} h_i h_j \partial_i \partial_j f(x).$$

Analog zum eindimensionalen Fall approximiert die Taylorformel eine $l+1$ mal differenzierbare Funktion also durch ein Polynom (in n Variablen) dessen Koeffizienten partielle Ableitungen von f an der Stelle x sind.

(ii) Analog zu Korollar 2.19 ist es im Prinzip möglich lokale Extrema von hinreichend differenzierbaren Funktionen mittels der Ableitungen zu charakterisieren. Es gilt zum Beispiel: Falls $\mathrm{D}f(x) = 0, \cdots, \mathrm{D}^{l-1} f(x) = 0$ und $\mathrm{D}^l f(x) \neq 0$ so liegt in x

- ein Minimum vor, falls l gerade und $\mathrm{D}^l f(x)$ positiv definit ist, das heißt $\mathrm{D}^l f(x)(h, \ldots, h) > 0$ für alle $h \in \mathbb{R}^n \setminus \{0\}$,

- ein Maximum vor, falls l gerade und $\mathrm{D}^l f(x)$ negativ definit ist,
- kein Extremum vor, falls l ungerade oder $\mathrm{D}^l f(x)$ indefinit ist, das heißt es existieren $h_1, h_2 \in \mathbb{R}^n$, so dass $\mathrm{D}^l(h_1, \ldots, h_1) > 0$ und $\mathrm{D}^l(h_2, \ldots, h_2) < 0$.

Im Allgemeinen ist diese Charakterisierung jedoch weniger hilfreich als im eindimensionalen Fall. Zum Einen ist es insbesondere für $l > 2$ relativ kompliziert zu bestimmen ob eine multilineare Abbildung positiv definit ist. Zum Anderen können die Ableitungen auch (positiv) semidefinit sein, also $\mathrm{D}^l(h, \ldots, h) \geq 0$ für alle $h \in \mathbb{R}^n \setminus \{0\}$ aber $\mathrm{D}^l(h, \ldots, h) = 0$ für bestimmte h.

2.4 Funktionenfolgen und Differenzierbarkeit

Beispiel 2.53. Betrachte die Funktionenfolge (f_n) gegeben durch

$$\begin{aligned} f_n &: \mathbb{R} \to \mathbb{R} \\ x &\mapsto \frac{1}{n} e^{\mathrm{i}nx}. \end{aligned}$$

Es gilt

$$\|f_n\|_\infty = \frac{1}{n}$$

also konvergiert die Folge (f_n) gleichmäßig gegen 0.

Die Folge der Ableitungen

$$f_n'(x) = \mathrm{i}e^{\mathrm{i}nx}$$

konvergiert jedoch nur für $x = 0$.

Das obige Beispiel suggeriert, dass Bedingungen an die Konvergenz der Funktionenfolge nicht ausreichen um gutes Verhalten bezüglich der Ableitungen zu erreichen. Das liegt daran, dass selbst Funktionen mit beliebig kleinen Normen sehr stark oszillieren können. Wir müssen also zusätzlich Bedinungen an die Konvergenz der Folge der Ableitungen stellen.

Bemerkung 2.54. Wir haben in Abschnitt 1.6 die Supremumnorm für zahlwertige Funktionen eingeführt und damit die gleichmäßige Konvergenz definiert. Falls $f : X \to V$ eine Abbildung von einem metrischen Raum in einen normierten Raum $(V, \|\cdot\|)$ ist, dann können wir analog definieren

$$\|f\|_\infty = \sup \{ \|f(x)\| \mid x \in X \}.$$

Auch die Räume der beschränkten $\mathbf{B}(X, V)$ beziehungsweise stetigen beschränkten $\mathbf{C}_b(X, V)$ Funktionen können damit analog definiert werden. Gleichmäßige Konvergenz einer Folge solcher Funktionen ist dann die Konvergenz bezüglich dieser Norm.

Satz 1.74 und Satz 1.77 gelten dann entsprechend, vorausgesetzt der Raum V ist vollständig. Dazu muss in den entsprechenden Beweisen nur jeweils der Betrag durch die Norm $\|\cdot\|$ ersetzt werden.

Satz 2.55. *Sei $U \subset \mathbb{R}^n$ und $f_n : U \to \mathbb{R}^m$ stetig differenzierbar für jedes $n \in \mathbb{N}$. Falls $f_n \to f$ und $\mathrm{D}f_n \to g$ gleichmäßig konvergieren, dann ist f differenzierbar und es gilt $\mathrm{D}f = g$.*

Beweis. Es genügt wiederum den Fall $m = 1$ zu betrachten, da wir im allgemeinen Fall den Satz komponentenweise anwenden können. Da die f_n stetig differenzierbar sind und $\mathrm{D}f_n$ gleichmäßig gegen g konvergieren, folgt aus Satz 1.77 die Stetigkeit von g.

Sei nun $x \in U$ fest und $\epsilon > 0$. Dann gibt es $\delta > 0$ so, dass $\|g(y) - g(x)\| < \frac{\epsilon}{3}$ falls $\|y - x\| < \delta$. Außerdem existiert $N \in \mathbb{N}$, so dass für alle $n > N$ gilt $\|\mathrm{D}f_n - g\|_\infty \leq \frac{\epsilon}{3}$.

Sei $h \in B_\delta(x)$. Dann existiert nach dem Mittelwertsatz für jedes n ein $\xi_n \in \{x + th | \, t \in (0,1)\}$, so dass $f_n(x+h) - f_n(x) = \mathrm{D}f_n(\xi_n)(h)$. Dann gilt für $n > N$

$$\begin{aligned}
\frac{1}{\|h\|} \left| f_n(x+h) - f_n(x) - \mathrm{D}f_n(x)(h) \right| &= \frac{1}{\|h\|} \left| \left(\mathrm{D}f_n(\xi_n) - \mathrm{D}f_n(x)\right)(h) \right| \\
&\leq \|\mathrm{D}f_n(\xi_n) - \mathrm{D}f_n(x)\| \\
&\leq \|\mathrm{D}f_n(\xi_n) - g(\xi_n)\| + \|g(\xi_n) - g(x)\| + \|g(x) - \mathrm{D}f_n(x)\| \leq \epsilon.
\end{aligned}$$

Für $n \to \infty$ erhalten wir

$$\frac{1}{\|h\|} \left| f(x+h) - f(x) - g(x)(h) \right| \leq \epsilon.$$

Da ϵ beliebig positiv gewählt war, konvergiert die linke Seite für $h \to 0$ gegen 0. Damit ist f in x (total) differenzierbar mit Ableitung $\mathrm{D}f(x) = g(x)$. □

Bemerkung 2.56. Wir können nun auf $\mathbf{C}^1_b(\mathbb{R}^n, \mathbb{R}^m)$, dem Raum der differenzierbaren, beschränkten Funktionen mit beschränkten Ableitungen, eine Norm wie folgt definieren:

$$\|f\|_{\mathbf{C}^1} := \|f\|_\infty + \|Df\|_\infty . \tag{2.7}$$

Konvergenz bezüglich dieser Norm bedeutet dann genau gleichmäßige Konvergenz und gleichmäßige Konvergenz der Ableitungen. Der obige Satz zeigt dann, dass $\mathbf{C}^1(\mathbb{R}^n, \mathbb{R}^m)$ vollständig ist.

Wir formulieren die folgenden Resultate über Potenzreihen lediglich für Entwicklungen um 0, sie gelten aber sinngemäß für beliebige Entwicklungspunkte.

Satz 2.57. *Sei $f(x) = \sum_{n=0}^\infty a_n x^n$ eine Potenzreihe mit positivem Konvergenzradius r. Dann ist f auf $B_r(0)$ differenzierbar und es gilt $f'(x) = \sum_{n=0}^\infty n a_n x^{n-1}$.*

Beweis. Wir untersuchen zunächst den Konvergenzradius der formal abgeleiteten Reihe $\sum_{n=0}^\infty n a_n x^{n-1}$:

$$\limsup_{n\to\infty} \sqrt[n]{(n+1)a_{n+1}} = \lim_{n\to\infty} \sqrt[n]{n+1} \limsup_{n\to\infty} \sqrt[n]{a_{n+1}} = \frac{1}{r}.$$

Diese Reihe hat also den selben Konvergenzradius wie die ursprüngliche. Setze nun für $x \in B_r(0)$

$$g(x) := \sum_{n=0}^\infty n a_n x^{n-1}.$$

Sei nun $x \in B_r(0)$ fest und wähle $a > 0$, so dass $|x| < a < r$. Nach Satz 1.82 konvergieren die Partialsummenfolgen

$$f_N := \sum_{n=0}^{N} a_n x^n$$
$$g_N := \sum_{n=0}^{N} n a_n x^{n-1}$$

auf $B_a(0)$ (sogar auf $K_a(0)$) gleichmäßig gegen f beziehungsweise g. Offensichtlich gilt $f_N' = g_N$. Damit folgt aus Satz 2.55, dass die Grenzfunktion f differenzierbar auf $B_a(0)$ ist und $f' = g$.

Da Differenzierbarkeit eine lokale Eigenschaft ist und x innerer Punkt von $B_a(0)$ ist gilt $f'(x) = g(x)$. Da $x \in B_r(0)$ beliebig war, gilt $f' = g$ auf $B_r(0)$. □

Korollar 2.58. *Sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe mit positivem Konvergenzradius. Dann ist f beliebig differenzierbar, und es gilt $a_k = \frac{f^{(k)}(0)}{k!}$.*

Beweis. Die beliebige Differenzierbarkeit folgt induktiv aus dem vorhergehenden Satz. Ebenfalls induktiv folgt, dass wir auch die höheren Ableitungen gliedweise berechnen dürfen. Es gilt also

$$f^{(k)}(x) = \sum_{n=0}^{\infty} n(n-1)\cdots(n-k) a_n x^{n-k}.$$

Für $x = 0$ folgt $f^{(k)}(0) = k! a_k$. □

Beispiel 2.59. Mit dieser Vorbereitung können wir jetzt die Ableitung der Exponentialfunktion berechnen:

$$(e^x)' = \sum_{n=0}^{\infty} \left(\frac{x^n}{n!}\right) = \sum_{n=0}^{\infty} \frac{n x^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = e^x.$$

Wir erhalten damit auch die Ableitungen von Sinus und Cosinus. Da beide Funktionen durch auf ganz $\mathbb{R}$ konvergente Potenzreihen gegeben sind, wissen wir bereits dass sie beliebig differenzierbar sind.

$$\begin{aligned}(e^{\mathrm{i}x})' &= (\cos(x) + \mathrm{i}\sin(x))' = (\cos(x))' + \mathrm{i}(\sin(x))' \\ &= \mathrm{i}e^{\mathrm{i}x} = \mathrm{i}\cos(x) - \sin(x).\end{aligned}$$

Vergleich von Real- und Imaginärteil ergibt

$$(\cos(x))' = -\sin(x) \qquad (\sin(x))' = \cos(x).$$

Definition 2.60. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und seien $a_\eta \in \mathbb{K}$ Koeffizienten für jedes $\eta \in \mathbb{N}_0^d$. Dann ist

$$\sum_{\eta \in \mathbb{N}_0^d} a_\eta x^\eta$$

eine (formale) Potenzreihe in d Variablen (wobei wir Multiindexnotation verwendet haben).

Eine Funktion $f : U \to \mathbb{K}$ für eine offene Nullumgebung $U \subset \mathbb{R}^d$ heißt analytisch in 0, wenn sie in eine Potenzreihe um 0 entwickelt werden kann, das heißt es existieren $\epsilon > 0$ und $a_\eta \in \mathbb{K}$ für jedes $\eta \in \mathbb{N}_0^d$, so dass

$$f(x) = \sum_{\eta \in \mathbb{N}_0^d} a_\eta x^\eta$$

für alle $x \in B_\epsilon(0)$.

Bemerkung 2.61. (i) Die Konvergenz der auftretenden Reihen kann zum Beispiel wie folgt definiert werden: wir nennen $\sum_{\eta \in \mathbb{N}_0^d} a_\eta x^\eta$ konvergent gegen S, falls für jedes $\epsilon > 0$ eine endliche Teilmenge $A \subset \mathbb{N}_0^d$ existiert, so dass für jede endliche Teilmenge $A \subset B \subset N_0^d$ gilt

$$\left| \sum_{\eta \in B} a_\eta x^\eta - S \right| < \epsilon.$$

(ii) Der Konvergenzbereich einer Potenzreihe in mehr als einer Variablen ist nicht mehr so einfach zu charakterisieren wie im eindimensionalen Fall (siehe Beispiel 2.62). Es gilt aber analog zum eindimensionalen Fall: falls $y = (y_1, \cdots, y_d) \in \mathbb{R}^d$ ein Punkt ist mit $y_1 \neq 0, \cdots, y_d \neq 0$, so dass

$$\sum_{\eta \in \mathbb{N}_0^d} a_\eta y^\eta$$

konvergiert, dann konvergiert die Reihe für jedes x im offenen Quader

$$\left\{ x \in \mathbb{R}^d \middle| \ |x_i| < |y_i| \ \text{ für alle } i \in \{1, \ldots, d\} \right\}$$

absolut, das heißt es gilt

$$\sum_{\eta \in \mathbb{N}_0^d} |a_\eta| \left(|x_1|, \cdots, |x_2| \right)^\eta < \infty.$$

(iii) Konvergiert eine Potenzreihe auf einer offenen Umgebung U von 0, so ist sie dort beliebig differenzierbar und es gilt

$$a_\eta = \frac{\partial^\eta f(0)}{\eta!}.$$

Dabei setzen wir

$$\partial^\eta f = \partial_1^{\eta_1} \cdots \partial_d^{\eta_d} f \text{ und } \eta! := \eta_1! \cdots \eta_d!.$$

(iv) Aus der vorhergehenden Bemerkung folgt insbesondere, dass die Koeffizienten der Potenzreihendarstellung einer Funktion (falls diese existiert) eindeutig bestimmt sind (vergleiche dazu auch Übungsaufgabe 3 vom Blatt 6). Die obige Gleichung ist selten nützlich um die Potenzreihendarstellung einer Funktion zu finden (falls diese existiert). Umgekehrt kann sie jedoch hilfreich sein, falls man viele (partielle) Ableitungen einer Funktion gleichzeitig berechnen will um zum Beispiel Extrema zu untersuchen oder die Taylorformel aufzustellen. Wir kennen bereits einige Reihendarstellungen, z.B. für $e^x, \sin(x), \cos(x), \frac{1}{1-x}$ sowie alle Polynome. In vielen Fällen kann man nun Reihendarstellungen für Funktionen finden, indem man sie mittels geschickter Umformungen als Summen oder Produkte (manchmal auch Verkettungen) dieser Funktionen darstellt:

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

(v) Es gilt

$$\mathbf{C}(U) \supset \mathbf{C}^1(U) \supset \mathbf{C}^2(U) \supset \cdots \supset \mathbf{C}^k(U) \supset \cdots \supset \mathbf{C}^\infty(U) \supset \mathbf{C}^\omega(U).$$

Dabei bezeichne $\mathbf{C}^\infty(U)$ den Raum der beliebig differenzierbaren Funktionen und $\mathbf{C}^\omega(U)$ den Raum der analytischen Funktionen. Alle Inklusionen in obiger Kette sind echt.

(vi) Die analytischen Funktionen sind abgeschlossen unter Summen, Produkten und Verkettungen. Eine Potenzreihe ist auf ihrem Konvergenzkreis analytisch. Die Beweise für diese Aussagen sind nicht trivial.

Beispiel 2.62. Die Potenzreihe in zwei reellen Variablen

$$\sum_{n=0}^{\infty} (xy)^n$$

konvergiert auf $\{(x, y) | \ |xy| < 1\}$ gegen $\frac{1}{1-xy}$. Diese Funktion ist also analytisch in 0.

2.5 Der Banachsche Fixpunktsatz und der Satz über implizite Funktionen

Theorem 2.63 (Banachscher Fixpunktsatz). *Sei (X, d) ein vollständiger metrischer Raum und $\varphi : X \to X$ eine streng kontraktive Abbildung, das heißt es gibt $c \in [0, 1)$, so dass*

$$d(\varphi(x), \varphi(y)) \leq cd(x, y)$$

für alle $x, y \in X$ gilt. Dann hat φ genau einen Fixpunkt, das heißt es existiert genau ein $x \in X$, so dass $\varphi(x) = x$.

Beweis. Zunächst stellen wir fest, dass φ insbesondere Lipschitz-stetig, also stetig ist. Sei $x_0 \in X$ und definiere rekursiv die Folge $x_n := \varphi(x_{n-1})$ für $n \in \mathbb{N}$. Dann gilt auf Grund der strengen Kontraktivität

$$d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq cd(x_n, x_{n-1})$$

und mit Induktion folgt

$$d(x_{n+k}, x_{n+k-1}) \leq c^k d(x_n, x_{n-1})$$

für alle $k, n \in \mathbb{N}$. Insbesondere folgt

$$d(x_n, x_{n-1}) \leq c^{n-1} d(x_1, x_0)$$

Aus der Dreiecksungleichung folgt dann

$$\begin{aligned} d(x_{n+k}, x_{n-1}) &\leq d(x_{n+k}, x_{n+k-1}) + d(x_{n+k-1}, x_{n+k-2}) + \cdots + d(x_n, x_{n-1}) \\ &\leq \left(c^k + c^{k-1} + \cdots + c + 1\right) d(x_n, x_{n-1}) \\ &= \frac{1 - c^{k+1}}{1 - c} d(x_n, x_{n-1}) \leq \frac{1 - c^{k+1}}{1 - c} c^{n-1} d(x_1, x_0) \\ &\leq \frac{c^{n-1}}{1 - c} d(x_1, x_0). \end{aligned}$$

Da der letzte Term unabhängig von k ist und für $n \to \infty$ gegen 0 konvergiert, haben wir damit die Cauchy-Eigenschaft der Folge (x_n) gezeigt. Nach Voraussetzung ist X vollständig, also konvergiert (x_n) gegen ein $x \in X$. Dann gilt

$$x = \lim_{n \to \infty} x_n = \lim_{n \to \infty} \varphi(x_{n-1}) = \varphi\left(\lim_{n \to \infty} x_{n-1}\right) = \varphi(x)$$

wobei wir im vorletzten Schritt die Stetigkeit von φ verwendet haben.

Es bleibt noch die Eindeutigkeit des Fixpunktes zu zeigen. Seien $x, y \in X$ Fixpunkte von φ. Dann gilt

$$d(x, y) = d(\varphi(x), \varphi(y)) \leq cd(x, y).$$

Da $c < 1$ folgt daraus $d(x, y) = 0$ und aus der Definitheit der Metrik $x = y$. □

Bemerkung 2.64. (i) Eine Folgerung aus dem Banachschen Fixpunktsatz: läßt man in diesem Hörsaal einen Stadtplan von Leipzig zufällig auf den Boden fallen, dann gibt es genau einen Punkt auf dem Stadtplan, der genau über dem entsprechenden Punkt der Stadt liegt.

(ii) Aus dem Beweis ist auch ein Verfahren zum Bestimmen des Fixpunktes ersichtlich: man wendet φ wiederholt auf einen beliebigen Startpunkt an. Läßt man in einer der obigen Abschätzungen $k \to \infty$ gehen, so erhält man

$$d(x, x_{n-1}) \leq \frac{1}{1-c} d(x_n, x_{n-1}),$$

also eine Abschätzung für den Fehler.

(iii) Viele Probleme lassen sich in Fixpunktprobleme überführen. Die Schwierigkeit liegt dann häufig darin, einen passenden metrischen Raum zu finden, so dass man den Satz anwenden kann.

Beispiel 2.65. Betrachte die Gleichung $x - y^2 = 0$. Für $x \in (0, \infty)$ gibt es die beiden Lösungen

$$y = \sqrt{x} \text{ beziehungsweise } y = -\sqrt{x}.$$

Für jeden Lösungspunkt (ξ, η) der Gleichung gibt es also eine Umgebung U und eine Funktion f, so dass alle in U liegenden Lösungen der Gleichung durch $(x, f(x))$ gegeben sind. Darüber hinaus ist f sogar stetig differenzierbar.

Bemerkung 2.66. In vielen Fällen wird es nicht mehr möglich sein, Formeln für die Lösungen von Gleichungssystemen anzugeben. Der Satz über implizite Funktionen erlaubt es unter bestimmten Voraussetzungen zumindest die Existenz und bestimmte Eigenschaften solcher Lösungen nachzuweisen.

Sei $F : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}^n$ eine Abbildung. Betrachten wir $x_1, \ldots, x_p$ als unabhängige Variablen und $y_1, \ldots, y_q$ als abhängige Variablen und setzen $x = (x_1, \ldots, x_p)$, $y = (y_1, \ldots, y_q)$, dann ist durch

$$F(x, y) = 0$$

ein Gleichungssystem (mit n Gleichungen) gegeben.

Ist so ein F gegeben, so bezeichnen wir mit $D_y F(x, \eta)$ die Ableitung der Funktion $\mathbb{R}^q \ni y \mapsto F(x, y)$ an der Stelle η, also für festes x. Analog ist $D_x F(\xi, y)$ die Ableitung der Funktion $\mathbb{R}^p \ni x \mapsto F(x, y)$ an der Stelle ξ für festes y.

Theorem 2.67 (Satz über implizite Funktionen). *Seien $U \subset \mathbb{R}^p$ und $V \subset \mathbb{R}^q$ offen. Sei $F : U \times V \to \mathbb{R}^q$ stetig differenzierbar und es seien $\xi \in U$ und $\eta \in V$ so, dass $F(\xi, \eta) = 0$ und $\mathrm{D}_y F(\xi, \eta)$ invertierbar ist.*

Dann existieren Umgebungen $\tilde{U} \subset U$ von ξ und $\tilde{V} \subset V$ von η und eine stetige Funktion $f : \tilde{U} \to \tilde{V}$, so dass $f(\xi) = \eta$ und $F(x, f(x)) = 0$ für alle $x \in \tilde{U}$.

Beweis. Wir setzen $D = \mathrm{D}_y F(\xi, \eta)$. Wir betrachten das Fixpunktproblem (auf Funktionen g)

$$g(x) - D^{-1}F(x, g(x)) = g(x) \qquad \forall x \in \tilde{U}.$$

Eine Funktion f ist Lösung dieses Fixpunktproblems genau dann wenn $F(x, f(x)) = 0$ (da D nach Voraussetzung invertierbar ist). Wir suchen also Fixpunkte der Abbildung

$$(\varphi(g))(x) := g(x) - D^{-1}F(x, g(x)).$$

Auf dieses Fixpunktproblem wollen wir den Banachschen Fixpunktsatz anwenden und müssen nun noch einen geeigneten Definitionsbereich für φ suchen.

Bezeichne $I : \mathbb{R}^q \to \mathbb{R}^q$ die identische Abbildung. Die Funktion

$$(x, y) \mapsto \left\| I - D^{-1}\mathrm{D}_y F(x, y) \right\|$$

ist nach Voraussetzung stetig und verschwindet im Punkt (ξ, η). Es existieren also $\delta, \epsilon > 0$ so, dass $B_\delta(\xi) \subset U$, $B_\epsilon(\eta) \subset V$ und

$$\left\| I - D^{-1}\mathrm{D}_y F(x, y) \right\| \leq \frac{1}{2}$$

für alle $x \in B_\delta(\xi)$ und $y \in \tilde{V} := B_\epsilon(\eta)$. Auf Grund der Stetigkeit von $x \mapsto \left\| D^{-1}F(x, \eta) \right\|$ können wir ein (möglicherweise kleineres) $\delta > 0$ wählen, so dass

$$\left\| D^{-1}F(x, \eta) \right\| \leq \frac{\epsilon}{4}$$

für alle $x \in \tilde{U} = B_\delta(\xi)$.

Sei nun X die Menge all der Funktionen $g : \tilde{U} \to \mathbb{R}^q$ die

(i) g ist stetig,

(ii) $g(\xi) = \eta$ und

(iii) $\|g(x) - \eta\| \leq \frac{\epsilon}{2}$ für alle $x \in \tilde{U}$

erfüllen. Die dritte Bedingung impliziert insbesondere, dass $g(x) \in \tilde{V}$ für alle $g \in X$ und $x \in \tilde{U}$. Damit ist X Teilmenge des Banachraumes $\mathbf{B}(\tilde{U}, \mathbb{R}^q)$ mit der Norm

$$\|g\|_\infty = \sup \left\{ \|g(x)\| \,\middle|\, x \in \tilde{U} \right\}.$$

Darüber hinaus ist X in diesem Raum abgeschlossen, da die drei obigen Bedingungen unter gleichmäßigen Grenzwerten erfüllt bleiben (siehe Bemerkung 2.54) also nach Satz 1.38 vollständig.

Betrachte nun für festes $x \in \tilde{U}$ die Abbildung

$$\begin{aligned} \Phi : \tilde{V} &\to \mathbb{R}^q \\ y &\mapsto y - D^{-1}F(x, y). \end{aligned}$$

Dann folgt mit dem allgemeinen Mittelwertsatz Satz 2.38 für alle $y, z \in V$

$$\|\Phi(y) - \Phi(z)\| \leq \sup_{v \in \tilde{V}} \left\| I - D^{-1} \mathrm{D}_y F(x, y) \right\| \|y - z\| \leq \frac{1}{2} \|y - z\| .$$

Insbesondere gilt für $g_1, g_2 \in X$ und $x \in \tilde{U}$

$$\|(\varphi(g_1))(x) - (\varphi(g_2))(x)\| = \|\Phi(g_1(x)) - \Phi(g_2(x))\| \leq \frac{1}{2} \|g_1(x) - g_2(x)\|$$

und damit ist φ eine strenge Kontraktion

$$\|\varphi(g_1) - \varphi(g_2)\|_\infty \leq \frac{1}{2} \|g_1 - g_2\|_\infty .$$

Für $g \in X$ ist $\varphi(g)$ stetig und $(\varphi(g))(\xi) = \eta$. Schließlich gilt

$$\|(\varphi(g))(x) - \eta\| \leq \|(\varphi(g))(x) - (\varphi(\eta))(x)\| + \|(\varphi(\eta))(x) - \eta\| \leq \frac{1}{2} \|g(x) - \eta\| + \frac{\epsilon}{4} \leq \frac{\epsilon}{2}.$$

Damit ist gezeigt, dass φ die Menge X in sich selbst abbildet.

Jetzt sind alle Voraussetzungen des Banachschen Fixpunktsatzes erfüllt. Die Abbildung φ hat also genau einen Fixpunkt $f \in X$ und für dieses f gilt $F(x, f(x)) = 0$ für alle $x \in \tilde{U}$. □

Für die folgenden Aussagen behalten wir die Voraussetzungen sowie die Notation des Theorems und seines Beweises bei.

Bemerkung 2.68. Die Eindeutigkeitsaussage des Banachschen Fixpunktsatzes besagt, dass es für fest gewähltes $\tilde{U}$ und $\tilde{V}$ genau eine Funktion in X gibt, die $F(x, f(x))$ erfüllt. Es gilt jedoch etwas stärker: für festes $x \in \tilde{U}$ ist $f(x)$ die einzige Lösung des Gleichungssystems $F(x, y) = 0$. Für jede Lösung y gilt

$$\|y - f(x)\| = \|\Phi(y) - \Phi(f(x))\| \leq \frac{1}{2} \|y - f(x)\| ,$$

was nur durch $y = f(x)$ zu erfüllen ist.

Satz 2.69. *Es gibt eine (möglicherweise kleinere) Umgebung $\tilde{\tilde{U}}$ von ξ auf der f stetig differenzierbar ist. Die Ableitung ist durch*

$$\mathrm{D}f(x) = -\left(\mathrm{D}_y F(x, f(x))\right)^{-1} \mathrm{D}_x F(x, f(x))$$

gegeben.

Ohne Beweis.

Korollar 2.70 (Satz über die Umkehrfunktion). *Sei $U \subset \mathbb{R}^n$ und $f : U \to \mathbb{R}^n$ stetig differenzierbar. Falls $\mathrm{D}f(\xi)$ invertierbar ist für ein $\xi \in U$, dann gibt es eine Umgebung $\tilde{U}$ von ξ und eine Umgebung $\tilde{V}$ von $\eta = f(\xi)$, so dass f die Umgebung $\tilde{U}$ bijektiv auf $\tilde{V}$ abbildet und dass die Umkehrfunktion*

$$\begin{aligned} g : \tilde{V} &\to \tilde{U} \\ y &\mapsto f^{-1}(y) \end{aligned}$$

stetig differenzierbar ist. Es gilt

$$\mathrm{D}g(\eta) = (\mathrm{D}f(\xi))^{-1}.$$

Beweisidee. Wende den Satz über implizite Funktionen an, um das Gleichungssystem

$$F(x, y) = f(x) - y = 0$$

(lokal) nach x aufzulösen. □

Beispiel 2.71. Betrachte die Exponentialfunktion

$$\begin{aligned} f : \mathbb{R} &\to (0, \infty) \\ x &\mapsto e^x. \end{aligned}$$

Sie ist streng monoton wachsend (Warum eigentlich?), also injektiv und stetig, also auf Grund des Zwischenwertsatzes auch surjektiv. Es existiert also die Umkehrfunktion

$$\begin{aligned} g : (0, \infty) &\to \mathbb{R} \\ y &\mapsto \ln(y). \end{aligned}$$

Es gilt $f'(x) = f(x) > 0$ für jedes $x \in \mathbb{R}$. Dann sagt der Satz über die Umkehrfunktion, dass f lokal invertierbar ist (das wissen wir bereits, sie ist ja sogar global invertierbar), dass die Inverse g differenzierbar ist und für die Ableitung gilt für $\eta = e^\xi$:

$$g'(\eta) = \frac{1}{f'(\xi)} = \frac{1}{e^\xi} = \frac{1}{\eta}.$$