Mathematik II für inf, swt, msv

Vorlesung Sommersemester 2022

Prof. Meinolf Geck, Lehrstuhl für Algebra, Universität Stuttgart
https://pnp.mathematik.uni-stuttgart.de/idsr/idsr1/geckmf

Dies ist das Skript zur Vorlesung Mathematik II (für inf, swt, msv) im Sommersemester 2022 (V4Ü2, 14 Wochen). Es ist eine Fortsetzung des Skripts zur Vorlesung Mathematik I (für inf, swt, msv) im Wintersemester 2021/22; insbesondere führen wir die Nummerierung der Kapitel, Abschnitte etc. aus dem letzten Semester fort.

Hauptthema sind nun algebraische und analytische Strukturen über den Körpern $\mathbb{R}$ und $\mathbb{C}$. Nach einem ersten Kapitel zu Matrizen mit Einträgen in $\mathbb{R}$ und $\mathbb{C}$ geht es dann um Funktionen von einer oder mehreren reellen Variablen (und ansatzweise auch mit komplexen Variablen). Zentral ist dabei der Begriff der Grenzwertbildung, der auf der Vollständigkeit von $\mathbb{R}$ beruht (also mit den Begriffen "sup" und "inf" zu tun hat, die wir bereits im 1. Semester kurz kennengelernt haben). Insbesondere wird die klassische Differential- und Integralrechnung behandelt, die fundamental in der Mathematik selbst und für zahlreiche Anwendungen in allen möglichen Wissenschaften ist.

Wie bereits im vorherigen Semester werden nicht alle Beweise im Detail ausgeführt, sondern einerseits nur einige beispielhafte Argumentationen in den Anfängen eines jeden Kapitels, und andererseits solche Beweise mit einer algorithmischen Komponente, die sich dann auch effizient programmieren lässt.

Im Durchschnitt werden wiederum pro Woche etwa 7 Seiten dieses Skriptes behandelt.

Mein Dank geht wieder an Herrn Rainer Häußling für die regelmäßigen Listen mit Druckfehlern und Verbesserungsvorschlägen.

Kommentare sehr willkommen! (Insbesondere Druckfehler, sonstige Unklarheiten, Verbesserungsvorschläge etc.)

Stuttgart, Juli 2022

Literatur

Besonders geeignet für diese Vorlesung:

- G. Teschl und S. Teschl, Mathematik für Informatiker. Band 1: Diskrete Mathematik und Lineare Algebra, 4. Auflage, Springer Vieweg, 2013.
- G. Teschl und S. Teschl, Mathematik für Informatiker. Band 2: Analysis und Stochastik, 3. Auflage, Springer Vieweg, 2014.

Skripte aus früheren Durchgängen an der Uni Stuttgart:

- P. Lesky, Mathematik 2 für inf, swt, msv. Skript zur Vorlesung im WiSe 2018/19; siehe `http://info.mathematik.uni-stuttgart.de/Mathe2InfSS19`.
- M. Künzer, Mathematik 2 für inf, swt, msv, dsc. Skript zum SoSe 2021; siehe `https://info.mathematik.uni-stuttgart.de/Mathe-2-Inf-SoSe21`.

Zum Auffrischen von Schulwisssen und Grundlagen:

- T. Glosauser, (Hoch)Schulmathematik, Ein Sprungbrett vom Gymnasium zur Uni. Springer-Spektrum, 2015.
- M. Liebeck, A Concise Introduction to Pure Mathematics. Chapman Hall/CRC Mathematics Series, CRC Press, 3rd edition 2010.
- MINT Kolleg Baden-Württemberg, Mathematik-Vorkurs (Online), siehe `http://www.mint-kolleg.de/stuttgart/angebote/online_kurse` (folge dort auch den Links für den Mathematik-Vorkurs und dann zu `Mathematik-Online`).

Frei verfügbare mathematische Software zum Ausprobieren/Experimentieren:

- GAP - Groups, Algorithms, and Programming, siehe `http://www.gap-system.org/` (Exaktes Rechnen mit Zahlen und diskreten algebraischen Strukturen.)
- SageMath, siehe `https://www.sagemath.org/` (Basiert auf der Programmiersprache Python; siehe `https://www.python.org/`)

Einige weiterführende Texte (wird laufend ergänzt):

- M. Artin, Algebra. Aus dem Englischen übersetzt von Annette A'Campo. Birkhäuser Verlag, 1993.
- S. Axler, Linear Algebra done right. Undergraduate Texts in Mathematics, Springer-Verlag, 2015.
- N. L. Biggs, Discrete Mathematics, 2nd Edition. Oxford University Press, 2002.

- O. Forster, Analysis 1, 12. Auflage, Grundkurs Mathematik, Springer–Sektrum, 2016; e-Book frei verfügbar über https://doi.org/10.1007/978-3-658-11545-6.
- O. Forster, Analysis 2, 9. Auflage, Grundkurs Mathematik, Vieweg+Teubner, 2010; e-Book frei verfügbar über https://doi.org/10.1007/978-3-8348-8103-8.
- R. Haggarty, Principles of Mathematical Analysis, 2nd Edition. Prentice Hall, Addison–Wesley, 1993.
- J. R. Hass, C. E. Heil and M. D. Weir, Thomas' Calculus, 14th Edition, Pearson, 2017.
- B. Huppert und W. Willems, Lineare Algebra: Mit zahlreichen Anwendungen in Kryptographie, Codierungstheorie, Mathematischer Physik und Stochastischen Prozessen, Vieweg + Teubner Verlag, 2. Auflage 2010.
- M. Koecher, Lineare Algebra und analytische Geometrie (Neuauflage überarbeitet, aktualisiert und ergänzt), Grundwissen Mathematik, Springer–Verlag, 1985.
- E. Zeidler (Hrsg.), Springer-Handbuch der Mathematik I, II; Springer-Verlag, 2013; e-Book frei verfügbar über https://doi.org/10.1007/978-3-658-00285-5 bzw. https://doi.org/10.1007/978-3-658-00297-8.

Inhaltsverzeichnis

Kapitel V: Matrizen über $\mathbb{R}$ und $\mathbb{C}$

In diesem Kapitel führen wir das Thema Matrizen aus dem letzten Semester fort, betrachten hier aber ausschließlich Matrizen mit Einträgen in $\mathbb{R}$ oder $\mathbb{C}$. Wir können dann die besonderen Eigenschaften dieser Körper benutzen, also die Vollständigkeit von $\mathbb{R}$ (siehe Kapitel II, §10) sowie die algebraische Abgeschlossenheit von $\mathbb{C}$ (siehe Kapitel II, §11). Dadurch ergeben sich einige starke Aussagen über die Diagonalisierbarkeit von Matrizen, die sowohl rein mathematisch sehr bemerkenswert sind, als auch in Anwendungen eine große Rolle spielen.

23. ***Spektralzerlegung***

In §19 (Kapitel IV) haben wir das Standard-Skalarprodukt auf $\mathbb{R}^n$ betrachtet. Wir wollen dies nun auf den Vektorraum $\mathbb{C}^n$ ausdehnen, wobei die komplexe Konjugation ins Spiel zu bringen ist. Für $n = 1$, ist der Absolutbetrag von $z \in \mathbb{C}$ durch $|z| := \sqrt{z\bar{z}}$ definiert. Für $n \geqslant 1$ beliebig definieren wir nun eine Abbildung $\langle\ ,\ \rangle\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ durch

$$\left\langle \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle := x_1\bar{y}_1 + \ldots + x_n\bar{y}_n \qquad \text{für alle } x_i, y_j \in \mathbb{C}.$$

Sei $x \in \mathbb{C}^n$ mit Komponenten $x_1, \ldots, x_n \in \mathbb{C}$. Dann definieren wir $\bar{x} \in \mathbb{C}^n$ als den Spaltenvektor mit Komponenten $\bar{x}_1, \ldots, \bar{x}_n \in \mathbb{C}$. Sei $y \in \mathbb{C}^n$ mit Komponenten $y_1, \ldots, y_n \in \mathbb{C}$. Dann gilt die Regel:

$$\overline{\langle x, y\rangle} = \overline{x_1\bar{y}_1 + \ldots + x_n\bar{y}_n} = \bar{x}_1\bar{\bar{y}}_1 + \ldots + \bar{x}_n\bar{\bar{y}}_n = y_1\bar{x}_1 + \ldots + y_n\bar{x}_n = \langle y, x\rangle.$$

Insbesondere folgt $\langle x, x\rangle = \overline{\langle x, x\rangle} = |x_1|^2 + \ldots + |x_n|^2 \in \mathbb{R}$ für alle $x \in \mathbb{C}^n$. Das Ergebnis ist sogar $\geqslant 0$, also können wir $\|x\| := \sqrt{\langle x, x\rangle}$ als die Norm von x definieren; beachte: Auch hier gilt $\|x\| = 0$ nur für $x = 0_n$. Wir bezeichnen $\langle\ ,\ \rangle$ als die standard-***hermitesche Form*** auf $\mathbb{C}^n$. Analog zu §19 (Kapitel IV) können wir dann auch schreiben:

$$\langle x, y\rangle = x^{\mathrm{tr}} \cdot \bar{y} \qquad \text{für alle } x, y \in \mathbb{C}^n,$$

wobei das Ergebnis eine 1×1-Matrix der Form $[u]$ mit $u \in \mathbb{C}$ ist, wofür wir nach unseren Konventionen einfach nur $u \in \mathbb{C}$ schreiben.

Satz 23.1 (***Gram–Schmidt–Orthogonalisierung über*** $\mathbb{C}$). *Sei* $U \subseteq \mathbb{C}^n$ *ein Teilraum mit* $m := \dim U \geqslant 1$ *und* $\{v_1, \ldots, v_m\}$ *eine Basis von* U. *Dann gibt es eine Basis* $\{w_1, \ldots, w_m\}$ *von* U *mit folgenden Eigenschaften:*

(a) $d_i := \|w_i\|^2 > 0$ *und* $\langle w_i, w_j\rangle = 0$ *für alle* $1 \leqslant i, j \leqslant m$ *mit* $i \neq j$*;*

(b) $w_1 = v_1$ *und* $w_k = v_k - \sum_{j=1}^{k-1} d_j^{-1}\langle v_k, w_j\rangle.w_j$ *für* $k = 2, 3, \ldots, m$.

Beweis. Völlig analog zu Satz 19.3 in Kapitel IV. □

Definition 23.2. (a) Sei $A = [a_{ij}] \in M_n(\mathbb{C})$ beliebig. Dann setzen wir $\bar{A} := [\bar{a}_{ij}] \in M_n(\mathbb{C})$ und $A^* := \bar{A}^{tr}$. Man sieht sofort $(A^*)^* = A$ und $(A \cdot B)^* = B^* \cdot A^*$ für $B \in M_n(\mathbb{C})$.
Ist $A = A^*$ so heißt A eine ***hermitesche Matrix***. (Ist also $A \in M_n(\mathbb{R})$, so ist A hermitesch genau dann, wenn A symmetrisch ist.)
(b) Eine Matrix $U \in M_n(\mathbb{C})$ heißt ***unitäre Matrix***, wenn $U^* \cdot U = I_n$ gilt. In diesem Fall ist also U invertierbar, und die inverse Matrix ist auf besonders einfache Weise gegeben, nämlich durch $U^{-1} = U^*$. (Es folgt dann auch $U \cdot U^* = I_n$.)

Lemma 23.3. *Sei $U \in M_n(\mathbb{C})$ eine beliebige Matrix; seien $v_1, \ldots, v_n \in \mathbb{C}^n$ die Spalten von U. Genau dann ist U unitär, wenn $B = \{v_1, \ldots, v_n\}$ eine Orthonormalbasis von $\mathbb{C}^n$ bezüglich der standard-hermiteschen Form $\langle\ ,\ \rangle\colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$, $(x, y) \mapsto x^{tr} \cdot \bar{y}$, ist.*

Beweis. Für $1 \leqslant i, j \leqslant n$ gilt $\langle v_i, v_j\rangle = \langle U \cdot e_i, U \cdot e_j\rangle = (U \cdot e_i)^{tr} \cdot \overline{U \cdot e_j} = e_i^{tr} \cdot (U^{tr} \cdot \bar{U}) \cdot \bar{e}_j =$ Eintrag an der Stelle (i, j) von $U^{tr} \cdot \bar{U} = (U^* \cdot U)^{tr}$. Also ist $U^* \cdot U = I_n$ genau dann, wenn $\langle v_i, v_i\rangle = 1$ und $\langle v_i, v_j\rangle = 0$ für $1 \leqslant i, j \leqslant n$ mit $i \neq j$ gilt. □

Gegeben seien Spaltenvektoren $v_1, \ldots, v_m \in \mathbb{C}^n$ (mit $m \geqslant 1$). Das Tupel $(v_1, \ldots, v_m)$ heißt ***orthonormal***, wenn $\|v_i\| = 1$ und $\langle v_i, v_j\rangle = 0$ für $1 \leqslant i, j \leqslant m$ und $i \neq j$ gilt.

Lemma 23.4 (***Orthogonale Basisergänzung***). *Sei $1 \leqslant m \leqslant n$ und $(v_1, \ldots, v_m)$ ein orthonormales Tupel wie oben. Dann ist $(v_1, \ldots, v_m)$ l.u. und es gibt $v_{m+1}, \ldots, v_n \in \mathbb{C}^n$ so dass $B = \{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ eine Orthormalbasis von $\mathbb{C}^n$ ist. Sind $v_1, \ldots, v_m$ in $\mathbb{R}^n$, so können auch $v_{m+1}, \ldots, v_n$ in $\mathbb{R}^n$ gewählt werden.*

Beweis. Zuerst zeigen wir, dass $(v_1, \ldots, v_m)$ l.u. ist. Seien also $z_1, \ldots, z_m \in \mathbb{C}$ gegeben mit $z_1 v_1 + \ldots + z_m v_m = 0_n$. Sei $i \in \{1, \ldots, m\}$ fest. Wegen $\langle v_i, v_i\rangle = 1$ und $\langle v_i, v_j\rangle = 0$ für $i \neq j$ folgt dann $0 = \langle z_1 v_1 + \ldots + z_m v_m, v_i\rangle = z_i\langle v_i, v_i\rangle = z_i$, wie gewünscht. Nach dem Basisergänzungssatz gibt es Spaltenvektoren $v'_{m+1}, \ldots, v'_n \in \mathbb{C}^n$, so dass $\{v_1, \ldots, v_m, v'_{m+1}, \ldots, v'_n\}$ eine Basis von $\mathbb{C}^n$ ist. Wir wenden das obige Gram–Schmidt–Verfahren auf diese Basis an und erhalten eine Orthogonalbasis $\{w_1, \ldots, w_n\}$ von $\mathbb{C}^n$. Setze $d_i := \|w_i\|^2$ für alle i. Die Formeln in Satz 23.1 zeigen sofort, dass $w_i = v_i$ für $1 \leqslant i \leqslant m$ gilt. Dazu: Zunächst ist $w_1 = v_1$. Sei nun $m \geqslant 2$; wegen $\langle v_2, v_1\rangle = 0$ folgt $w_2 = v_2 - d_1^{-1}\langle v_2, w_1\rangle w_1 = v_2 - d_1^{-1}\langle v_2, v_1\rangle v_1 = v_2$. Sei nun $m \geqslant 3$; dann folgt analog

$$w_3 = v_3 - d_1^{-1}\langle v_3, w_1\rangle w_1 - d_2^{-1}\langle v_3, w_2\rangle w_2 = v_3 - d_1^{-1}\underbrace{\langle v_3, v_1\rangle}_{=0} v_1 - d_2^{-1}\underbrace{\langle v_3, v_2\rangle}_{=0} v_2 = v_3,$$

und so weiter bis $w_m = v_m$. Setzen wir schließlich $v_i := \|w_i\|^{-1} w_i$ für $i = m+1, m+2, \ldots, n$, so ist $\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ eine Orthonormalbasis von $\mathbb{C}^n$. Sind $v_1, \ldots, v_m$ in $\mathbb{R}^n$, so ergänzen wir zunächst $v_1, \dots, v_m$ zu einer Basis von $\mathbb{R}^n$ und wenden dann wieder das Gram–Schmidt–Verfahren in Satz 19.3 (Kapitel IV) an. $\square$

Beispiel 23.5. Für $\theta \in \mathbb{R}$, $0 \leqslant \theta < 2\pi$, und $z, w \in \mathbb{C}$ mit $|z|^2 + |w|^2 = 1$ setzen wir

$$U := \begin{bmatrix} w & z \\ -u(\theta)\bar{z} & u(\theta)\bar{w} \end{bmatrix} \in M_2(\mathbb{C}) \qquad \text{wobei} \qquad u(\theta) := \cos(\theta) + \sin(\theta)i \in \mathbb{C}.$$

Dann prüft man leicht nach, dass U unitär ist. Umgekehrt kann man zeigen, dass jede unitäre Matrix in $M_2(\mathbb{C})$ obige Form hat. (Dies ist eine sehr gute Übungsaufgabe.)

Eine Matrix $A \in M_n(\mathbb{C})$ heißt ***unitär diagonalisierbar***, wenn es eine unitäre Matrix $U \in M_n(\mathbb{C})$ gibt, so dass $U^{-1} \cdot A \cdot U = U^* \cdot A \cdot U$ eine Diagonalmatrix ist. Wir wollen herausfinden, unter welchen Bedingungen dies möglich ist.

Bemerkung 23.6. Sei $A \in M_n(\mathbb{C})$ unitär diagonalisierbar. Dann gilt $A \cdot A^* = A^* \cdot A$.

Denn sei $U \in M_n(\mathbb{C})$ unitär, so dass $D := U^* \cdot A \cdot U$ eine Diagonalmatrix ist. Dann ist $A = U \cdot D \cdot U^*$ und $A^* = (U \cdot D \cdot U^*)^* = (U^*)^* \cdot D^* \cdot U^* = U \cdot D^* \cdot U^*$. Es folgt $A \cdot A^* = (U \cdot D \cdot U^*) \cdot (U \cdot D^* \cdot U^*) = U \cdot (D \cdot D^*) \cdot U^*$ und analog $A^* \cdot A = U \cdot (D^* \cdot D) \cdot U^*$. Da D und D^* beide Diagonalmatrizen sind, gilt $D \cdot D^* = D^* \cdot D$ und damit auch $A \cdot A^* = A^* \cdot A$.

Ist $A \in M_n(\mathbb{C})$ und gilt $A \cdot A^* = A^* \cdot A$, so heißt A eine ***normale Matrix***. Zum Beispiel sind hermitesche Matrizen automatisch auch normale Matrizen.

Lemma 23.7. *Ist $A \in M_n(\mathbb{C})$ eine normale Matrix, so gilt $\|A \cdot x\| = \|A^* \cdot x\|$ für $x \in \mathbb{C}^n$.*

Beweis. Sei $x \in \mathbb{C}^n$. Dann ist $(A^* \cdot x)^{tr} = (\bar{A}^{tr} \cdot x)^{tr} = x^{tr} \cdot \bar{A}$ und damit

$$\|A^* \cdot x\|^2 = \langle A^* \cdot x, A^* \cdot x \rangle = (A^* \cdot x)^{tr} \cdot \overline{A^* \cdot x} = x^{tr} \cdot (\bar{A} \cdot \overline{A^*}) \cdot \bar{x} = x^{tr} \cdot (\overline{A \cdot A^*}) \cdot \bar{x};$$

andererseits gilt auch $\|A \cdot x\|^2 = \langle A \cdot x, A \cdot x \rangle = (A \cdot x)^{tr} \cdot \overline{A \cdot x} = x^{tr} \cdot (A^{tr} \cdot \bar{A}) \cdot \bar{x} = x^{tr} \cdot (\overline{A^* \cdot A}) \cdot \bar{x}$. Nach Voraussetzung sind die beiden rechten Seiten gleich. $\square$

Satz 23.8 (***Spektralzerlegung***). *Sei $A \in M_n(\mathbb{C})$ eine normale Matrix, also $A \cdot A^* = A^* \cdot A$. Dann ist A unitär diagonalisierbar; es gibt also eine unitäre Matrix $U \in M_n(\mathbb{C})$ so dass $D := U^{-1} \cdot A \cdot U = U^* \cdot A \cdot U \in M_n(\mathbb{C})$ eine Diagonalmatrix ist.*

Beweis. (Vollständige Induktion nach n.) Ist $n = 1$, so ist $A = [a]$ mit $a \in \mathbb{C}$ und die Aussage gilt mit $U = [1]$, $D = [a]$. Sei nun $n \geqslant 2$ und die Aussage bereits für alle normalen Matrizen in $M_{n-1}(\mathbb{C})$ gezeigt. Da $\mathbb{C}$ algebraisch abgeschlossen ist, hat das Minimalpolynom $\mu_A \in \mathbb{C}[X]$ eine Nullstelle, also gibt es einen Eigenwert $\lambda_1 \in \mathbb{C}$; sei $w_1 \in \mathbb{C}^n$ ein zugehöriger Eigenvektor, d.h., $w_1 \neq 0_n$ und $A \cdot w_1 = \lambda_1 w_1$. Setzen wir $v_1 := \|w_1\|^{-1} w_1 \in \mathbb{C}^n$, so ist auch v_1 ein Eigenvektor (mit Eigenwert λ_1) und $\|v_1\| = 1$. Nach Lemma 23.4 gibt es $v_2, \dots, v_n \in \mathbb{C}^n$ so dass $\{v_1, v_2, \dots, v_n\}$ eine Orthonormalbasis von $\mathbb{C}^n$ ist. Sei $U_1 \in M_n(\mathbb{C})$

die unitäre Matrix, deren Spalten durch $v_1, \ldots, v_n$ gegeben sind (siehe Lemma 23.3). Wir setzen $B := U_1^* \cdot A \cdot U_1 \in M_n(\mathbb{C})$. Wir zeigen nun, dass auch B normal ist. Dazu: Es gilt $B^* = (U_1^* \cdot A \cdot U_1)^* = U_1^* \cdot A^* \cdot (U_1^*)^* = U_1^* \cdot A^* \cdot U_1$. Mit $A \cdot A^* = A^* \cdot A$ erhalten wir $B \cdot B^* = (U_1^* \cdot A) \cdot (U_1 \cdot U_1^*) \cdot (A^* \cdot U_1) = U_1^* \cdot (A \cdot A^*) \cdot U_1 = U_1^* \cdot (A^* \cdot A) \cdot U_1 = (U_1^* \cdot A^*) \cdot (U_1 \cdot U_1^*) \cdot (A \cdot U_1) = B^* \cdot B$, wie behauptet.

Nun ist $U_1 \cdot e_1 = v_1$, also ist die erste Spalte von B gegeben durch $B \cdot e_1 = (U_1^* \cdot A) \cdot (U_1 \cdot e_1) = U_1^* \cdot (A \cdot v_1) = \lambda_1 (U_1^* \cdot v_1) = \lambda_1 (U_1^{-1} \cdot v_1) = \lambda_1 e_1$. Die Einträge in der ersten Spalte von B sind also $\lambda_1, 0, \ldots, 0$; insbesondere ist $\|B \cdot e_1\| = |\lambda_1|$. Seien $b_1, \ldots, b_n \in \mathbb{C}$ die Einträge in der ersten Zeile von B; insbesondere $b_1 = \lambda_1$. Wegen $B^* = \overline{B}^{\mathrm{tr}}$ sind dann $\bar{b}_1, \ldots, \bar{b}_n$ die Einträge in der ersten Spalte von B^*. Mit Lemma 23.7 folgt

$$|\lambda_1|^2 + |b_2|^2 + \ldots + |b_n|^2 = |b_1|^2 + |b_2|^2 + \ldots + |b_n|^2 = \|B^* \cdot e_1\|^2 = \|B \cdot e_1\|^2 = |\lambda_1|^2,$$

und damit $b_2 = \ldots = b_n = 0$. Also hat B eine Blockdiagonalgestalt wie folgt:

$$U_1^* \cdot A \cdot U_1 = B = \left[\begin{array}{c|c} \lambda_1 & 0_{n-1}^{\mathrm{tr}} \\ \hline 0_{n-1} & A_1 \end{array}\right] \qquad \text{mit} \qquad A_1 \in M_{n-1}(\mathbb{C}).$$

Da B normal ist, folgt mit einer leichten Rechnung, dass auch A_1 normal ist. Nach Induktion gibt es also eine unitäre Matrix $U_2 \in M_{n-1}(\mathbb{C})$, so dass $D_1 := U_2^* \cdot A_1 \cdot U_2 \in M_{n-1}(\mathbb{C})$ eine Diagonalmatrix ist. Setzen wir

$$U := U_1 \cdot U_2' \qquad \text{mit} \qquad U_2' := \left[\begin{array}{c|c} 1 & 0_{n-1}^{\mathrm{tr}} \\ \hline 0_{n-1} & U_2 \end{array}\right] \in M_n(\mathbb{C})$$

so sieht man sofort, dass auch U unitär ist; außerdem gilt

$$U^* \cdot A \cdot U = (U_2')^* \cdot \big(U_1^* \cdot A \cdot U_1\big) \cdot U_2' = (U_2')^* \cdot \left[\begin{array}{c|c} \lambda_1 & 0_{n-1}^{\mathrm{tr}} \\ \hline 0_{n-1} & A_1 \end{array}\right] \cdot U_2' = \left[\begin{array}{c|c} \lambda_1 & 0_{n-1}^{\mathrm{tr}} \\ \hline 0_{n-1} & U_2^* \cdot A_1 \cdot U_2 \end{array}\right],$$

und die rechte Seite ist eine Diagonalmatrix, deren Diagonaleinträge durch λ_1 und die Diagonaleinträge von D_1 gegeben sind. □

Ist $A \in M_n(\mathbb{C})$ normal und $D = U^* \cdot A \cdot U$ die Spektralzerlegung wie oben, so sind die Diagonaleinträge von D genau die Eigenwerte von A. Für hermitesche Matrizen ergibt sich:

Lemma 23.9. *Sei $A = [a_{ij}] \in M_n(\mathbb{C})$ beliebig.*
(a) *Ist $\lambda \in \mathbb{C}$ ein Eigenwert von A, so ist $\bar{\lambda} \in \mathbb{C}$ ein Eigenwert von $\bar{A}$.*
(b) *Ist A hermitesch (also $A = A^*$), so sind alle Eigenwerte von A reell.*

Beweis. (a) Sei $v \in \mathbb{C}^n$ ein zu λ gehöriger Eigenvektor, also $v \neq 0_n$ und $A \cdot v = \lambda v$. Seien $z_1, \ldots, z_n \in \mathbb{C}$ die Komponenten von v. Für $i \in \{1, \ldots, n\}$ ist dann die i-te Komponente von $A \cdot v$ gleich λz_i. Sei $\bar{v} \in \mathbb{C}^n$ der Spaltenvektor mit Komponenten $\bar{z}_1, \ldots, \bar{z}_n$. Die i-te

Komponente von $\bar{A} \cdot \bar{v}$ ist dann gegeben durch
$$\sum_{j=1}^{n} \bar{a}_{ij}\bar{z}_j = \sum_{j=1}^{n} \overline{a_{ij}z_j} = \overline{\sum_{1 \leqslant j \leqslant n} a_{ij}z_j} = \overline{\lambda z_i} = \bar{\lambda}\bar{z}_i;$$
also gilt $\bar{A} \cdot \bar{v} = \bar{\lambda}\bar{v}$ und damit ist $\bar{\lambda}$ ein Eigenwert von $\bar{A}$, mit Eigenvektor $\bar{v}$.

(b) Sei λ ein Eigenwert von A, mit Eigenvektor $0_n \neq v \in \mathbb{C}^n$. Nach (a) ist $\bar{\lambda}$ ein Eigenwert von $\bar{A}$, mit Eigenvektor $\bar{v} \in \mathbb{C}^n$. Wegen $\bar{A} = A^{\mathrm{tr}}$ folgt einerseits
$$v^{\mathrm{tr}} \cdot \bar{A} \cdot \bar{v} = (v^{\mathrm{tr}} \cdot A^{\mathrm{tr}}) \cdot \bar{v} = (A \cdot v)^{\mathrm{tr}} \cdot \bar{v} = (\lambda v^{\mathrm{tr}}) \cdot \bar{v} = \lambda(v^{\mathrm{tr}} \cdot \bar{v}) = \lambda\|v\|^2.$$
Andererseits ist die linke Seite auch gleich $v^{\mathrm{tr}} \cdot (\bar{A} \cdot \bar{v}) = v^{\mathrm{tr}} \cdot (\bar{\lambda}\bar{v}) = \bar{\lambda}(v^{\mathrm{tr}} \cdot \bar{v}) = \bar{\lambda}\|v\|^2$.
Also folgt $\bar{\lambda}\|v\|^2 = \lambda\|v\|^2$. Wegen $v \neq 0_n$ ist $\|v\| \neq 0$ und damit $\lambda = \bar{\lambda}$. □

Bemerkung 23.10. Sei $A \in M_n(\mathbb{C})$ eine normale Matrix. Um die Spektralzerlegung in Satz 23.8 zu berechnen, kann man rekursiv wie im obigen Beweis vorgehen. Alternativ geht dies auch wie folgt:

1. Schritt. Berechne das charakteristische Polynom $\chi_A \in \mathbb{C}[X]$ und finde die Nullstellen dieses Polynoms. Da $\mathbb{C}$ algebraisch abgeschlossen ist, gilt $\chi_A = (-1)^n(X-\lambda_1)^{n_1} \cdots (X-\lambda_r)^{n_r}$, wobei $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ paarweise verschieden sind und $n_i \geqslant 1$ für $1 \leqslant i \leqslant r$.
2. Schritt. Für $1 \leqslant i \leqslant r$ sei $E_i = N\big(A - \lambda_i I_n\big) \subseteq \mathbb{C}^n$ der Eigenraum zu λ_i. Da A diagonalisierbar ist, gilt $\dim E_i = n_i$. Bestimme eine Basis B_i von E_i; dazu muss man das homogene LGS mit Matrix $A - \lambda_i I_n \in M_n(\mathbb{C})$ lösen.
3. Schritt. Für $1 \leqslant i \leqslant r$ bestimme eine Orthonormalbasis C_i von E_i (zum Beispiel mit dem Gram–Schmidt–Verfahren). Dann ist $C := C_1 \cup \ldots \cup C_r$ eine Orthonormalbasis von $\mathbb{C}^n$.

Sei schließlich $U \in M_n(\mathbb{C})$ die Matrix, deren Spalten durch die Spaltenvektoren in C gegeben sind; dann ist U unitär und $D := U^* \cdot A \cdot U$ eine Diagonalmatrix.

Beispiel 23.11. Sei $A = A^* = \begin{bmatrix} 1 & i & i & 1 \\ -i & 1 & 1 & -i \\ -i & 1 & 3 & i \\ 1 & i & -i & 3 \end{bmatrix} \in M_4(\mathbb{C})$.

Wir erhalten $\chi_A = \det\big(A - XI_4\big) = \ldots = X^2(X-4)^2$; es gibt also die beiden Eigenwerte $\lambda_1 = 0$ und $\lambda_2 = 4$. Mit dem obigen Verfahren ergibt sich mit
$$U := \frac{1}{2\sqrt{3}} \begin{bmatrix} -i\sqrt{6} & -\sqrt{3} & i\sqrt{2} & 1 \\ \sqrt{6} & i\sqrt{3} & \sqrt{2} & -i \\ 0 & -i\sqrt{3} & 2\sqrt{2} & i \\ 0 & \sqrt{3} & 0 & 3 \end{bmatrix} \in M_4(\mathbb{C}),$$
dass U unitär ist und $D := U^* \cdot A \cdot U$ eine Diagonalmatrix mit Einträgen $0, 0, 4, 4$ auf der Diagonalen; Details in den Übungen (oder selbst).

24. ***Der Spekralsatz über*** $\mathbb{R}$

In vielen Anwendungen spielt eine Version der Spektralzerlegung über $\mathbb{R}$ eine wichtige Rolle. Wir können dies nun leicht aus der Spektralzerlegung über $\mathbb{C}$ herleiten. In Definition 23.2 haben wir bereits bemerkt, dass eine Matrix mit Einträgen in $\mathbb{R}$ genau dann hermitesch ist, wenn sie symmetrisch ist. Für solche Matrizen gilt die folgende zentrale Aussage:

Bemerkung 24.1. *Sei* $A \in M_n(\mathbb{R})$ *symmetrisch. Dann besitzt* A *einen Eigenwert* $\lambda \in \mathbb{R}$.
Denn: Da $\mathbb{C}$ algebraisch abgeschlossen ist, gibt es einen Eigenwert $\lambda \in \mathbb{C}$. Da A hermitesch ist, folgt $\lambda \in \mathbb{R}$ mit Lemma 23.9(b). — Viel später in diesem Semester werden wir noch einen anderen Beweis sehen, der nicht den Fundamentalsatz der Algebra benutzt (siehe §37).

Eine unitäre Matrix mit Einträgen in $\mathbb{R}$ heißt eine ***orthogonale Matrix***. Also: $T \in M_n(\mathbb{R})$ ist eine orthogonale Matrix genau dann, wenn $T^{\mathrm{tr}} \cdot T = I_n$ gilt. (Wegen $T \in M_n(\mathbb{R})$ ist $T^* = T^{\mathrm{tr}}$.) Analog zu Lemma 23.3 ist T orthogonal genau dann, wenn die Spalten von T eine Orthonormalbasis bezüglich des Standardskalarprodukts auf $\mathbb{R}^n$ bilden.

Bemerkung 24.2. Sei $T \in M_n(\mathbb{R})$ eine orthogonale Matrix. Ist $x \in \mathbb{R}^n$, so gilt
$$\|T \cdot x\|^2 = \langle T \cdot x, T \cdot x\rangle = (T \cdot x)^{\mathrm{tr}} \cdot (T \cdot x) = x^{\mathrm{tr}} \cdot (T^{\mathrm{tr}} \cdot T) \cdot x = x^{\mathrm{tr}} \cdot x = \|x\|^2,$$
und damit $\|T \cdot x\| = \|x\|$, d.h., T erhält die Norm von Vektoren. Etwas allgemeiner heißt eine Abbildung $\varphi \colon \mathbb{R}^n \to \mathbb{R}^n$ eine ***Bewegung***, wenn es eine orthogonale Matrix $T \in M_n(\mathbb{R})$ und einen Spaltenvektor $u \in \mathbb{R}^n$ gibt mit $\varphi(x) = T \cdot x + u$ für alle $x \in \mathbb{R}^n$. In diesem Fall gilt
$$\|\varphi(x) - \varphi(y)\| = \|T \cdot x - T \cdot y\| = \|T \cdot (x - y)\| = \|x - y\|,$$
für alle $x, y \in \mathbb{R}^n$, d.h., φ erhält die Abstände zwischen Vektoren.

Beispiel 24.3. Sei $n = 2$ und $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_2(\mathbb{R})$ orthogonal, also $A^{\mathrm{tr}} \cdot A = I_2$.
Ausmultiplizieren ergibt die Bedingungen $a^2 + c^2 = 1$, $b^2 + d^2 = 1$ und $ab + cd = 0$.
1. Fall: Ist $a = 0$, so folgt $c = \pm 1$ und dann auch $d = 0$, $c = \pm 1$.
2. Fall: Sei $a \neq 0$. Aus $ab + cd = 0$ erhalten wir $b = -cd/a$ und dann $1 = b^2 + d^2 = c^2d^2/a^2 + d^2 = (a^2 + c^2)d^2/a^2 = d^2/a^2$, also $d = \pm a$ und dann auch $c = \mp b$. Also ist $A = \begin{bmatrix} a & b \\ \mp b & \pm a \end{bmatrix}$ mit $a^2 + b^2 = 1$. Umgekehrt sieht man sofort, dass eine Matrix dieser Form orthogonal ist. Der 1. Fall passt auch in dieses Schema, mit $a = 0$ und $b = \pm 1$. Schließlich beachte: Sind $a, b \in \mathbb{R}$ gegeben mit $a^2 + b^2 = 1$, so gibt es ein eindeutiges $\theta \in [0, 2\pi)$ mit $a = \cos(\theta)$ und $b = \sin(\theta)$. Also gilt:
$$A \in M_2(\mathbb{R}) \text{ orthogonal} \quad \Leftrightarrow \quad A = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\delta \sin(\theta) & \delta \cos(\theta) \end{bmatrix} \text{ mit } \delta = \pm 1, \ \theta \in [0, 2\pi).$$
Ist $\delta = 1$, so ist $\det(A) = 1$ und die Abbildung $\mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto A \cdot x$, ist eine Drehung (im Uhrzeigersinn) um den Winkel θ. Sei nun $\delta = -1$. Dann ist $\det(A) = -1$ und

$$\chi_A = (\cos(\theta) - X) * (-\cos(\theta) - X) - \sin(\theta)^2 = X^2 - \cos(\theta)^2 - \sin(\theta)^2 = X^2 - 1.$$

Es gibt also die beiden Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = -1$. Mit Hilfe der Additionstheoreme für sin und cos in §10, Kapitel II, findet man zugehörige Eigenvektoren:

$$v_1 = \begin{bmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{bmatrix} \in \mathbb{R}^2 \qquad \text{und} \qquad v_2 = \begin{bmatrix} -\sin(\theta/2) \\ \cos(\theta/2) \end{bmatrix} \in \mathbb{R}^2.$$

In der reellen Ebene ist also $\mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto A \cdot x$, die Spiegelung an der Geraden durch den Ursprung und den durch v_1 gegebenen Punkt. Der Vektor $v_2 \in \mathbb{R}^2$ steht senkrecht auf dieser Geraden und wird auf $-v_2$ abgebildet.

Eine allgemeine Bewegung $\varphi \colon \mathbb{R}^2 \to \mathbb{R}^2$ wie in Bemerkung 24.2 ist damit eine Drehung oder Spiegelung, gefolgt von einer Verschiebung um einen festen Vektor $u \in \mathbb{R}^2$.

Lemma 24.4. *Seien $v, w \in \mathbb{R}^n$ mit $\|v\| = \|w\| > 0$. Dann gibt es eine orthogonale Matrix $T \in M_n(\mathbb{R})$ mit $T \cdot w = v$. Ist $r \in \{0, 1, \dots, n\}$ so, dass die ersten r Komponenten von v und von w gleich 0 sind, so kann T so gewählt werden, dass $T \cdot e_i = e_i$ für $1 \leqslant i \leqslant r$ gilt.*

Beweis. Sei $\beta := \|v\| = \|w\| > 0$. Sei $v' \in \mathbb{R}^{n-r}$ der Spaltenvektor mit Komponenten $\beta^{-1}v_{r+1}, \dots, \beta^{-1}v_n$ und $w' \in \mathbb{R}^{n-r}$ der Spaltenvektor mit Komponenten $\beta^{-1}w_{r+1}, \dots, \beta^{-1}w_n$. Dann gilt $\|v'\|' = \|w'\|' = 1$, wobei $\| \ \|'$ die Norm zum Standard-Skalarprodukt auf $\mathbb{R}^{n-r}$ bezeichne. Nach Lemma 23.4 können wir v'_1 zu einer Orthonormalbasis von $\mathbb{R}^{n-r}$ ergänzen; es gibt also eine orthogonale Matrix $T_1 \in M_{n-r}(\mathbb{R})$ mit $T_1 \cdot e'_1 = v'$ (wobei e'_1 der erste Standard-Basisvektor von $\mathbb{R}^{n-r}$ sei). Analog gibt es eine orthogonale Matrix $T_2 \in M_{n-r}(\mathbb{R})$ mit $T_2 \cdot e'_1 = w'$. Nun sind Produkte und Inverse von orthogonalen Matrizen wieder orthogonal, also ist auch $T' := T_1 \cdot T_2^{-1} \in M_{n-r}(\mathbb{R})$ orthogonal. Es gilt $T' \cdot w' = T_1 \cdot (T_2^{-1} \cdot w') = T_1 \cdot e'_1 = v'$. Nun bilden wir die Blockmatrix

$$T := \left[\begin{array}{c|c} I_r & 0_{n-r}^{tr} \\ \hline 0_{n-r} & T' \end{array}\right] \in M_n(\mathbb{R}).$$

Man sieht sofort, dass T orthogonal ist und die gewünschten Bedingungen erfüllt. □

Satz 24.5 (***Spektralsatz über*** $\mathbb{R}$)**.** *Sei $A \in M_n(\mathbb{R})$ symmetrisch, also $A^{tr} = A$. Dann gibt es eine orthogonale Matrix $T \in M_n(\mathbb{R})$ so dass $D := T^{-1} \cdot A \cdot T = T^{tr} \cdot A \cdot T \in M_n(\mathbb{R})$ eine Diagonalmatrix ist. Insbesondere ist A diagonalisierbar.*

Ab hier Woche 2

Beweis. Wir gehen noch einmal den Beweis von Satz 23.8 durch und achten darauf, dass wir in jedem Schritt im Bereich der reellen Zahlen bleiben können. Wegen $A^{tr} = A$ hat A nach Bemerkung 24.1 einen Eigenwert $\lambda_1 \in \mathbb{R}$. Es gibt dann auch einen zugehörigen Eigenvektor $w_1 \in \mathbb{R}^n$, also $w_1 \neq 0_n$ und $A \cdot w_1 = \lambda_1 w_1$. Setzen wir $v_1 := \|w_1\|^{-1} w_1 \in \mathbb{R}^n$, so ist auch v_1

ein Eigenvektor (mit Eigenwert λ_1) und $\|v_1\| = 1$. Nach Lemma 24.4 gibt es eine orthogonale Matrix $T_1 \in M_n(\mathbb{R})$ mit $T_1 \cdot e_1 = v_1$. Wie im Beweis von Satz 23.8 gilt dann

$$T_1^{tr} \cdot A \cdot T_1 = B = \left[\begin{array}{c|c} \lambda_1 & 0_{n-1}^{tr} \\ \hline 0_{n-1} & A_1 \end{array}\right] \qquad \text{mit} \qquad A_1 \in M_{n-1}(\mathbb{R}),$$

und man sieht leicht, dass auch A_1 symmetrisch ist. Wir können dann wieder mit Induktion fortfahren. Es gibt also eine orthogonale Matrix $T_2 \in M_{n-1}(\mathbb{R})$, so dass $D_1 := T_2^{tr} \cdot A_1 \cdot T_2$ eine Diagonalmatrix ist. Wir bilden $T_2' \in M_n(\mathbb{R})$ (analog wie im Beweis von Satz 23.8) und setzen $T := T_1 \cdot T_2' \in M_n(\mathbb{R})$. Dann ist T orthogonal und $T^{tr} \cdot A \cdot T$ eine Diagonalmatrix. □

Beispiel 24.6. Sei $A \in M_n(\mathbb{R})$ symmetrisch. Um eine orthogonale Matrix $T \in M_n(\mathbb{R})$ wie oben zu finden, können wir in 3 Schritten wie in Bemerkung 23.10 vorgehen: Zuerst Eigenwerte bestimmen, dann Basen der zugehörigen Eigenräume, und schließlich Transformation dieser Basen in Orthonormalbasen.

Konkretes Beispiel: Sei $A = A^{tr} = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \in M_3(\mathbb{R})$. Wir erhalten

$\chi_A = \det\left(\begin{bmatrix} -X & 1 & -1 \\ 1 & -X & 1 \\ -1 & 1 & -X \end{bmatrix}\right) = (X+2)(X-1)^2$; es gibt die Eigenwerte $\lambda_1 = -2$ und $\lambda_2 = 1$.

Der Eigenraum zu $\lambda_1 = -2$ hat Dimension 1, mit Basis $v_1 := \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$; es gilt $\langle v_1, v_1 \rangle = 3$.

Der Eigenraum zu $\lambda_2 = 1$ hat Dimension 2, mit Basis $\{v_2, v_3\}$ wobei $v_2 := \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$.

Hier ist $\langle v_2, v_2 \rangle = 2$ und $\langle v_2, v_3 \rangle = -1$; die Basis $\{v_2, v_3\}$ ist also noch keine Orthogonalbasis des Eigenraums. Gemäß Gram–Schmidt–Verfahren setzen wir daher $v_3' := v_3 + \frac{1}{2} v_2 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}$; dann folgt $\langle v_2, v_3' \rangle = 0$, $\langle v_3', v_3' \rangle = 3/2$ und $\{v_2, v_3'\}$ ist eine Orthogonalbasis des Eigenraums. Sei $T \in M_3(\mathbb{R})$ die Matrix mit Spalten gegeben durch $\sqrt{3}^{-1} v_1$, $\sqrt{2}^{-1} v_2$, $\sqrt{3/2}^{-1} v_3'$. Dann ist T orthogonal und es gilt $T^{tr} \cdot A \cdot T =$ Diagonalmatrix mit Diagonaleinträgen $-2, 1, 1$.

Bemerkung 24.7. Sei K beliebiger Körper und $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \in M_2(K)$ symmetrisch.

(a) Sei $K = \mathbb{R}$. Dann ist $\chi_A = \ldots = X^2 - (a+c)X + (ac - b^2) = \left((X - \frac{1}{2}(a+c)\right)^2 + \Delta \in \mathbb{R}[X]$ mit $\Delta = ac - b^2 - \frac{1}{4}(a+c)^2 = -b^2 - \frac{1}{4}(a-c)^2 \leqslant 0$; also zerfällt χ_A tatsächlich über $\mathbb{R}$.

(b) Ist $K = \mathbb{C}$ und zum Beispiel $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$, so ist $\chi_A = X^2$, also $\lambda = 0$ der einzige Eigenwert. Aber A ist nicht diagonalisierbar.

(c) Ist $K = \mathbb{F}_2$ und zum Beispiel $A = \begin{bmatrix} \bar{0} & \bar{1} \\ \bar{1} & \bar{0} \end{bmatrix}$, so ist $\chi_A = (X + \bar{1})^2$, also $\lambda = \bar{1}$ der einzige Eigenwert. Aber A ist wiederum nicht diagonalisierbar.

Also: Für die Gültigkeit des Spektralsatzes ist es wesentlich, dass $K = \mathbb{R}$ der Grundkörper ist.

25. *Hauptachsentransformation*

Als eine klassische Anwendung des Spektralsatzes (über $\mathbb{R}$) behandeln wir in diesem Abschnitt einen kleinen Höhepunkt der elementaren analytischen Geometrie. Dabei geht es um Lösungsmengen nicht nur von linearen Gleichungen, sondern von quadratischen Gleichungen in n Variablen. Für $n = 2$ hat eine solche Gleichung die allgemeine Form

$$f(x_1, x_2) := \alpha_1 x_1^2 + \alpha_2 x_2^2 + \alpha_{12} x_1 x_2 + \beta_1 x_1 + \beta_2 x_2 + \gamma = 0,$$

wobei die Koeffizienten $\alpha_1, \alpha_2, \alpha_{12}, \beta_1, \beta_2, \gamma \in \mathbb{R}$ fest vorgegeben und alle $x_1, x_2 \in \mathbb{R}$ gesucht sind, so dass $f(x_1, x_2) = 0$ gilt. Zum Beispiel ist für $f(x_1, x_2) = x_1^2 + x_2^2 - 1$ die Lösungsmenge ein Kreis in der reellen Ebene mit Radius 1 und Mittelpunkt im Ursprung $(0, 0)$. Gibt es mehr Terme in der Gleichung mit Koeffizienten ungleich 0, so wird es schon schwieriger, gleich zu sehen, um welche geometrische Figur es sich handelt; Beispiele:

$$f(x_1, x_2) = 5x_1^2 + 2\sqrt{3}x_1 x_2 + 7x_2^2 - 16 = 0,$$
$$f(x_1, x_2) = x_1^2 + 2x_1 x_2 + x_2^2 + 3x_1 - 4x_2 + 1 = 0.$$

Die Idee ist nun, eine Koordinatentransformation vorzunehmen, so dass in den “neuen” Koordinaten die Gleichung “möglichst einfach” wird. Behandeln wir dies gleich für beliebiges $n \geqslant 1$. Die allgemeine Form einer quadratischen Gleichung ist

$$f(x_1, \ldots, x_n) := \underbrace{\alpha_1 x_1^2 + \ldots + \alpha_n x_n^2}_{\text{rein quadratisch}} + \underbrace{\sum_{1 \leqslant i < j \leqslant n} \alpha_{ij} x_i x_j}_{\text{gemischt quadratisch}} + \underbrace{\beta_1 x_1 + \ldots + \beta_n x_n}_{\text{linear}} + \gamma = 0,$$

mit vorgegebenen Koeffizienten $\alpha_i, \alpha_{ij}, \beta_j, \gamma \in \mathbb{R}$. Um triviale Sonderfälle zu vermeiden, nehmen wir an, dass mindestens ein α_i oder ein α_{ij} ungleich 0 ist, also quadratische Terme (rein oder gemischt) tatsächlich vorkommen. Die Lösungsmenge

$$Q := \big\{(x_1, \ldots, x_n) \in \mathbb{R}^{1\times n} \,\big|\, f(x_1, \ldots, x_n) = 0\big\}$$

heißt ***Quadrik*** oder auch ***Hyperfläche zweiten Grades***; für $n = 2$ spricht man auch von ***Kegelschnitten***. Siehe zum Beispiel `https://de.wikipedia.org/wiki/Quadrik` für mehr Hintergrund dazu. Als geeignete Koordinatentransformation betrachten wir eine Bewegung $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$; es gibt also eine orthogonale Matrix $T = [t_{ij}] \in M_n(\mathbb{R})$ und ein $u \in \mathbb{R}^n$ mit $\varphi(y) = T \cdot y + u$ für alle $y \in \mathbb{R}^n$. Seien $u_1, \ldots, u_n \in \mathbb{R}$ die Komponenten von u. Jedes

$x \in \mathbb{R}^n$ können wir schreiben als $x = \varphi(y) = T \cdot y + u$ mit $y \in \mathbb{R}^n$. Sind $x_1, \ldots, x_n \in \mathbb{R}$ die Komponenten von x und $y_1, \ldots, y_n \in \mathbb{R}$ die Komponenten von y, so gilt also

$$x_i = \Big(\sum_{j=1}^{n} t_{ij} y_j\Big) + u_i \qquad \text{für } 1 \leqslant i \leqslant n.$$

Setzen wir diese Ausdrücke in die Gleichung $f(x_1, \ldots, x_n) = 0$ ein, so erhalten wir eine neue quadratische Gleichung $f'(y_1, \ldots, y_n) = 0$ in den Variablen $y_1, \ldots, y_n$.

Satz 25.1 (***Normalformen von Quadriken***). *Es gibt eine Bewegung* $\varphi\colon \mathbb{R}^n \to \mathbb{R}^n$, *so dass die ursprüngliche quadratische Gleichung* $f(x_1, \ldots, x_n) = 0$ *mittels der Transformation* $x = \varphi(y)$ *in eine der folgenden Gleichungen in* $y_1, \ldots, y_n$ *überführt werden kann:*

1) $\lambda_1 y_1^2 + \ldots + \lambda_r y_r^2 + \gamma = 0$ *mit* $1 \leqslant r \leqslant n$, *oder*
2) $\lambda_1 y_1^2 + \ldots + \lambda_r y_r^2 + \beta y_n = 0$ *mit* $1 \leqslant r < n$,

wobei jeweils $\lambda_i, \beta, \gamma \in \mathbb{R}$ *sowie* $\beta \neq 0$ *und* $\lambda_i \neq 0$ *für* $1 \leqslant i \leqslant r$ *gilt.*

Eine Bewegung, die die ursprüngliche Gleichung $f(x_1, \ldots, x_n) = 0$ in eine der Gleichungen 1) oder 2) überführt, heißt ***Hauptachsentransformation***.

Beweis. Wir definieren eine Matrix $A = [a_{ij}] \in M_n(\mathbb{R})$ durch $a_{ij} := \begin{cases} \alpha_i & \text{für } i = j, \\ \alpha_{ij}/2 & \text{für } i < j, \\ \alpha_{ji}/2 & \text{für } i > j; \end{cases}$
dann ist $A \neq 0_{n\times n}$ (weil mindestens ein α_i oder ein α_{ij} ungleich 0 ist). Sei $b \in \mathbb{R}^{1\times n}$ der Zeilenvektor mit Komponenten $\beta_1, \ldots, \beta_n$. Dann ist A symmetrisch und es gilt

$$f(x_1, \ldots, x_n) = x^{\mathrm{tr}} \cdot A \cdot x + b \cdot x + \gamma \qquad \text{für alle } x \in \mathbb{R}^n.$$

Wir zeigen nun, dass man in maximal 4 Schritten diese Gleichung auf die Form 1) oder 2) transformieren kann.

<u>1. Schritt</u>. Nach dem Spektralsatz gibt es eine orthogonale Matrix $T \in M_n(\mathbb{R})$, so dass $D := T^{-1} \cdot A \cdot T = T^{\mathrm{tr}} \cdot A \cdot T$ eine Diagonalmatrix ist; seien $\lambda_1 \ldots, \lambda_n \in \mathbb{R}$ die Diagonaleinträge von D. Wir können diese so anordnen, dass $\lambda_i \neq 0$ für $1 \leqslant i \leqslant r$ und $\lambda_{r+1} = \ldots = \lambda_n = 0$ gilt (wobei $r \in \{1, \ldots, n\}$, weil $A \neq 0_{n\times n}$). Betrachte die Bewegung $\varphi_1\colon \mathbb{R}^n \to \mathbb{R}^n$, $y \mapsto T \cdot y$; schreibe $x = \varphi_1(y)$ und setze $b' := b \cdot T \in \mathbb{R}^{1\times n}$; dann hat die neue Gleichung die Form

$$\begin{aligned} f'(y_1, \ldots, y_n) &= (T \cdot y)^{\mathrm{tr}} \cdot A \cdot (T \cdot y) + b \cdot (T \cdot y) + \gamma \\ &= y^{\mathrm{tr}} \cdot D \cdot y + (b \cdot T) \cdot y + \gamma = \lambda_1 y_1^2 + \ldots + \lambda_r y_r^2 + b' \cdot y + \gamma; \end{aligned}$$

d.h., es gibt keine gemischt quadratischen Terme mehr. — Je nach Fall werden wir jetzt noch weitere Bewegungen anwenden müssen, um 1) oder 2) zu erreichen.

<u>2. Schritt</u>. Damit wir nicht ständig neue Namen für die Variablen einführen müssen, nehmen wir nun an, dass die Ursprungsgleichung die gerade erreichte Form hat, also

$$f(x_1, \ldots, x_n) = \lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 + \beta_1 x_1 + \cdots + \beta_n x_n + \gamma,$$

wobei $\lambda_i, \beta_i, \gamma \in \mathbb{R}$ und $\lambda_i \neq 0$ für $1 \leqslant i \leqslant r$. Betrachte nun die Bewegung $\varphi_2\colon \mathbb{R}^n \to \mathbb{R}^n$, $y \mapsto y + c$, wobei $c \in \mathbb{R}^n$ ein Spaltenvektor mit Komponenten $c_1, \ldots, c_r, 0, \ldots, 0$ sei (und

$c_1, \ldots, c_r \in \mathbb{R}$ noch zu bestimmen sind). Schreibe $x_i = y_i + c_i$ für $1 \leqslant i \leqslant r$ und $x_i = y_i$ für $r+1 \leqslant i \leqslant n$. Dann ergibt sich als neue Gleichung

$$\begin{aligned} f'(y_1, \ldots, y_n) &= \sum_{i=1}^{r} \lambda_i (y_i + c_i)^2 + \sum_{i=1}^{r} \beta_i (y_i + c_i) + \sum_{i=r+1}^{n} \beta_i y_i + \gamma \\ &= \sum_{i=1}^{r} \lambda_i y_i^2 + \sum_{i=1}^{r} (2\lambda_i c_i + \beta_i) y_i + \sum_{i=r+1}^{n} \beta_i y_i + \underbrace{\sum_{i=1}^{r} (\lambda_i c_i^2 + \beta_i c_i) + \gamma}_{\gamma' :=} \end{aligned}$$

Mit $c_i := -\beta_i/(2\lambda_i)$ für $1 \leqslant i \leqslant r$ erhalten wir:

$$f'(y_1, \ldots, y_n) = \sum_{i=1}^{r} \lambda_i y_i^2 + \sum_{i=r+1}^{n} \beta_i y_i + \gamma',$$

d.h., die Variablen in den quadratischen und den linearen Termen wurden getrennt.

3. Schritt. Nehmen wir nun wieder an, dass die Ursprungsgleichung die gerade erreichte Form hat, also $f(x_1, \ldots, x_n) = \lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 + \beta_{r+1} x_{r+1} + \ldots + \beta_n x_n + \gamma$, wobei $\lambda_i, \beta_i, \gamma \in \mathbb{R}$ und $\lambda_i \neq 0$ für $1 \leqslant i \leqslant r$. Ist $r = n$ oder $\beta_{r+1} = \ldots = \beta_n = 0$, so haben wir 1) erreicht. Nehmen wir schließlich an, es sei $r < n$ und es gibt ein $k \in \{r+1, \ldots, n\}$ mit $\beta_k \neq 0$. Betrachte nun die Bewegung $\varphi_3 \colon \mathbb{R}^n \to \mathbb{R}^n$, $y \mapsto y + c e_k$, wobei $c \in \mathbb{R}$ noch zu bestimmen ist. Schreibe $x_k = y_k + c$ und $x_i = y_i$ für $i \neq k$. Dann ergibt sich als neue Gleichung

$$f'(y_1, \ldots, y_n) = \sum_{i=1}^{r} \lambda_i y_i^2 + \sum_{i=r+1}^{n} \beta_i y_i + \underbrace{\beta_k c + \gamma}_{\gamma :=}.$$

Mit $c := -\gamma/\beta_k$ erhalten wir $\gamma' = 0$, d.h., es gibt keinen konstanten Term mehr.

4. Schritt. Nehmen wir wiederum an, dass die Ursprungsgleichung die gerade erreichte Form hat, also $f(x_1, \ldots, x_n) = \lambda_1 x_1^2 + \ldots + \lambda_r x_r^2 + \beta_{r+1} x_{r+1} + \ldots + \beta_n x_n$, wobei $\lambda_i, \beta_i \in \mathbb{R}$ und $\lambda_i \neq 0$ für $1 \leqslant i \leqslant r$; außerdem ist mindestens ein β_i ungleich 0. Um 2) zu erreichen, müssen wir noch eine Bewegung $\varphi_4 \colon \mathbb{R}^n \to \mathbb{R}^n$ finden, so dass nach der Transformation $x = \varphi_4(y)$ nur der lineare Term in y_n übrig bleibt.

Dazu: Sei $v \in \mathbb{R}^n$ der Spaltenvektor mit Komponenten $0, \ldots, 0, \beta_{r+1}, \ldots, \beta_n$. Dann ist $v \neq 0_n$; sei $\beta := \|v\| > 0$. Sei $w := \beta e_n \in \mathbb{R}^n$; dann ist $\|w\| = \beta = \|v\|$. Sei dann $\varphi_4 \colon \mathbb{R}^n \to \mathbb{R}^n$, $y \mapsto T \cdot y$, wobei $T \in M_n(\mathbb{R})$ eine orthogonale Matrix wie in Lemma 24.4 ist. (Beachte, dass die Voraussetzung an die ersten r Komponenten von v und w erfüllt ist.) Schreibe wieder $x = \varphi_4(y)$; wegen $T \cdot e_i = e_i$ ist dann $x_i = y_i$ für $1 \leqslant i \leqslant r$. Als neue Gleichung ergibt sich:

$$f'(y_1, \ldots, y_n) = \sum_{i=1}^{r} \lambda_i y_i^2 + \sum_{i=r+1}^{n} \beta_i \Big(\sum_{j=1}^{n} t_{ij} y_j \Big) = \sum_{i=1}^{r} \lambda_i y_i^2 + \underbrace{\sum_{j=1}^{n} \Big(\sum_{i=r+1}^{n} \beta_i t_{ij} \Big) y_j}_{=(v^{\mathrm{tr}} \cdot T) \cdot y}.$$

Nun ist $T \cdot w = v$, also $v^{\mathrm{tr}} = w^{\mathrm{tr}} \cdot T^{\mathrm{tr}} = w^{\mathrm{tr}} \cdot T^{-1}$ und damit $v^{\mathrm{tr}} \cdot T = w^{\mathrm{tr}} = \beta e_n^{\mathrm{tr}}$. Schließlich folgt $(v^{\mathrm{tr}} \cdot T) \cdot y = \beta e_n^{\mathrm{tr}} \cdot y = \beta y_n$, wie gewünscht. $\square$

Beispiel 25.2. Sei $f(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2 + 3x_1 - 4x_2 + 1$. Wir gehen die obigen Schritte durch. Die Matrix A ist gegeben durch $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, mit $\chi_A = \det(A - XI_2) = X^2 - 2X$; es gibt also die beiden Eigenwerte $\lambda_1 = 2$ und $\lambda_2 = 0$. Wie in Beispiel 24.6 finden wir

$$T^{-1}AT = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \qquad \text{mit der orthogonalen Matrix } T = \sqrt{2}^{-1} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Die Transformation $x_1 = \sqrt{2}^{-1}(y_1 + y_2)$, $x_2 = \sqrt{2}^{-1}(y_1 - y_2)$ ergibt die neue Gleichung

$$f'(y_1, y_2) = 2y_1^2 - \sqrt{2}^{-1} y_1 + 7\sqrt{2}^{-1} y_2 + 1 = 0.$$

Schreibe nun $y_1 = z_1 + u$, $y_2 = z_2$ wobei $u \in \mathbb{R}$ noch zu bestimmen ist. Einsetzen ergibt

$$2(z_1+u)^2 - \sqrt{2}^{-1}(z_1+u) + 7\sqrt{2}^{-1} z_2 + 1 = 2z_1^2 + 4z_1u + 2u^2 - \sqrt{2}^{-1} z_1 - \sqrt{2}^{1} u + 7\sqrt{2}^{-1} z_2 + 1.$$

Mit $u := (4\sqrt{2})^{-1}$ erhalten wir die Gleichung $f''(z_1, z_2) = 2z_1^2 + 7\sqrt{2}^{-1} z_2 + 15/16 = 0$. Schließlich schreibe $z_1 = w_1$, $z_2 = w_2 + r$ wobei $r \in \mathbb{R}$ so zu bestimmen ist, dass der konstante Term verschwindet. Mit $r = -15\sqrt{2}/112$ erhalten wir die Gleichung $f'''(w_1, w_2) = 2w_1^2 + 7\sqrt{2}^{-1} w_2 = 0$, was in der reellen Ebene $\mathbb{R}^2$ (mit Koordinatenachsen w_1, w_2) einer Parabel entspricht.

26. *Singulärwertzerlegung*

Als eine weitere Folgerung aus dem Spektralsatz betrachten wir die ***Singulärwertzerlegung*** einer Matrix (englisch: ***singular value decomposition***, kurz ***SVD***). Diese wird in zahlreichen Anwendungen aus den unterschiedlichsten Bereichen eingesetzt; siehe

C. D. Martin and M. A. Porter, The extraordinary SVD, Amer. Math. Monthly **119** (2012), 838–851; http://dx.doi.org/10.4169/amer.math.monthly.119.10.838

für einen Überblick mit vielen weiterführenden Referenzen. Außerdem ist dies eine schöne Illustration, wie die bisher erzielten Ergebnisse zusammenwirken.

Lemma 26.1. *Sei* $A \in \mathbb{R}^{m \times n}$ *beliebig und* $r := \mathrm{Rang}(A)$. *Dann gilt:*

(a) *Es ist auch* $\mathrm{Rang}(A^{\mathrm{tr}} \cdot A) = r$.

(b) *Die Eigenwerte von* $A^{\mathrm{tr}} \cdot A \in M_n(\mathbb{R})$ *sind reell und* $\geqslant 0$; *es gibt* r *Eigenwerte* $\neq 0$.

Beweis. (a) Sei $x \in N(A)$, also $A \cdot x = 0_m$. Dann ist auch $(A^{\mathrm{tr}} \cdot A) \cdot x = 0_n$, also $x \in N(A^{\mathrm{tr}} \cdot A)$. Umgekehrt: Sei $x \in N(A^{\mathrm{tr}} \cdot A)$, also $(A^{\mathrm{tr}} \cdot A) \cdot x = 0_n$. Dann ist auch $\|A \cdot x\|^2 = \langle A \cdot x, A \cdot x \rangle = (A \cdot x)^{\mathrm{tr}} \cdot (A \cdot x) = x^{\mathrm{tr}} \cdot (A^{\mathrm{tr}} \cdot A) \cdot x = 0$, also $A \cdot x = 0_m$ und damit $x \in N(A)$. Damit ist $N(A) = N(A^{\mathrm{tr}} \cdot A)$ gezeigt, also folgt mit Lemma 20.6 (Kapitel IV):

$$\mathrm{Rang}(A) = n - \dim N(A) = n - \dim N(A^{\mathrm{tr}} \cdot A) = \mathrm{Rang}(A^{\mathrm{tr}} \cdot A).$$

(b) Sei $B := A^{\mathrm{tr}} \cdot A \in M_n(\mathbb{R})$; dann ist B symmetrisch. Nach dem Spekralsatz gibt es eine orthogonale Matrix $V \in M_n(\mathbb{R})$ so dass $D := V^{-1} \cdot B \cdot V = V^{\mathrm{tr}} \cdot B \cdot V$ eine Diagonalmatrix ist.

Nun ist V ein Produkt von Elementarmatrizen; damit entsteht D aus B durch eine endliche Anzahl von elementaren Zeilen- und Spaltenumformungen. Dabei ändert sich der Rang nicht (siehe die Bemerkungen in Definition 20.6, Kapitel IV), also gilt $\mathrm{Rang}(D) = \mathrm{Rang}(B) = r$, d.h., genau r der Diagonaleinträge von D sind ungleich 0. Sei schließlich $\lambda \in \mathbb{R}$ ein Eigenwert von $B = A^{tr} \cdot A$ (also ein Diagonaleintrag von D) und $v \in \mathbb{R}^n$ ein zugehöriger Eigenvektor, also $v \neq 0_n$ und $B \cdot v = \lambda v$. Dann gilt $\|A \cdot v\|^2 = \langle A \cdot v, A \cdot v\rangle = (A \cdot v)^{tr} \cdot (A \cdot v) = v^{tr} \cdot (B \cdot v) = v^{tr} \cdot (\lambda v) = \lambda \|v\|^2$. Wegen $\|v\| > 0$ und $\|A \cdot v\| \geqslant 0$ folgt dann auch $\lambda \geqslant 0$. □

Satz 26.2 (***SVD***). *Sei* $A \in \mathbb{R}^{m\times n}$ *mit* $m, n \geqslant 1$ *beliebig; sei* $r = \mathrm{Rang}(A) \geqslant 0$. *Dann gibt es orthogonale Matrizen* $U \in M_m(\mathbb{R})$ *und* $V \in M_n(\mathbb{R})$ *mit* $A = U \cdot S \cdot V^{tr}$, *wobei* $S = [s_{ij}] \in \mathbb{R}^{m\times n}$ *Diagonalgestalt hat, mit* $s_{ij} = 0$ *für alle* $i \neq j$ *sowie* $s_{11} \geqslant \ldots \geqslant s_{rr} > 0$ *und* $s_{ii} = 0$ *für* $i > r$.

Die Diagonaleinträge $s_{11}, s_{22}, \ldots$ von S heißen ***Singulärwerte*** von A.

Beweis. Sei $B := A^{tr} \cdot A \in M_n(\mathbb{R})$. Wie oben im Beweis gibt es eine orthogonale Matrix $V \in M_n(\mathbb{R})$ so dass $V^{-1} \cdot B \cdot V = V^{tr} \cdot B \cdot V$ eine Diagonalmatrix ist; seien $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ die Einträge auf der Diagonalen. Wir könnnen dies so einrichten, dass $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$ gilt. Nach Lemma 26.1 gilt $\lambda_i \geqslant 0$ für alle i; außerdem ist $\lambda_i = 0$ für $i > r$. Setze $\alpha_i := \sqrt{\lambda_i} \in \mathbb{R}$ für $1 \leqslant i \leqslant n$; dann ist auch $\alpha_1 \geqslant \ldots \geqslant \alpha_r > \alpha_{r+1} = \ldots = \alpha_n = 0$. Seien $v_1, \ldots, v_n \in \mathbb{R}^n$ die Spalten von V; dann gilt $B \cdot v_i = \lambda_i v_i$ für $1 \leqslant i \leqslant n$. Wir setzen $u_j := \alpha_j^{-1} A \cdot v_j \in \mathbb{R}^m$ für $1 \leqslant j \leqslant r$. Weil V eine orthogonale Matrix ist, ist $\{v_1, \ldots, v_n\}$ eine Orthonormalbasis von $\mathbb{R}^n$. Damit folgt für $1 \leqslant i, j \leqslant r$:

$$\begin{aligned}\langle u_i, u_j\rangle &= \alpha_i^{-1}\alpha_j^{-1}\langle A \cdot v_i, A \cdot v_j\rangle = \alpha_i^{-1}\alpha_j^{-1}\big(v_i^{tr} \cdot A^{tr} \cdot A \cdot v_j\big) = \alpha_i^{-1}\alpha_j^{-1}\big(v_i^{tr} \cdot (B \cdot v_j)\big)\\ &= \alpha_i^{-1}\alpha_j^{-1}\big(v_i^{tr} \cdot (\lambda_j v_j)\big) = \alpha_i^{-1}\alpha_j^{-1}\lambda_j\langle v_i, v_j\rangle.\end{aligned}$$

Ist $i \neq j$, so ist $\langle v_i, v_j\rangle = 0$, also auch $\langle u_i, u_j\rangle = 0$. Ist $i = j$, so ist $\lambda_j = \alpha_i\alpha_j$ und $\langle v_i, v_i\rangle = 1$, also auch $\langle u_i, u_i\rangle = 1$. Nach Lemma 23.4 gibt es $u_{r+1}, \ldots, u_m \in \mathbb{R}^m$ so dass $\{u_1, \ldots, u_m\}$ eine Orthonormalbasis von $\mathbb{R}^m$ ist. Sei $U \in M_m(\mathbb{R})$ die orthogonale Matrix mit Spalten $u_1, \ldots, u_m$. Jetzt betrachte $S = [s_{ij}] := U^{-1} \cdot A \cdot V = U^{tr} \cdot A \cdot V$. Für $1 \leqslant i \leqslant m$ und $1 \leqslant j \leqslant n$ gilt also $s_{ij} = u_i^{tr} \cdot A \cdot v_j$. Für i beliebig und $1 \leqslant j \leqslant r$ gilt $A \cdot v_j = \alpha_j u_j$ und damit $s_{ij} = u_i^{tr} \cdot A \cdot v_j = \alpha_j\langle u_i, u_j\rangle = \alpha_j\delta_{ij}$. Für $j > r$ ist $\|A \cdot v_j\|^2 = \langle A \cdot v_j, A \cdot v_j\rangle = v_j^{tr} \cdot (A^{tr} \cdot A) \cdot v_j = v_j^{tr} \cdot (B \cdot v_j) = v_j^{tr} \cdot 0_n = 0$, also $A \cdot v_j = 0_m$ und damit auch $s_{ij} = u_i^{tr} \cdot A \cdot v_j = 0$ für alle i. Folglich ist $A = U \cdot S \cdot V^{-1} = U \cdot S \cdot V^{tr}$ und S hat die gewünschte Diagonalgestalt. □

Beispiel 26.3. Sei $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 2}$. Dann ist $B := A^{tr} \cdot A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \in M_2(\mathbb{R})$ mit $\chi_B = (X - 3)(X - 1)$. Mit dem Verfahren in Beispiel 24.6 finden wir eine orthogonale Matrix $V \in M_2(\mathbb{R})$, so dass $V^{tr} \cdot B \cdot V$ eine Diagonalmatrix ist; in diesem Fall:

$$V^{\mathrm{tr}} \cdot B \cdot V = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \qquad \text{mit} \qquad V = \sqrt{2}^{-1} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Die Singulärwerte von A sind also $\alpha_1 = \sqrt{3}$ und $\alpha_2 = 1$. Seien $v_1, v_2 \in \mathbb{R}^2$ die beiden Spalten von V. Wie oben im Beweis setzen wir anschließend

$$u_1 := \alpha_1^{-1} A \cdot v_1 = \sqrt{6}^{-1} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \qquad \text{und} \qquad u_2 := \alpha_2^{-1} A \cdot v_2 = \sqrt{2}^{-1} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}.$$

Zum Schluss müssen wir noch einen Spaltenvektor $u_3 \in \mathbb{R}^3$ finden, so dass die Matrix $U \in M_3(\mathbb{R})$ mit Spalten u_1, u_2, u_3 eine orthogonale Matrix ist; in diesem Fall geht dies mit

$$u_3 = \sqrt{3}^{-1} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \qquad \text{und damit} \qquad U := \sqrt{6}^{-1} \begin{bmatrix} 1 & \sqrt{3} & \sqrt{2} \\ 2 & 0 & -\sqrt{2} \\ 1 & -\sqrt{3} & \sqrt{2} \end{bmatrix}.$$

Mit $S := \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 2}$ erhalten wir also die Singulärwertzerlegung $A = U \cdot S \cdot V^{\mathrm{tr}}$.

Bemerkung 26.4. Sei $A = U \cdot S \cdot V^{\mathrm{tr}}$ und $r = \mathrm{Rang}(A)$, wie in Satz 26.2. Für $1 \leqslant k \leqslant r$ sei $S_{(k)} \in \mathbb{R}^{m \times n}$ die Matrix in Diagonalgestalt mit Diagonaleinträgen $s_{11}, \ldots, s_{kk}$, d.h., $S_{(k)}$ entsteht aus S, indem wir die ersten k Diagonaleinträge in S beibehalten und die restlichen Diagonaleinträge an den Positionen $k+1, \ldots, r$ alle gleich 0 setzen. Dann heißt

$$A_{(k)} := U \cdot S_{(k)} \cdot V^{\mathrm{tr}} \in \mathbb{R}^{m \times n} \qquad \text{“}\textbf{\textit{Rang-}}k\textbf{\textit{-Approximation}}\text{” von } A.$$

Beachte, dass tatsächlich $\mathrm{Rang}(A_{(k)}) = k$ gilt; man kann zeigen, dass $A_{(k)}$ die “beste” Approximation von A durch eine Matrix mit Rang gleich k ist (Satz von Eckart–Young; siehe §1 im oben zitierten Artikel von Martin und Porter). Für $1 \leqslant i \leqslant k$ sei $u_i \in \mathbb{R}^m$ die i-te Spalte von U und $v_i \in \mathbb{R}^n$ die i-te Spalte von V. Dann ist $u_i \cdot v_i^{\mathrm{tr}} \in \mathbb{R}^{m \times n}$ und die Definition der Matrixmultiplikation zeigt sofort folgende Formel:

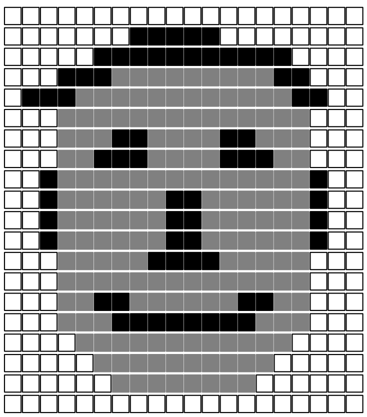
$$A_{(k)} := \sum_{i=1}^{k} s_{ii} \big(u_i \cdot v_i^{\mathrm{tr}}\big).$$

Die mn Einträge von $A_{(k)}$ sind also bestimmt durch die insgesamt $k(m+n+1)$ Einträge in den Spaltenvektoren $u_1, \ldots, u_k \in \mathbb{R}^m$, $v_1, \ldots, v_k \in \mathbb{R}^n$ plus den k Singulärwerten $d_{11}, \ldots, d_{kk}$.

Ab hier Woche 3

Dies findet Anwendungen zum Beispiel im Bereich der ***Datenkompression***, wo man es mit dem Problem der Speicherung einer großen Matrix A zu tun hat. Wenn es nicht auf die exakte Kenntnis jedes einzelnen Eintrags von A ankommt, so dürfen die Werte der gespeicherten Matrix innerhalb einer vorgegebenen Fehlergrenze von den tatsächlichen Werten abweichen, ohne dass das Gesamtbild zu stark verfälscht wird. (Denken Sie zum Beispiel an die Millionen von Farb- und Helligkeitswerten in einem digitalen Photo, das man auf einem kleinen Bildschirm betrachten möchte; kleinere Abweichungen von den Ursprungswerten werden dann nicht ins Auge fallen.) Ersetzt man A durch eine Rang-k-Approximation, so ergibt sich je nach Wahl von k unter Umständen eine enorme Ersparnis im Speicherplatz.

Beispiel 26.5. Zur Veranschaulichung betrachten wir ein extrem vereinfachtes Beispiel, mit digitalen Schwarz-Weiss-Photos in einem 20×20 Raster; jeder Punkt im Raster ist entweder weiss, grau oder schwarz. Wir stellen dies mit einer Matrix $A = [a_{ij}] \in M_{20}(\mathbb{R})$ dar, wo a_{ij} gleich 0, 1 oder 2 ist, je nach dem, ob der entsprechende Punkt im Raster weiss, grau oder schwarz ist. Für ein solches Photo müssen wir also die 400 Einträge von A speichern. (Die Photos mit heutigen Smartphones brauchen natürlich viel mehr Speicherplatz.) Beispiel:



$$A = \begin{bmatrix}
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&2&2&2&2&2&0&0&0&0&0&0&0&0\\
0&0&0&0&0&2&2&2&2&2&2&2&2&2&2&2&0&0&0&0\\
0&0&0&2&2&2&1&1&1&1&1&1&1&1&1&2&2&0&0&0\\
0&2&2&2&1&1&1&1&1&1&1&1&1&1&1&1&2&2&0&0\\
0&0&0&1&1&1&1&1&1&1&1&1&1&1&1&1&1&0&0&0\\
0&0&0&1&1&1&2&2&1&1&1&1&2&2&1&1&1&0&0&0\\
0&0&0&1&1&2&2&2&1&1&1&1&2&2&2&1&1&0&0&0\\
0&0&2&1&1&1&1&1&1&1&1&1&1&1&1&1&1&2&0&0\\
0&0&2&1&1&1&1&1&1&2&2&1&1&1&1&1&1&2&0&0\\
0&0&2&1&1&1&1&1&1&2&2&1&1&1&1&1&1&2&0&0\\
0&0&2&1&1&1&1&1&1&2&2&1&1&1&1&1&1&2&0&0\\
0&0&0&1&1&1&1&1&2&2&2&2&1&1&1&1&1&0&0&0\\
0&0&0&1&1&1&1&1&1&1&1&1&1&1&1&1&1&0&0&0\\
0&0&0&1&1&2&2&1&1&1&1&1&1&2&2&1&1&0&0&0\\
0&0&0&1&1&1&2&2&2&2&2&2&2&2&1&1&1&0&0&0\\
0&0&0&0&1&1&1&1&1&1&1&1&1&1&1&1&0&0&0&0\\
0&0&0&0&0&1&1&1&1&1&1&1&1&1&1&0&0&0&0&0\\
0&0&0&0&0&0&1&1&1&1&1&1&1&1&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0
\end{bmatrix}.$$

Sei $A = U \cdot S \cdot V^{\mathrm{tr}}$ die SVD von A. Zum Beispiel mit Sage erhalten wir diese wie folgt:

```
SageMath version 9.3, Release Date: 2021-05-09
Using Python 3.9.2. Type "help()" for help.
sage: A=matrix(RDF,[                           # obige Matrix A
  [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0],
  [0,0,0,0,0,0,0,2,2,2,2,2,0,0,0,0,0,0,0,0],
  [0,0,0,0,0,2,2,2,2,2,2,2,2,2,2,2,0,0,0,0],
  ...
  [0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0]])
sage: U,S,V=A.SVD()
```

Der Rang von A ist gleich 12; die Singulärwerte ungleich 0 (also die ersten 12 Diagonaleinträge von S) sind wie folgt in Dezimaldarstellung gegeben, gerundet auf 2 Nachkommastellen:

$$18{,}88 \quad 6{,}02 \quad 4{,}71 \quad 3{,}06 \quad 2{,}27 \quad 1{,}79 \quad 1{,}26 \quad 1{,}15 \quad 0{,}79 \quad 0{,}69 \quad 0{,}56 \quad 0{,}40.$$

Die Rang-4-Approximation von A ist in Tabelle 1 angegeben, auf eine Nachkommastelle gerundet. Runden wir die Werte jeweils zu 0, 1, 2 auf oder ab, so erhalten wir eine Annäherung an das ursprüngliche Bild; siehe ebenfalls Tabelle 1.

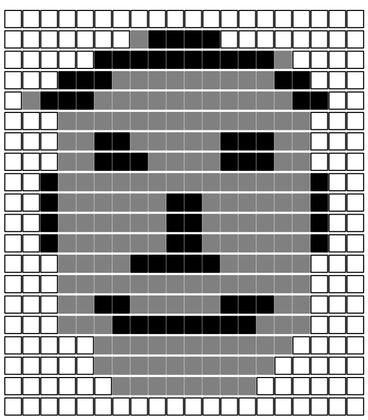
Bei der Rang-4-Approximation werden insgesamt nur sehr wenige Rasterpunkte nicht korrekt dargestellt, nämlich zum Beispiel die Einträge:

$1{,}4$ an der Stelle $(2, 8)$ (wo wir auf 1 abrunden aber die Farbe schwarz sein sollte);

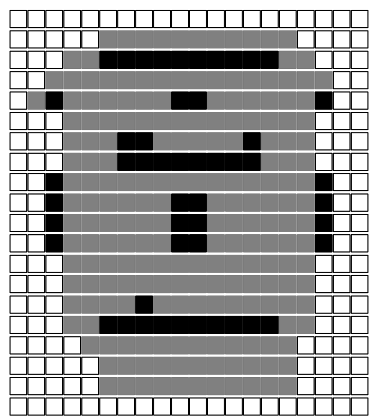
$1{,}6$ an der Stelle $(7, 6)$ (wo wir auf 2 aufrunden aber die Farbe grau sein sollte);

Tabelle 1. Approximation an A

$$\begin{bmatrix}
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & -0.2 & 0.0 & 0.0 & 0.1 & -0.2 & 0.1 & 1.4 & 2.0 & 2.1 & 2.1 & 2.0 & 0.3 & 0.1 & -0.1 & 0.2 & 0.0 & 0.0 & 0.0 & 0.0 \\
0.0 & -0.3 & 0.0 & 0.2 & 0.4 & 1.6 & 2.3 & 2.2 & 1.9 & 2.0 & 2.0 & 1.9 & 2.1 & 2.3 & 1.8 & 1.1 & 0.2 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.4 & 0.0 & 2.1 & 1.9 & 1.6 & 1.2 & 1.0 & 1.1 & 0.9 & 0.9 & 1.1 & 1.1 & 1.2 & 1.1 & 1.6 & 2.1 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.6 & 2.0 & 1.9 & 1.5 & 1.2 & 0.9 & 0.7 & 0.8 & 1.3 & 1.3 & 0.8 & 0.9 & 0.9 & 1.0 & 1.3 & 1.9 & 2.0 & 0.0 & 0.0 \\
0.0 & 0.1 & 0.1 & 1.0 & 1.0 & 1.1 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.9 & 1.0 & 1.0 & 0.1 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.8 & 0.9 & 1.6 & 1.8 & 1.4 & 1.1 & 1.1 & 1.1 & 1.1 & 1.6 & 1.8 & 1.5 & 1.2 & 0.8 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.9 & 1.0 & 1.9 & 2.1 & 1.5 & 1.1 & 1.0 & 1.0 & 1.1 & 1.9 & 2.1 & 1.9 & 1.4 & 0.9 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.4 & 1.9 & 1.0 & 0.9 & 1.0 & 1.0 & 0.9 & 0.7 & 1.3 & 1.3 & 0.7 & 1.0 & 1.0 & 1.0 & 0.9 & 1.0 & 1.9 & 0.0 & 0.0 \\
0.0 & 0.4 & 2.0 & 1.0 & 0.9 & 0.9 & 1.0 & 1.1 & 1.2 & 1.8 & 1.8 & 1.2 & 1.0 & 1.0 & 1.0 & 0.9 & 1.0 & 2.0 & 0.0 & 0.0 \\
0.0 & 0.4 & 2.0 & 1.0 & 0.9 & 0.9 & 1.0 & 1.1 & 1.2 & 1.8 & 1.8 & 1.2 & 1.0 & 1.0 & 1.0 & 0.9 & 1.0 & 2.0 & 0.0 & 0.0 \\
0.0 & 0.4 & 2.0 & 1.0 & 0.9 & 0.9 & 1.0 & 1.1 & 1.2 & 1.8 & 1.8 & 1.2 & 1.0 & 1.0 & 1.0 & 0.9 & 1.0 & 2.0 & 0.0 & 0.0 \\
0.0 & 0.1 & 0.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.6 & 1.9 & 1.9 & 1.9 & 1.9 & 1.0 & 1.0 & 0.8 & 1.0 & 1.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.1 & 0.1 & 1.0 & 1.0 & 1.1 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 0.9 & 1.0 & 1.0 & 0.1 & 0.0 & 0.0 \\
0.0 & 0.1 & 0.0 & 1.0 & 1.0 & 1.7 & 1.9 & 1.3 & 0.9 & 0.9 & 0.9 & 0.9 & 1.7 & 1.9 & 1.6 & 1.3 & 1.0 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.9 & 1.0 & 1.5 & 1.7 & 2.0 & 2.0 & 2.0 & 2.0 & 2.0 & 1.7 & 1.7 & 1.4 & 1.3 & 0.9 & 0.0 & 0.0 & 0.0 \\
0.0 & -0.1 & 0.0 & 0.4 & 0.4 & 0.9 & 1.1 & 1.1 & 1.0 & 1.0 & 1.0 & 1.0 & 1.0 & 1.1 & 0.9 & 0.7 & 0.4 & 0.0 & 0.0 & 0.0 \\
0.0 & -0.2 & 0.0 & -0.1 & 0.1 & 0.7 & 1.1 & 1.1 & 0.9 & 1.0 & 1.0 & 0.9 & 1.0 & 1.1 & 0.8 & 0.5 & -0.1 & 0.0 & 0.0 & 0.0 \\
0.0 & -0.2 & 0.0 & -0.1 & -0.1 & 0.4 & 0.7 & 1.1 & 1.0 & 1.1 & 1.1 & 1.0 & 0.7 & 0.7 & 0.5 & 0.3 & -0.1 & 0.0 & 0.0 & 0.0 \\
0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0
\end{bmatrix}$$



Rang-4-Approximation $A_{(4)}$



Rang-2-Approximation $A_{(2)}$

$0{,}4$ an der Stelle $(17, 5)$ (wo wir auf 0 abrunden aber die Farbe grau sein sollte).

Nun beachte, dass $A_{(4)}$ gegeben ist durch zwei Matrizen der Größe 20×4, plus die 4 Singulärwerte, also brauchen wir für $A_{(4)}$ nur noch $164 = 2 \cdot 80 + 4$ Zahlen zu speichern — deutlich weniger als die Hälfte der ursprünglichen Anzahl! Für die Rang-2-Approximation bräuchten wir noch viel weniger Speicherplatz, aber die Qualität des Bildes ist dagegen ziemlich schlecht; hier bleibt im Wesentlichen nur die ungefähre Form erhalten.

Experimentieren Sie selbst mit der `SVD`–Funktion, um zu sehen, wie sich die Qualität des Bildes für unterschiedliche Werte von $k \in \{1, 2, \ldots, 12\}$ ändert.

Zu den vielen weiteren Anwendungen des Spektralsatzes und der Singulärwertzerlegung erwähnen wir hier nur noch das ***Moore–Penrose Inverse*** einer (nicht-invertierbaren oder nicht-quadratischen) Matrix; siehe https://de.wikipedia.org/wiki/Pseudoinverse.

Kapitel VI: Folgen und Reihen

In diesem und den folgenden Kapiteln geht es um reelle Funktionen der Form $f\colon D \to \mathbb{R}$, wobei typischerweise D ein Intervall oder ganz $\mathbb{R}$ ist. Einige Beispiele für solche Funktionen, die Sie vermutlich schon in der einen oder anderen Form gesehen haben, sind:

- Polynomfunktionen $f(x) = a_n x^n + \ldots + a_1 x + a_0$ auf $D = \mathbb{R}$ (wie im ersten Semester);
- Wurzelfunktionen $f(x) = \sqrt[n]{x}$ auf $D = \{x \in \mathbb{R} \mid x \geqslant 0\}$;
- die trigonometrischen Funktionen $\sin(x)$ und $\cos(x)$ auf $D = [0, 2\pi]$;
- die Exponential-Funktion $\exp(x) = e^x$ auf $D = \mathbb{R}$, wobei e die Eulersche Zahl ist.

Mit Ausnahme der Polynomfunktionen gibt es für die anderen Funktionen keine elementaren Beschreibungen, sondern deren Existenz beruht wesentlich auf der Möglichkeit, Grenzwerte bilden zu können. (Das Gleiche trifft schon auf e oder die Kreiszahl π zu.) Wir behandeln in diesem Kapitel die Grundlagen dieser Grenzwertbildung; ebenso auch die Frage, wie man die damit zusammenhängenden unendlichen Prozesse rechnerisch unter Kontrolle bringen kann.

27. *Grenzwerte von Folgen*

Zur Erinnerung: Ist eine Funktion $f\colon \mathbb{N}_0 \to \mathbb{R}$ gegeben, so setzen wir $a_n := f(n)$ für alle $n \in \mathbb{N}_0$ und schreiben die Funktion einfach als $f = (a_n)_{n \in \mathbb{N}_0}$; dies wird dann auch als Folge von reellen Zahlen bezeichnet. Man interessiert sich nun dafür, wie sich die Folgenglieder a_n bei immer größer werdendem Index n verhalten.

Betrachte zum Beispiel die Zahlenfolge $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$, also $a_n = \frac{1}{n+1}$ für alle $n \in \mathbb{N}_0$. Offenbar werden die Folgenglieder mit wachsendem n immer kleiner und nähern sich der Zahl 0 an (ohne 0 jemals exakt zu erreichen). Dies ist der Prototyp einer "konvergenten" Folge. Hier ist die präzise mathematische Formulierung dieser Idee:

Definition 27.1. Sei $(a_n)_{n \in \mathbb{N}_0}$ eine Folge von reellen Zahlen und $\alpha \in \mathbb{R}$. Die Folge $(a_n)_{n \in \mathbb{N}_0}$ heißt ***konvergent***, mit ***Grenzwert*** α, wenn folgende Bedingung gilt:

$(*)$ *Zu jedem* $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, *gibt es ein* $n_0 \in \mathbb{N}_0$ *mit* $|a_n - \alpha| < \varepsilon$ *für alle* $n \geqslant n_0$.

(Anders formuliert: $\forall \varepsilon > 0\ \exists\, n_0 \in \mathbb{N}_0\ \forall n \geqslant n_0 :\ |a_n - \alpha| < \varepsilon$.) In diesem Fall schreiben wir $\lim\limits_{n\to\infty}(a_n) = \alpha$, oder auch manchmal $(a_n) \to \alpha$ (für $n \to \infty$). Gibt es kein α mit $(*)$, so heißt die Folge $(a_n)_{n \in \mathbb{N}_0}$ ***divergent***.

Anschaulich gesprochen bedeutet obige Bedingung $(*)$: Egal wie klein man $\varepsilon > 0$ wählt, es gibt immer einen Index n_0 (der von ε abhängen kann), so dass alle Folgenglieder a_n ab diesem Index höchstens den Abstand ε von α haben. Oder noch einmal anders: Für größer werdendes n nähern sich die Folgenglieder a_n immer mehr dem Wert α an. — Dennoch ist die obige Definition sicherlich gewöhnungsbedürftig. Die folgenden Beispiele und Aussagen illustrieren, dass man damit jedenfalls präzise arbeiten und Grenzwerte bestimmen kann.

Beispiel 27.2. (a) Sei $c \in \mathbb{R}$ fest und $a_n := c$ für alle $n \in \mathbb{N}_0$. Dann konvergiert $(a_n)_{n\in\mathbb{N}_0}$ mit Grenzwert c. Denn sei $\varepsilon > 0$. Dann gilt $|a_n - c| = 0 < \varepsilon$ für alle $n \in \mathbb{N}_0$; die Bedingung $(*)$ ist also mit $n_0 = 0$ erfüllt.

(b) Die Folge $(a_n)_{n\in\mathbb{N}_0}$ heißt ***Null-Folge***, wenn sie gegen $\alpha = 0$ konvergiert. Sei zum Beispiel $c \in \mathbb{R}$ fest und $a_n := \frac{c}{n+1}$ für alle $n \in \mathbb{N}_0$. Behauptung: Dies ist eine Null-Folge. Denn sei $\varepsilon > 0$. Nach Lemma 10.4 (Kapitel II) gibt es ein $n_0 \in \mathbb{N}$ mit $n_0\varepsilon > |c|$. Ist $n \geqslant n_0$, so folgt auch $(n+1)\varepsilon > n\varepsilon \geqslant n_0\varepsilon > |c|$ und damit $|a_n| = \frac{|c|}{n+1} < \varepsilon$, wie gewünscht.

(c) Sei $a_n := (-1)^n$ für alle $n \in \mathbb{N}_0$. Diese Folge ist divergent. Annahme, sie wäre konvergent, mit einem Grenzwert α. Sei $\varepsilon = \frac{1}{2} > 0$. Dann gibt es ein $n_0 \in \mathbb{N}_0$ mit $|a_n - \alpha| < \frac{1}{2}$ für alle $n \geqslant n_0$. Daraus folgt $-\frac{1}{2} < \alpha - a_n < \frac{1}{2}$, also $-\frac{1}{2} + a_n < \alpha < \frac{1}{2} + a_n$ für alle $n \geqslant n_0$. Nun sei $n \geqslant n_0$ gerade. Dann ist $a_n = (-1)^n = 1$ und die obige Ungleichung zeigt $\frac{1}{2} < \alpha < \frac{3}{2}$. Ist $n \geqslant n_0$ ungerade, so ist $a_n = (-1)^n = -1$ und die obige Ungleichung zeigt $-\frac{3}{2} < \alpha < -\frac{1}{2}$. Also folgt $\alpha < 0$ und $\alpha > 0$, Widerspruch.

(d) Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge und $d \in \mathbb{N}$ fest. Dann definiere die Folge $(b_n)_{n\in\mathbb{N}_0}$ mit $b_n := a_{d+n}$ für alle $n \in \mathbb{N}_0$. (Wir streichen also die ersten d Folgenglieder $a_0, a_1, \ldots, a_{d-1}$.) Man sieht sofort: Genau dann ist $(a_n)_{n\in\mathbb{N}_0}$ konvergent, wenn $(b_n)_{n\in\mathbb{N}_0}$ konvergent ist; und in diesem Fall sind die Grenzwerte gleich.

Bemerkung 27.3. Sei $d \in \mathbb{N}$ und $D := \{n \in \mathbb{N}_0 \mid n \geqslant d\}$. Ist $f\colon D \to \mathbb{R}$ eine Funktion, so setzen wir $a_n := f(n)$ für $n \geqslant d$, erhalten also eine Folge $(a_n)_{n\geqslant d}$. Solche Situationen, also Folgen, die erst ab einem bestimmten Index d definiert sind, werden wir öfter in Beispielen sehen. Wir können aber stets zur Folge $(\tilde{a}_n)_{n\in\mathbb{N}_0}$ übergehen, mit $\tilde{a}_n := a_n$ für $n \geqslant d$, und $\tilde{a}_n := 0$ für $n = 0, 1, \ldots, d-1$. Nach Beispiel 27.2(d) ändert dies nichts am Grenzwertverhalten für $n \to \infty$. Die allgemeinen Sätze für Folgen mit Definitionsbereich $\mathbb{N}_0$ behalten also auch ihre Gültigkeit für Folgen, die erst ab einem bestimmten Index d definiert sind.

Lemma 27.4. *Sei $(a_n)_{n\in\mathbb{N}_0}$ eine konvergente Folge. Dann gilt:*
(a) *Der Grenzwert α der Folge ist eindeutig bestimmt.*
(b) *Die Menge $\{a_n \mid n \in \mathbb{N}_0\} \subseteq \mathbb{R}$ ist nach oben und nach unten beschränkt.*

Beweis. (a) Nehmen wir an, es gibt $\alpha, \alpha' \in \mathbb{R}$ mit $\alpha < \alpha'$, so dass die Bedingung $(*)$ sowohl mit α als auch mit α' gilt. Sei $\varepsilon := (\alpha' - \alpha)/2 > 0$. Dann gibt es ein $n_1 \in \mathbb{N}_0$ mit $|a_n - \alpha| < \varepsilon$ für alle $n \geqslant n_1$, und es gibt ein $n_2 \in \mathbb{N}_0$ mit $|a_n - \alpha'| < \varepsilon$ für alle $n \geqslant n_2$. Nun sei $n \geqslant \max\{n_1, n_2\}$. Dann ist $-\varepsilon < a_n - \alpha < \varepsilon$, also $a_n < \alpha + \varepsilon$. Es gilt auch $-\varepsilon < a_n - \alpha' < \varepsilon$, also $\alpha' - \varepsilon < a_n$. Zusammengenommen erhalten wir $a_n < \alpha + \varepsilon = \alpha + (\alpha' - \alpha)/2 = (\alpha + \alpha')/2 = \alpha' - (\alpha' - \alpha)/2 = \alpha' - \varepsilon < a_n$, Widerspruch.

(b) Sei $\alpha \in \mathbb{R}$ Grenzwert der Folge. Zu $\varepsilon = 1$ gibt es ein $n_0 \in \mathbb{N}_0$ mit $|a_n - \alpha| < 1$ für alle $n \geqslant n_0$, also $\alpha - 1 < a_n < \alpha + 1$ für alle $n \geqslant n_0$. Sei $C := \max\{\alpha + 1, a_0, a_1, \ldots, a_{n_0-1}\} \in \mathbb{R}$; dann gilt $a_n \leqslant C$ für alle $n \in \mathbb{N}$. Also ist C eine obere Schranke für $\{a_n \mid n \in \mathbb{N}_0\}$. Analog ist $\min\{\alpha - 1, a_0, a_1, \ldots, a_{n_0-1}\} \in \mathbb{R}$ eine untere Schranke. □

Lemma 27.5 (Cauchy)**.** *Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge reeller Zahlen. Ist diese Folge konvergent, so gibt es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}_0$ mit $|a_n - a_m| < \varepsilon$ für alle $n \geqslant m \geqslant n_0$.*

Beweis. Sei $(a_n)_{n\in\mathbb{N}_0}$ konvergent, mit Grenzwert $\alpha \in \mathbb{R}$. Zu $\varepsilon > 0$ gibt es dann ein $n_0 \in \mathbb{N}_0$ mit $|a_n - \alpha| < \varepsilon/2$ für alle $n \geqslant n_0$. Sei $n \geqslant m \geqslant n_0$. Mit der Dreiecksungleichung folgt $|a_n - a_m| = |a_n - \alpha + \alpha - a_m| \leqslant |a_n - \alpha| + |a_m - \alpha| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. □

Es gilt auch die Umkehrung der obigen Aussage, aber der Beweis ist komplizierter; siehe zum Beispiel §5, Satz 3, im 1. Band des Buches von Forster. (Das Problem ist, dass man erstmal einen Kandidaten für einen Grenzwert benötigt, um überhaupt die Bedingung (∗) in Definition 27.1 testen zu können.)

Satz 27.6 (***Einschließungs-Kriterium***)**.** *Gegeben seien Folgen $(a_n)_{n\in\mathbb{N}_0}$, $(b_n)_{n\in\mathbb{N}_0}$ und $(c_n)_{n\in\mathbb{N}_0}$. Es gebe ein $N \in \mathbb{N}_0$ mit $a_n \leqslant c_n \leqslant b_n$ für alle $n \geqslant N$. Sind $(a_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$ konvergent mit gleichem Grenzwert $\alpha \in \mathbb{R}$, so ist auch $(c_n)_{n\in\mathbb{N}_0}$ konvergent mit Grenzwert α.*

Beweis. Sind $x, y, z \in \mathbb{R}$ gegeben mit $x \leqslant y \leqslant z$, so gilt $-|x| \leqslant x \leqslant y \leqslant z \leqslant |z|$ und damit $|y| \leqslant \max\{|x|, |z|\}$. Nach Voraussetzung ist nun $a_n \leqslant c_n \leqslant b_n$ für alle $n \geqslant N$, also auch $a_n - \alpha \leqslant b_n - \alpha \leqslant c_n - \alpha$ und damit

$$|b_n - \alpha| \leqslant \max\{|a_n - \alpha|, |c_n - \alpha|\} \qquad \text{für alle } n \geqslant N.$$

Sei $\varepsilon > 0$ beliebig. Dann gibt es nach Voraussetzung ein $n_1 \in \mathbb{N}_0$ mit $|a_n - \alpha| < \varepsilon$ für alle $n \geqslant n_1$, und ein $n_2 \in \mathbb{N}_0$ mit $|c_n - \alpha| < \varepsilon$ für alle $n \geqslant n_2$. Sei $n_0 := \max\{n_1, n_2\}$. Für $n \geqslant n_0$ ist dann $|a_n - \alpha| < \varepsilon$ und $|c_n - \alpha| < \varepsilon$, also auch $|b_n - \alpha| \leqslant \max\{|a_n - \alpha|, |c_n - \alpha|\} < \varepsilon$. □

Beispiel 27.7. Sei $c \in \mathbb{R}$ fest und $a_n := c^n$ für $n \in \mathbb{N}$. Ist $|c| < 1$, so ist $(a_n)_{n\in\mathbb{N}}$ eine Null-Folge. Denn: Ist $c = 0$, so ist dies klar. Sei nun $c \neq 0$, $|c| < 1$ und setze $x := \frac{1-|c|}{|c|} \in \mathbb{R}$; wegen $|c| < 1$ ist $x > 0$. Nun ist $|c| = \frac{1}{1+x}$, also folgt mit der Bernoulli–Ungleichung

$$|c|^n = \frac{1}{(1+x)^n} \leqslant \frac{1}{1+nx} \leqslant \frac{1}{nx} \qquad \text{für alle } n \in \mathbb{N}.$$

Damit gilt $0 \leqslant |c|^n \leqslant c_n$ mit $c_n := \frac{1}{nx}$ für alle $n \in \mathbb{N}$. Die konstante Folge $(0)_{n\in\mathbb{N}}$ konvergiert mit Grenzwert 0, die Folge $(c_n)_{n\in\mathbb{N}}$ ist eine Null–Folge; siehe Beispiel 27.2. Mit dem Einschließungs-Kriterium folgt, dass die Folge $(|c|^n)_{n\in\mathbb{N}}$ konvergiert, also auch $(a_n)_{n\in\mathbb{N}}$, mit dem gleichen Grenzwert 0.

Satz 27.8 (***Grenzwert-Regeln***)**.** *Gegeben seien konvergente Folgen $(a_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$. Sei $\alpha := \lim\limits_{n\to\infty}(a_n)$ und $\beta := \lim\limits_{n\to\infty}(b_n)$.*

(1) *Seien* $r, s \in \mathbb{R}$ *fest und* $(c_n)_{n\in\mathbb{N}_0}$ *die Folge mit* $c_n := ra_n + sb_n$. *Dann konvergiert* $(c_n)_{n\in\mathbb{N}_0}$ *mit Grenzwert* $r\alpha + s\beta$.
(2) *Sei* $(c_n)_{n\in\mathbb{N}_0}$ *die Folge mit* $c_n := a_n \cdot b_n$. *Dann konvergiert* $(c_n)_{n\in\mathbb{N}_0}$ *mit Grenzwert* $\alpha \cdot \beta$.
(3) *Sei* $\beta \neq 0$. *Dann gibt es ein* $n_0 \in \mathbb{N}_0$ *mit* $b_n \neq 0$ *für alle* $n \geqslant n_0$. *Sei* $(c_n)_{n\in\mathbb{N}_0}$ *die Folge mit* $c_n := a_n/b_n$ *für* $n \geqslant n_0$, *und* $c_n := 0$ *sonst. Dann konvergiert* $(c_n)_{n\in\mathbb{N}_0}$ *mit Grenzwert* α/β.
(4) *Gibt es ein* $n_0 \in \mathbb{N}_0$ *mit* $a_n \leqslant b_n$ *für alle* $n \geqslant n_0$, *so ist auch* $\alpha \leqslant \beta$.

Beweis. Zuerst (4): Annahme, es wäre $\alpha > \beta$. Sei dann $\varepsilon := \alpha - \beta > 0$. Nach Voraussetzung gibt es ein $n_1 \in \mathbb{N}_0$ mit $|a_n - \alpha| < \varepsilon/2$ für alle $n \geqslant n_1$, sowie ein $n_2 \in \mathbb{N}_0$ mit $|b_n - \beta| < \varepsilon/2$ für alle $n \geqslant n_2$. Dann ist $-\varepsilon/2 < a_n - \alpha < \varepsilon/2$, also $a_n > \alpha - \varepsilon/2$ für $n \geqslant n_1$; analog $-\varepsilon/2 < b_n - \beta < \varepsilon/2$, also $\beta > b_n - \varepsilon/2$ für $n \geqslant n_2$. Sei nun $n \geqslant \max\{n_1, n_2\}$. Wegen $\alpha = \beta + \varepsilon$ folgt dann $a_n > \alpha - \varepsilon/2 = \beta + \varepsilon - \varepsilon/2 > (b_n - \varepsilon/2) + \varepsilon - \varepsilon/2 = b_n$, Widerspruch. Die Aussagen (1), (2), (3) werden auf ähnliche Weise gezeigt; siehe etwa §4.1.2 im Buch von Glosauer für die Details. □

Bemerkung 27.9. Wir werden später sehen, dass es analoge Regeln auch für das Zusammenspiel von Grenzwerten mit geeigneten Funktionen $f\colon D \to \mathbb{R}$ gibt (wie etwa sin, cos, Polynomfunktionen, Wurzelfunktionen $\sqrt[n]{x}$ für $x \geqslant 0$, ..., oder Kombinationen von solchen Funktionen). D.h., ist $(a_n)_{n\in\mathbb{N}_0}$ konvergent mit $a_n \in D$ für alle $n \in \mathbb{N}_0$ und $\alpha := \lim_{n\to\infty}(a_n)$, so wird auch die Folge $(f(a_n))_{n\in\mathbb{N}_0}$ konvergent sein, mit $f(\alpha) = \lim_{n\to\infty}(f(a_n))$.

Beispiel 27.10. Seien $a, b \in \mathbb{R}$ und $(a_n)_{n\in\mathbb{N}_0}$ eine Null-Folge. Es gebe ein $n_0 \in \mathbb{N}_0$ mit $a \leqslant a_n + b$ für alle $n \geqslant n_0$. Dann ist $a \leqslant b$. Dazu: Setze $b_n := a_n + b$ für $n \in \mathbb{N}_0$. Nach Beispiel 27.2(a) und Satz 27.8(1) konvergiert $(b_n)_{n\in\mathbb{N}_0}$, mit Grenzwert $0 + b = b$. Nun ist $a \leqslant b_n$ für alle $n \geqslant n_0$. Nach Satz 27.8(4) folgt also auch $a \leqslant b$.

Beispiel 27.11. (a) Sei $a_n := \dfrac{n^2+2n+3}{4n^2+5n+6}$ für $n \in \mathbb{N}$. (Der Nenner ist stets $\neq 0$, der Bruch also definiert.) Berechne $\lim_{n\to\infty}(a_n)$ (falls der Grenzwert existiert).

Solche Beispiele lassen sich mit folgendem Trick behandeln. Wir teilen Zähler und Nenner durch die höchste vorkommende Potenz von n, in diesem Fall also n^2. Dies ergibt

$$\frac{n^2+2n+3}{4n^2+5n+6} = \frac{1+\frac{2}{n}+\frac{3}{n^2}}{4+\frac{5}{n}+\frac{6}{n^2}} \longrightarrow \frac{1+0+0}{4+0+0} = \frac{1}{4} \qquad (\text{für } n \to \infty),$$

wobei wir Beispiel 27.2 und die Regeln in Satz 27.8 benutzen. Also ist die Folge $(a_n)_{n\in\mathbb{N}}$ konvergent, mit Grenzwert $\alpha = \frac{1}{4}$.

(b) Sei $r \in \mathbb{R}$, $r > 0$, fest. Betrachte die rekursiv definierte Folge $(a_n)_{n\in\mathbb{N}}$ mit:

$$a_1 = \max\{r, 1\} \qquad \text{und} \qquad a_{n+1} = \frac{1}{2}\Big(a_n + \frac{r}{a_n}\Big) \qquad \text{für } n \geqslant 1.$$

Beachte: Man sieht sofort, dass $a_n > 0$ für alle $n \in \mathbb{N}$ gilt, der Bruch r/a_n ist also in jedem Fall definiert. Nehmen wir an, wir wüssten bereits, dass diese Folge konvergiert; sei $\alpha \in \mathbb{R}$ der Grenzwert. Nun gilt $a_1 \geqslant \sqrt{r}$ und

$$a_{n+1} - \sqrt{r} = \frac{1}{2}\Big(a_n + \frac{r}{a_n}\Big) - \sqrt{r} = \frac{a_n^2 - 2\sqrt{r}\,a_n + r}{2a_n} = \frac{(a_n - \sqrt{r})^2}{2a_n} \geqslant 0$$

für alle $n \geqslant 1$. Also gilt $a_n \geqslant \sqrt{r}$ für alle $n \in \mathbb{N}$, und damit auch $\alpha \geqslant \sqrt{r} > 0$; siehe Satz 27.8(4). Nach Beispiel 27.2(d) ist die Folge $(a_{n+1})_{n\in\mathbb{N}}$ ebenfalls konvergent, mit Grenzwert α. Mit den weiteren Regeln in Satz 27.8 folgt nun

$$\alpha = \lim_{n\to\infty}(a_{n+1}) = \lim_{n\to\infty} \frac{1}{2}\Big(a_n + \frac{r}{a_n}\Big) = \frac{1}{2}\Big(\alpha + \frac{r}{\alpha}\Big)$$

und damit $2\alpha^2 = \alpha^2 + r$, d.h., $\alpha^2 = r$. Wegen $\alpha > 0$ erhalten wir also $\alpha = \sqrt{r}$. — Wir müssen uns aber noch davon überzeugen, dass $(a_n)_{n\in\mathbb{N}}$ überhaupt konvergiert; siehe weiter unten.

(c) Betrachte die rekursiv definierte Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_1 := 1$ und $a_{n+1} := 2a_n + 3$ für $n \geqslant 1$. Nehmen wir wieder an, dass diese Folge konvergiert; sei $\alpha \in \mathbb{R}$ der Grenzwert. Wie oben erhalten wir dann $\alpha = 2\alpha + 3$, also $\alpha = -3$; Widerspruch, weil $a_n > 0$ für alle $n \in \mathbb{N}$ gilt, und damit $\alpha \geqslant 0$. — Also war die Annahme falsch: Die Folge konvergiert gar nicht!

Der anschließende Satz ist ein sehr nützliches Hilfsmittel, um zu zeigen, dass eine Folge konvergiert (ohne irgendetwas über den Grenzwert zu wissen). Im Beweis geht wesentlich ein, dass $\mathbb{R}$ ein ***vollständiger Körper*** ist; siehe §10, Kapitel II. Vorab noch eine allgemeine Definition, die nicht nur für Folgen gültig ist, sondern beliebige reelle Funktionen.

Definition 27.12 (Monotonie von Funktionen)**.** Sei $f\colon D \to \mathbb{R}$ eine Funktion mit $D \subseteq \mathbb{R}$.

$$f \text{ heißt } \left\{\begin{array}{c} \text{monoton wachsend} \\ \text{streng monoton wachsend} \\ \text{monoton fallend} \\ \text{streng monoton fallend} \end{array}\right\}, \quad \text{falls} \quad \left\{\begin{array}{c} f(x) \leqslant f(y) \\ f(x) < f(y) \\ f(x) \geqslant f(y) \\ f(x) > f(y) \end{array}\right\}$$

für alle $x, y \in D$ mit $x < y$ gilt.

Beachte: Ist f streng monoton wachsend oder streng monoton fallend, so ist f injektiv.

Einfaches Beispiel: $f(x) = x^2$ ist streng monoton wachsend auf $D = \{x \in \mathbb{R} \mid x \geqslant 0\}$ und streng monoton fallend auf $D = \{x \in \mathbb{R} \mid x \leqslant 0\}$; aber $f(x) = x^2$ ist weder monoton wachsend noch monoton fallend auf $D = \mathbb{R}$.

Satz 27.13 (***Monotonie-Prinzip***)**.** *Eine monoton wachsende und nach oben beschränkte Folge ist konvergent. Genauer: Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge von reellen Zahlen. Gibt es ein $C \in \mathbb{R}$ und ein $N \in \mathbb{N}_0$ mit $a_n \leqslant C$ und $a_n \leqslant a_{n+1}$ für alle $n \geqslant N$, so ist die Folge konvergent und es gilt $\lim_{n\to\infty}(a_n) = \sup\{a_n \mid n \geqslant N\} \leqslant C$.*

Beweis. Nach Voraussetzung ist C eine obere Schranke für die Menge $S := \{a_n \mid n \geqslant N\} \subseteq \mathbb{R}$. Also existiert $\alpha := \sup(S) \in \mathbb{R}$ und es gilt $\alpha \leqslant C$. Sei nun $\varepsilon > 0$ beliebig. Dann ist $\alpha - \varepsilon$ keine obere Schranke für S, also gibt es ein $n_0 \in \mathbb{N}_0$ mit $n_0 \geqslant N$ und $a_{n_0} > \alpha - \varepsilon$. Wegen $a_{n_0} \leqslant \alpha$ ist also $0 \leqslant \alpha - a_{n_0} < \varepsilon$. Sei nun $n \geqslant n_0 \geqslant N$ beliebig. Nach Voraussetzung ist $a_n \geqslant a_{n_0}$; außerdem ist $a_n \in S$, also $a_n \leqslant \alpha$ und $0 \leqslant \alpha - a_n \leqslant \alpha - a_{n_0} < \varepsilon$. Damit gilt $|a_n - \alpha| < \varepsilon$ für alle $n \geqslant n_0$, also ist $\lim\limits_{n\to\infty}(a_n) = \alpha$. □

Analog gibt es auch eine Version für monoton fallende Folgen, also: *Sei* $(a_n)_{n\in\mathbb{N}_0}$ *eine Folge von reellen Zahlen. Gibt es ein* $C \in \mathbb{R}$ *und ein* $N \in \mathbb{N}_0$ *mit* $a_n \geqslant C$ *und* $a_n \geqslant a_{n+1}$ *für alle* $n \geqslant N$*, so ist die Folge konvergent und es gilt* $\lim\limits_{n\to\infty}(a_n) = \inf\{a_n \mid n \geqslant N\} \geqslant C$.

Ab hier Woche 4

Beispiel 27.14. Gegeben sei eine Folge von abgeschlossenen Intervallen $I_1, I_2, I_3, \ldots \subseteq \mathbb{R}$ mit $I_1 \supseteq I_2 \supseteq I_3 \supseteq \ldots$ und $\inf(\{\ell(I_n) \mid n \in \mathbb{N}\}) = 0$, wie in Satz 10.9, Kapitel II (Intervallschachtelung). Es gibt dann genau ein $x \in \mathbb{R}$ mit $x \in I_n$ für alle $n \in \mathbb{N}$. Schreibe $I_n = [a_n, b_n]$ mit $a_n, b_n \in \mathbb{R}$, $a_n < b_n$. Dann ist $a_n \leqslant a_{n+1}$ und $b_n \geqslant b_{n+1}$ für alle $n \in \mathbb{N}$. In Satz 10.9 wurde gezeigt, dass $\{a_n \mid n \in \mathbb{N}\}$ nach oben beschränkt ist und $x := \sup(\{a_n \mid n \in \mathbb{N}\}) \in \mathbb{R}$ gilt. Kombiniert mit Satz 27.13 erhalten wir nun, dass die Folge $(a_n)_{n\in\mathbb{N}}$ konvergiert, mit Grenzwert x. Analog konvergiert die Folge $(b_n)_{n\in\mathbb{N}}$, ebenfalls mit Grenzwert x.
Beispiel: In §10, Kapitel II, haben wir die ***Eulersche Zahl*** $e \in \mathbb{R}$ mit Hilfe einer Intervallschachtelung eingeführt. Es folgt nun: $e = \lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1}$.

Beispiel 27.15. Betrachten wir noch einmal die rekursiv definierte Folge $(a_n)_{n\in\mathbb{N}}$ in Beispiel 27.11(b). Wir haben dort bereits gesehen, dass $a_n \geqslant \sqrt{r}$ für alle $n \in \mathbb{N}$ gilt, die Folge ist also nach unten beschränkt. Außerdem gilt:

$$a_n - a_{n+1} = a_n - \frac{1}{2}\Big(a_n + \frac{r}{a_n}\Big) = \frac{2a_n^2 - (a_n^2 + r)}{2a_n} = \frac{a_n^2 - r}{2a_n} \geqslant 0$$

für alle $n \in \mathbb{N}$, d.h., die Folge ist monoton fallend. Mit Satz 27.13 können wir jetzt schließen, dass $(a_n)_{n\in\mathbb{N}}$ konvergiert, und mit der Rechnung in Beispiel 27.11(b) folgt $\lim\limits_{n\to\infty}(a_n) = \sqrt{r}$.

Wir haben damit erstens gezeigt, dass jedes $r \in \mathbb{R}$ mit $r > 0$ eine Quadratwurzel in $\mathbb{R}$ besitzt; zweitens können wir $\sqrt{r}$ näherungsweise bestimmen, indem wir die Folgenglieder bis zu einem bestimmten Index ausrechnen. Für $r = 2$ erhalten wir zum Beispiel mit `Sage`:

```
sage: def sqrt(r,n):      # n = Anzahl der Iterationen
...:     a1=max(1,r)
...:     for i in range(n-1): a1=(a1+r/a1)/2
...:  return a1
sage: sqrt(2,2)           # r = 2, n = 2
3/2
sage: sqrt(2,6)           # r = 2, n = 6
886731088897/627013566048
```

```
sage: dec_exp(_)          # '_' = was zuletzt berechnet wurde
'1.4142135623730950'      # Wurzel(2) bereits korrekt auf 16 Stellen!
```

Ist die Eingabe von `sqrt` ein $r \in \mathbb{Q}$, so erfolgen die Rechnungen innerhalb der Funktion ebenfalls in $\mathbb{Q}$ (also exakt, nicht näherungsweise). Um das Ergebnis nicht als Bruch, sondern als Dezimalzahl zu sehen, benutzen wir am Ende die Funktion `dec_exp` in Tabelle 2. Diese implementiert die Prozedur in Beispiel 10.7, Kapitel II, und druckt die Dezimalentwicklung von $x \in \mathbb{Q}$ bis zu einer vorgegebenen Anzahl von Nachkommastellen aus.

Tabelle 2. Sage–Programm zur Dezimalentwicklung von $x \in \mathbb{Q}$

```
def dec_exp(x,nachkomma=16):            # nachkomma = Anzahl korrekter
  if x==0: return '0.0'                 # Nachkommastellen (default=16)
  if x<0: return '-'+dec_exp(-x,max)
  z=0
  while z+1<=x: z+=1
  d=str(z)+'.'; y=10*(x-z)
  for i in range(nachkomma):
    z=0
    while z+1<=y: z+=1
    d+=str(z); y=10*(y-z)
  return d                              # Ausgabe ist ein string
```

Definition 27.16. Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge reeller Zahlen. Dann heißt diese Folge ***bestimmt divergent*** gegen $+\infty$, wenn es zu jedem $C \in \mathbb{R}$ ein $n_0 \in \mathbb{N}_0$ gibt mit $a_n > C$ für alle $n \geqslant n_0$. Wir schreiben dann auch $\lim\limits_{n\to\infty}(a_n) = +\infty$.
Analog heißt die Folge ***bestimmt divergent*** gegen $-\infty$, wenn es zu jedem $C \in \mathbb{R}$ ein $n_0 \in \mathbb{N}_0$ gibt mit $a_n < C$ für alle $n \geqslant n_0$. Wir schreiben dann auch $\lim\limits_{n\to\infty}(a_n) = -\infty$.

Zum Beispiel ist die Folge $(n)_{n\in\mathbb{N}_0}$ bestimmt divergent gegen $+\infty$; die Folge $\big((-1)^n n\big)_{n\in\mathbb{N}_0}$ ist weder konvergent noch bestimmt divergent.

Lemma 27.17. *Ist $(a_n)_{n\in\mathbb{N}_0}$ bestimmt divergent gegen $+\infty$ oder $-\infty$, so ist $(1/a_n)_{n\in\mathbb{N}_0}$ eine Null-Folge.*

Beweis. Sei zuerst $(a_n)_{n\in\mathbb{N}_0}$ bestimmt divergent gegen $+\infty$. Zu jedem $\varepsilon > 0$ gibt es dann ein $n_0 \in \mathbb{N}_0$ mit $a_n > 1/\varepsilon > 0$ für alle $n \geqslant n_0$. Also ist $1/a_n < \varepsilon$ für alle $n \geqslant n_0$. Das Argument ist analog, wenn $(a_n)_{n\in\mathbb{N}_0}$ bestimmt divergent gegen $-\infty$ ist. □

28. *Unendliche Reihen*

Betrachten wir als erstes Beispiel die Folge $1, \frac{1}{2}, \frac{1}{4}, \ldots$, also $a_n := \frac{1}{2^n}$ für $n \in \mathbb{N}_0$. Wir wollen dann so etwas wie die unendliche Summe $\sum_{n=0}^{\infty} a_n = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots$ bilden — was natürlich zunächst noch keinen Sinn ergibt. Aber man kann dies wie folgt präzisieren:

Definition 28.1. Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge reeller Zahlen. Für jedes $n \in \mathbb{N}_0$ definieren wir die ***Partialsumme***
$$A_n := \sum_{k=0}^{n} a_k = a_0 + a_1 + \ldots + a_n \in \mathbb{R}.$$
Wir können dann die Folge $(A_n)_{n\in\mathbb{N}_0}$ bilden; diese Folge heißt ***unendliche Reihe*** mit Gliedern a_n und wird mit $\sum_{n=0}^{\infty} a_n$ bezeichnet. Konvergiert die Folge $(A_n)_{n\in\mathbb{N}_0}$ (im Sinne von Definition 27.1), so wird ihr Grenzwert ebenfalls mit $\sum_{n=0}^{\infty} a_n$ bezeichnet und heißt dann Summe der Reihe. — Das Symbol $\sum_{n=0}^{\infty} a_n$ hat also zwei Bedeutungen: Erstens ist es nur eine andere Schreibweise für die Folge $(A_n)_{n\in\mathbb{N}_0}$ (egal ob diese konvergiert oder nicht), und zweitens bezeichnet es im Fall der Konvergenz den Grenzwert $\lim\limits_{n\to\infty}(A_n)$.

Beispiel 28.2. Sei $c \in \mathbb{R}$ und $a_n := c^n$ für alle $n \in \mathbb{N}_0$. Setze $A_n := \sum_{k=0}^{n} a_k$. Konvergiert die Reihe $\sum_{n=0}^{\infty} a_n$, also die Folge $(A_n)_{n\in\mathbb{N}_0}$? Ist $c = 1$, so ist $a_n = 1$ für alle n, also $A_n = n$ für alle n; die Folge $(A_n)_{n\in\mathbb{N}_0}$ ist also divergent. Nun sei $c \neq 1$. Dann gilt $A_n = (c^{n+1}-1)/(c-1)$ für alle $n \in \mathbb{N}_0$, wie man leicht mit vollständiger Induktion zeigt. Ist $|c| < 1$, so ist $(c^n)_{n\in\mathbb{N}_0}$ eine Null-Folge (siehe Beispiel 27.7). Mit Satz 27.8 folgt also, dass $(A_n)_{n\in\mathbb{N}_0}$ konvergiert, mit
$$\sum_{n=0}^{\infty} c^n = \lim_{n\to\infty}(A_n) = \frac{0-1}{c-1} = \frac{1}{1-c} \qquad \text{“\textbf{\textit{geometrische Reihe}}”.}$$
Für unser Eingangsbeispiel mit $c = \frac{1}{2}$ erhalten wir jetzt $1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots = \sum_{n=0}^{\infty} \frac{1}{2^n} = 2$. Ist $|c| \geqslant 1$ und $c \neq 1$, so kann man sich überlegen, dass die Folge $(A_n)_{n\in\mathbb{N}_0}$ nicht konvergiert.

Lemma 28.3. *Ist die unendliche Reihe $\sum_{n=0}^{\infty} a_n$ konvergent, so ist $(a_n)_{n\in\mathbb{N}_0}$ eine Null-Folge.*

Beweis. Für $n \in \mathbb{N}_0$ sei $A_n = a_0+a_1+\cdots+a_n$. Nach Voraussetzung existiert $\alpha := \lim\limits_{n\to\infty}(A_n) \in \mathbb{R}$. Sei $\varepsilon > 0$. Nach Lemma 27.5 gibt es ein $n_0 \in \mathbb{N}_0$ mit $|A_n - A_m| < \varepsilon$ für alle $n \geqslant m \geqslant n_0$. Für $n > n_0$ und $m := n-1 \geqslant n_0$ ist $a_n = A_n - A_{n-1}$ und damit $|a_n| = |A_n - A_{n-1}| < \varepsilon$. Also ist $(a_n)_{n\in\mathbb{N}_0}$ eine Null-Folge. □

Bemerkung 28.4. Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge und $(A_n)_{n\in\mathbb{N}_0}$ die zugehörige Folge der Partialsummen, also $A_n = \sum_{k=0}^{n} a_k$ für alle $n \in \mathbb{N}_0$.

(a) In konkreten Beispielen sind manchmal die Folgenglieder a_n erst ab einem bestimmten Index $d \in \mathbb{N}$ definiert. Man kann dann aber stets zur Folge $(\tilde{a}_n)_{n\in\mathbb{N}_0}$ übergehen, wie in Bemerkung 27.3; das Verhalten der Partialsummen beim Grenzübergang $n \to \infty$ ändert sich nicht. Wir schreiben dann auch einfach $\sum_{n=d}^{\infty} a_n := \sum_{n=0}^{\infty} \tilde{a}_n$.

(b) Sei $m \in \mathbb{N}$ fest. Sei $a'_n := a_{m+n}$ für $n \in \mathbb{N}_0$ und $(A'_n)_{n\in\mathbb{N}_0}$ die zugehörige Folge der Partialsummen, also $A'_n = \sum_{k=0}^{n} a'_k = \sum_{k=m}^{m+n} a_k$ für alle $n \in \mathbb{N}_0$. Wir schreiben dann auch einfach $\sum_{n=m}^{\infty} a_n := \sum_{n=0}^{\infty} a'_n$. Man sieht leicht: Die Reihe $\sum_{n=0}^{\infty} a_n$ konvergiert genau dann,

wenn die Reihe $\sum_{n=m}^{\infty} a_n$ konvergiert; in diesem Fall gilt
$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + \ldots + a_{m-1} + \sum_{n=m}^{\infty} a_n.$$

Beispiel 28.5. Die unendliche Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ konvergiert nicht.
Dazu: Wäre die Reihe konvergent, so konvergiert also die Folge der zugehörigen Partialsummen $(A_n)_{n\in\mathbb{N}}$; diese müsste also inbesondere nach oben beschränkt sein (Lemma 27.4), d.h., es gibt ein $C \in \mathbb{R}$ mit $A_n \leqslant C$ für alle $n \in \mathbb{N}$. Sei nun aber $n = 2^k$ für $k \in \mathbb{N}$. Dann gilt
$$\begin{aligned} A_{2^k} &= 1 + \frac{1}{2} + \Big(\frac{1}{3} + \frac{1}{4}\Big) + \Big(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\Big) + \cdots + \Big(\frac{1}{2^{k-1}+1} + \frac{1}{2^{k-1}+2} + \cdots + \frac{1}{2^k}\Big) \\ &\geqslant 1 + \frac{1}{2} + \Big(\frac{1}{4} + \frac{1}{4}\Big) + \Big(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\Big) + \cdots + \Big(\frac{1}{2^k} + \frac{1}{2^k} + \cdots + \frac{1}{2^k}\Big) \\ &= 1 + \frac{1}{2} + 2\cdot\frac{1}{4} + 4\cdot\frac{1}{8} + \cdots + 2^{k-1}\cdot\frac{1}{2^k} = 1 + \frac{1}{2}k. \end{aligned}$$
Ist also $k > 2(C-1)$, so folgt $A_{2^k} > C$, Widerspruch.

Beispiel 28.6. Es gibt eine Darstellung für die ***Eulersche Zahl*** e als unendliche Reihe:
$$e = \lim_{n\to\infty}\Big(1 + \frac{1}{n}\Big)^n = \sum_{n=0}^{\infty} \frac{1}{n!}.$$
Dazu: Für $n \in \mathbb{N}_0$ sei $A_n := \sum_{k=0}^{n} \frac{1}{k!}$. Dann ist $A_{n+1} = A_n + \frac{1}{(n+1)!} \geqslant A_n$ für alle $n \in \mathbb{N}_0$. Um die Konvergenz der Folge $(A_n)_{n\in\mathbb{N}_0}$ zu zeigen, genügt es also nach dem Monotonie-Prinzip, eine obere Schranke zu finden. Mit dem Binomischen Lehrsatz erhalten wir
$$\Big(1 + \frac{1}{n}\Big)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^{n} \frac{n(n-1)\cdots(n-k+1)}{k!}\frac{1}{n^k} = \sum_{k=0}^{n} \frac{n(n-1)\cdots(n-k+1)}{n^k}\frac{1}{k!}$$
für alle $n \in \mathbb{N}$. Für $0 \leqslant k \leqslant n$ setzen wir zur Abkürzung
$$d(n,k) := \frac{n(n-1)\cdots(n-k+1)}{n^k} = \frac{n}{n}\cdot\frac{n-1}{n}\cdot\ldots\cdot\frac{n-k+1}{n}.$$
Mit dem "Trick" in Beispiel 27.11(a) folgt sofort $\lim\limits_{n\to\infty}\big(d(n,k)\big) = 1$. Wir benutzen dies nun, um A_n nach oben abzuschätzen. Sei dazu $m \in \mathbb{N}_0$ fest und $n > m$. Dann ist
$$\Big(1 + \frac{1}{n}\Big)^n = \sum_{k=0}^{n} d(n,k)\frac{1}{k!} = \sum_{k=0}^{m} d(n,k)\frac{1}{k!} + \underbrace{\sum_{k=m+1}^{n} d(n,k)\frac{1}{k!}}_{\geqslant 0} \geqslant \sum_{k=0}^{m} d(n,k)\frac{1}{k!}.$$
Mit den Regeln in Satz 27.8 folgt, dass die rechte Seite für $n \to \infty$ gegen $\sum_{k=0}^{m} \frac{1}{k!} = A_m$ konvergiert. Die linke Seite konvergiert gegen e, also folgt $A_m \leqslant e$; dies gilt für alle $m \in \mathbb{N}_0$. Insbesondere ist die Folge $(A_m)_{m\in\mathbb{N}_0}$ nach oben beschränkt und konvergiert. Es folgt dann auch $\tilde{e} := \lim\limits_{m\to\infty}(A_m) \leqslant e$. Andererseits ist $0 < d(n,k) \leqslant 1$, also $(1 + \frac{1}{n})^n \leqslant \sum_{k=0}^{n} \frac{1}{k!} = A_n$ für alle $n \in \mathbb{N}$, und damit $e = \lim\limits_{n\to\infty}(1 + \frac{1}{n})^n \leqslant \lim\limits_{n\to\infty}(A_n) = \tilde{e}$. Also schließlich $e = \tilde{e}$. □

Lemma 28.7. *Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ konvergente Reihen und $r, s \in \mathbb{R}$. Dann konvergiert auch die unendliche Reihe $\sum_{n=0}^{\infty}(ra_n + sb_n)$, mit Grenzwert $r\big(\sum_{n=0}^{\infty} a_n\big) + s\big(\sum_{n=0}^{\infty} b_n\big)$.*

Beweis. Dies folgt sofort mit Satz 27.8 angewandt auf die zugehörigen Partialsummen. □

Für Produkte von unendlichen Reihen ist die Situation komplizierter. Um dies zu behandeln (siehe Satz 28.13 unten), benötigen wir zunächst den folgenden Begriff.

Definition 28.8. Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge reeller Zahlen. Die unendliche Reihe $\sum_{n=0}^{\infty} a_n$ heißt ***absolut konvergent***, wenn die Reihe der Absolutbeträge $\sum_{n=0}^{\infty} |a_n|$ konvergiert.

Für die Partialsummen $A_n^* := |a_0| + |a_1| + \ldots + |a_n|$ gilt dann $A_{n+1}^* = A_n^* + |a_{n+1}| \geqslant A_n^*$ für alle $n \in \mathbb{N}_0$. Also ist $\sum_{n=0}^{\infty} |a_n|$ nach dem Monotonie-Prinzip konvergent, wenn $(A_n^*)_{n\in\mathbb{N}_0}$ nach oben beschränkt ist.

Beispiel 28.9. Für $n \in \mathbb{N}$ sei $a_n := \frac{1}{n^2}$. Konvergiert die unendliche Reihe $\sum_{n=1}^{\infty} a_n$? Dazu zeigt man am besten, dass die Folge der Partialsummen nach oben beschränkt ist, zum Beispiel durch 2 (siehe Übungen). Da $|a_n| = a_n \geqslant 0$ für alle $n \in \mathbb{N}$ gilt, folgt mit dem Monotonie-Prinzip, dass $\sum_{n=1}^{\infty} \frac{1}{n^2}$ konvergiert (siehe obige Bemerkung). Was ist der Grenzwert? Der geniale Euler zeigte 1735, dass $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$ gilt (was selbst heutzutage nicht so ganz einfach ist ...).

Satz 28.10 (***Majoranten-Kriterium***)**.** *Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge reeller Zahlen.*
(a) *Sei $(b_n)_{n\in\mathbb{N}_0}$ eine weitere Folge mit $0 \leqslant a_n \leqslant b_n$ für alle $n \in \mathbb{N}_0$. Ist die unendliche Reihe $\sum_{n=0}^{\infty} b_n$ konvergent, so konvergiert die unendliche Reihe $\sum_{n=0}^{\infty} a_n$.*
(b) *Wenn die unendliche Reihe $\sum_{n=0}^{\infty} |a_n|$ konvergiert, so auch die unendliche Reihe $\sum_{n=0}^{\infty} a_n$.*

Beweis. (a) Für $n \in \mathbb{N}_0$ sei $A_n := a_0 + \cdots + a_n$ und $B_n := b_0 + \cdots + b_n$. Wenn der Grenzwert $\sum_{n=0}^{\infty} b_n \in \mathbb{R}$ existiert, so ist die Folge $(B_n)_{n\in\mathbb{N}_0}$ konvergent, also gibt es nach Lemma 27.4 ein $C \in \mathbb{R}$ mit $B_n \leqslant C$ für alle $n \in \mathbb{N}_0$. Wegen $a_k \leqslant b_k$ für alle $k \in \mathbb{N}_0$ folgt $A_n = a_0 + \cdots + a_n \leqslant b_0 + \cdots + b_n \leqslant B_n \leqslant C$ für alle $n \in \mathbb{N}_0$, also ist $\{A_n \mid n \in \mathbb{N}_0\}$ nach oben beschränkt. Nun ist auch $a_n \geqslant 0$ für alle $n \in \mathbb{N}_0$, also $A_{n+1} = A_n + a_{n+1} \geqslant A_n$ für alle $n \in \mathbb{N}_0$. Nach dem Monotonie-Prinzip konvergiert $(A_n)_{n\in\mathbb{N}_0}$, also die Reihe $\sum_{n=0}^{\infty} a_n$.

(b) Wir definieren zwei neue Folgen durch

$$a_n^+ := \begin{cases} a_n & \text{falls } a_n \geqslant 0 \\ 0 & \text{falls } a_n < 0 \end{cases} \qquad \text{und} \qquad a_n^- = \begin{cases} -a_n & \text{falls } a_n \leqslant 0 \\ 0 & \text{falls } a_n > 0 \end{cases}.$$

Dann gilt $|a_n| = a_n^+ + a_n^-$ für alle $n \in \mathbb{N}_0$; außerdem $0 \leqslant a_n^+ \leqslant |a_n|$ und $0 \leqslant a_n^- \leqslant |a_n|$. Sei nun $\sum_{n=0}^{\infty} |a_n|$ konvergent. Nach (a) konvergieren dann auch $\sum_{n=0}^{\infty} a_n^+$ und $\sum_{n=0}^{\infty} a_n^-$. Nun gilt weiterhin $a_n = a_n^+ - a_n^-$ für alle $n \in \mathbb{N}_0$. Also konvergiert auch $\sum_{n=0}^{\infty} a_n$ nach Lemma 28.7. □

Beispiel 28.11. Sei $\mathcal{A}$ die Menge aller Folgen $\alpha = (z_n)_{n\in\mathbb{N}_0}$ mit $z_0 \in \mathbb{N}_0$ and $z_n \in \{0, 1, \ldots, 9\}$ für $n \geqslant 1$. Für $\alpha = (z_n)_{n\in\mathbb{N}_0}$ bilden wir die unendliche Reihe

$$\hat{\alpha} := \sum_{n=0}^{\infty} \frac{z_n}{10^n} = a_0 + \sum_{n=1}^{\infty} \frac{z_n}{10^n}.$$

Diese konvergiert absolut, denn: Es gilt $0 \leqslant \frac{z_n}{10^n} \leqslant b_n := 9(\frac{1}{10})^n$ für alle $n \in \mathbb{N}$; nach Beispiel 28.2 (mit $c = \frac{1}{10}$) und Lemma 28.7 konvergiert $\sum_{n=1}^{\infty} b_n$; nach Satz 28.10(a) konvergiert also auch $\hat{\alpha}$. Damit definiert jede Folge $\alpha \in \mathcal{A}$ eine reelle Zahl $\hat{\alpha} \in \mathbb{R}$; wir schreiben

$$\boxed{\hat{\alpha} = z_0{,}z_1z_2z_3\cdots \qquad \text{“Dezimalentwicklung”}}$$

Umgekehrt gibt es zu jeder reellen Zahl $x \in \mathbb{R}$, $x \geqslant 0$, ein $\alpha \in \mathcal{A}$ mit $x = \hat{\alpha}$. Dazu definiert man eine Folge $(z_n)_{n\in\mathbb{N}_0}$ wie in Beispiel 10.7, Kapitel II; nach Konstruktion ist $\alpha := (z_n)_{n\in\mathbb{N}_0} \in \mathcal{A}$. Mit $a_n := \sum_{k=0}^{n} \frac{z_k}{10^k}$ gilt außerdem $0 \leqslant x - a_n < \frac{1}{10^n}$ für alle $n \in \mathbb{N}_0$. Also konvergiert $(a_n)_{n\in\mathbb{N}_0}$, mit Grenzwert x, d.h., es ist $x = \hat{\alpha}$.

Folgerung 28.12. *Zu jedem $x \in \mathbb{R}$ gibt es eine konvergente Folge $(a_n)_{n\in\mathbb{N}_0}$ mit $x = \lim_{n\to\infty}(a_n)$ und $a_n \in \mathbb{Q}$ für alle $n \in \mathbb{N}_0$.*

Beweis. Sei zuerst $x \geqslant 0$. Nach Beispiel 28.11 gibt es ein $\alpha = (z_n)_{n\in\mathbb{N}_0} \in \mathcal{A}$ mit $x = \hat{\alpha}$. Dann ist die gewünschte Folge gegeben durch $a_n := \sum_{k=0}^{n} \frac{z_k}{10^k} \in \mathbb{Q}$ für alle $n \in \mathbb{N}_0$. Ist $x < 0$, so wende das vorherige Argument auf $-x$ an. □

Satz 28.13 (***Cauchy–Produkt von Reihen***)**.** *Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ absolut konvergente Reihen. Für $n \in \mathbb{N}_0$ definiere $c_n := \sum_{k=0}^{n} a_k b_{n-k}$ (wie beim Produkt von Polynomen, siehe §9, Kapitel II). Dann ist auch die Reihe $\sum_{n=0}^{\infty} c_n$ konvergent und es gilt*

$$\sum_{n=0}^{\infty} c_n = \Big(\sum_{n=0}^{\infty} a_n\Big) \cdot \Big(\sum_{n=0}^{\infty} b_n\Big).$$

Beweis. Da die Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ absolut konvergieren, so existieren auch $\alpha := \sum_{n=0}^{\infty} a_n \in \mathbb{R}$ und $\beta := \sum_{n=0}^{\infty} b_n \in \mathbb{R}$; siehe Satz 28.10(b). Für $N \in \mathbb{N}_0$ setze $C_N := \sum_{n=0}^{N} c_n$. Sei $\varepsilon > 0$. Wir müssen ein $N_0 \in \mathbb{N}$ finden, so dass $|C_N - \alpha\cdot\beta| < \varepsilon$ für alle $N \geqslant N_0$ gilt. Wir gehen dazu in mehreren Schritten vor. Zunächst einige Bezeichnungen: Für $N \in \mathbb{N}_0$ setzen wir auch $A_N := \sum_{n=0}^{N} a_n$ und $B_N := \sum_{n=0}^{N} b_n$, sowie

$$\Delta_N := \{(k,l) \in \mathbb{N}_0 \times \mathbb{N}_0 \mid k + l \leqslant N\} \quad \text{und} \quad Q_N := \{(k,l) \in \mathbb{N}_0 \times \mathbb{N}_0 \mid k \leqslant N \text{ und } l \leqslant N\};$$

außerdem setze $A_N^* := \sum_{n=0}^{N} |a_n|$ und $B_N^* := \sum_{n=0}^{N} |b_n|$. Dann gilt

$$C_N = \sum_{(k,l)\in\Delta_N} a_k b_l, \qquad A_N \cdot B_N = \sum_{(k,l)\in Q_N} a_k b_l, \qquad A_N^* \cdot B_N^* = \sum_{(k,l)\in Q_N} |a_k| \cdot |b_l|.$$

Nun beachte:

$$|C_N - \alpha\cdot\beta| = |C_N - A_N \cdot B_N + A_N \cdot B_N - \alpha\cdot\beta| \leqslant |A_N \cdot B_N - C_N| + |A_N \cdot B_N - \alpha\cdot\beta|.$$

Nach Satz 27.8 konvergiert $(A_N \cdot B_N)_{N\in\mathbb{N}_0}$, mit Grenzwert $\alpha\cdot\beta$. Also gibt es ein $N_1 \in \mathbb{N}_0$ mit $|A_N \cdot B_N - \alpha\cdot\beta| < \varepsilon/2$ für alle $N \geqslant N_1$. Weiterhin: Es gilt offenbar $\Delta_N \subseteq Q_N$, also folgt

$$|A_N \cdot B_N - C_N| = \Big| \sum_{(k,l)\in Q_N\setminus\Delta_N} a_k b_l \Big| \leqslant \sum_{(k,l)\in Q_N\setminus\Delta_N} |a_k|\cdot|b_l|.$$

Sei nun $N' := N/2$ falls N gerade ist, und $N' := (N-1)/2$ falls N ungerade ist. In jedem Fall gilt also $2N' \leqslant N$. Dann ist $Q_{N'} \subseteq \Delta_N$. (Denn für $(k,l) \in Q_{N'}$ ist $k \leqslant N'$ und $l \leqslant N'$, also $k+l \leqslant 2N' \leqslant N$ und damit $(k,l) \in \Delta_N$.) Es folgt $Q_N \setminus \Delta_N \subseteq Q_N \setminus Q_{N'}$, also

$$|A_N \cdot B_N - C_N| \leqslant \sum_{(k,l)\in Q_N\setminus Q_{N'}} |a_k|\cdot|b_l| = A_N^* \cdot B_N^* - A_{N'}^* \cdot B_{N'}^*.$$

Mit $(A_N^*)_{N\in\mathbb{N}_0}$ und $(B_N^*)_{N\in\mathbb{N}_0}$ konvergiert auch $(A_N^*\cdot B_N^*)_{N\in\mathbb{N}_0}$ (siehe Satz 27.8). Nach Lemma 27.5 gibt es zu unserem $\varepsilon > 0$ ein $N_2 \in \mathbb{N}_0$ mit $|A_N^*\cdot B_N^* - A_M^*\cdot B_M^*| < \varepsilon/2$ für $N \geqslant M \geqslant N_2$. Nun sei $N \geqslant N_0 := \max\{N_1, 2N_2+1\}$. Dann ist $N \geqslant N' \geqslant N_2$ (siehe Definition von N') und damit $|A_N \cdot B_N - C_N| < \varepsilon/2$. Insgesamt also $|C_N - \alpha\cdot\beta| < \varepsilon/2 + \varepsilon/2 = \varepsilon$ für alle $N \geqslant N_0$. □

Es folgen nun noch einige weitere Konvergenz-Kriterien.

Satz 28.14 (***Leibniz–Konvergenz-Kriterium***)**.** *Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Null-Folge mit $a_n \geqslant 0$ und $a_n \geqslant a_{n+1}$ für alle $n \in \mathbb{N}_0$. Dann konvergiert die unendliche Reihe $\sum_{n=0}^\infty (-1)^n a_n$.*

Beweis. Siehe zum Beispiel §7, Satz 4, im 1. Band des Buches von Forster. □

Beispiel 28.15. (a) Nach Satz 28.14 konvergiert die folgende unendliche Reihe:

$$s = 4\sum_{n=0}^\infty \frac{(-1)^n}{2n+1} = 4\cdot\Big(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \cdots\Big).$$

Hier sind einige Werte für die zugehörigen Partilalsummen A_n:

n	0	1	2	3	$\cdots$	1000	100000	$\cdots$
A_n	4	$\approx 2{,}6667$	$\approx 3{,}4667$	$\approx 2{,}8952$	$\cdots$	$\approx 3{,}1406$	$\approx 3{,}1415$	$\cdots$

(Schreiben Sie dazu ein kleines Programm in Sage, wie in Beispiel 27.15.) Es sieht so aus, als würde $(A_n)_{n\in\mathbb{N}_0}$ gegen die Kreiszahl π konvergieren! (Mehr dazu in §34.)

(b) Nach Satz 28.14 konvergiert die folgende unendliche Reihe:

$$s = \sum_{n=1}^\infty (-1)^{n-1}\frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} \pm \cdots.$$

Versuchen wir, s auszurechnen durch kleveres Umsortieren der Terme in der Summation:

$$\begin{aligned} s &= 1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \cdots \qquad \text{(jeder “+” Term gefolgt von zwei “−” Termen)} \\ &= \underbrace{\Big(1 - \frac{1}{2}\Big)}_{=1/2} - \frac{1}{4} + \underbrace{\Big(\frac{1}{3} - \frac{1}{6}\Big)}_{1/6} - \frac{1}{8} + \underbrace{\Big(\frac{1}{5} - \frac{1}{10}\Big)}_{=1/10} - \frac{1}{12} + \cdots = \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \cdots = \frac{s}{2}. \end{aligned}$$

Also $s = 0$. Andererseits ist $s = \big(1 - \frac{1}{2}\big) + \big(\frac{1}{3} - \frac{1}{4}\big) + \big(\frac{1}{5} - \frac{1}{6}\big) + \cdots = \frac{1}{2} + \frac{1}{12} + \frac{1}{30} + \cdots \geqslant \frac{1}{2}$, Widerspruch. — Es ist also nicht erlaubt, die Terme in einer solchen unendlichen Reihe beliebig umzusortieren! Ein solches Verhalten kann nicht passieren bei absolut konvergenten Reihen (siehe zum Beispiel §7, Satz 8 im 1. Band des Buches von Forster).

Das Beispiel der obigen Reihe $\sum_{n=1}^{\infty}(-1)^{n-1}/n$ zeigt, dass eine konvergente Reihe nicht absolut konvergent zu sein braucht (siehe Beispiel 28.5).

Satz 28.16 (***Quotienten-Kriterium***). *Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge. Es gebe ein $n_0 \in \mathbb{N}_0$ mit $a_n \neq 0$ für alle $n \geqslant n_0$, sowie ein $c \in \mathbb{R}$ mit $0 \leqslant c < 1$ und $|a_{n+1}/a_n| \leqslant c$ für alle $n \geqslant n_0$. Dann konvergiert die Reihe $\sum_{n=0}^{\infty} a_n$ absolut.*

Beweis. Siehe zum Beispiel §7, Satz 7, im 1. Band des Buches von Forster. □

Beispiel: Sei $a_n := \frac{n^2}{2^n}$ für alle $n \in \mathbb{N}_0$. Dann gilt für $n \geqslant 3$: $|\frac{a_{n+1}}{a_n}| = \frac{(n+1)^2 2^n}{2^{n+1}n^2} = \frac{1}{2}(1+\frac{1}{n})^2 \leqslant \frac{1}{2}(1+\frac{1}{3})^2 = \frac{8}{9} =: c < 1$. Also ist die unendliche Reihe $\sum_{n=0}^{\infty} a_n$ absolut konvergent.

Siehe §7 im 1. Band des Buches von Forster für einige weitere Konvergenz-Kriterien.

Ab hier Woche 5

29. *Potenzreihen und die Exponential–Funktion*

Sei $(a_n)_{n\in\mathbb{N}_0}$ eine beliebige Folge reeller Zahlen und $D \subseteq \mathbb{R}$ die Menge aller $x \in \mathbb{R}$ so dass die "***Potenzreihe***" $\sum_{n=0}^{\infty} a_n x^n$ konvergiert. Dadurch erhalten wir also eine Funktion

$$f\colon D \to \mathbb{R} \qquad \text{mit} \qquad f(x) := \sum_{n=0}^{\infty} a_n x^n \in \mathbb{R} \qquad \text{für alle } x \in D.$$

Beachte: Es gilt stets $0 \in D$, mit $f(0) = a_0$. Solche Potenzreihen sind von zentraler Bedeutung für die gesamte Analysis. Der folgende Satz bestimmt die Menge D genauer.

Satz 29.1. *Seien $(a_n)_{n\in\mathbb{N}_0}$ und $f\colon D \to \mathbb{R}$ wie oben. Dann ist entweder $D = \mathbb{R}$ oder $D = \{0\}$ oder es gibt ein $R_0 \in \mathbb{R}$, $R_0 > 0$, mit $(-R_0, R_0) \subseteq D \subseteq [-R_0, R_0]$. Für alle $x \in (-R_0, R_0)$ (bzw. für alle $x \in \mathbb{R}$ im Fall $D = \mathbb{R}$) konvergiert die Reihe $f(x) = \sum_{n=0}^{\infty} a_n x^n$ absolut.*

Hier heißt R_0 der ***Konvergenzradius*** von f. (Setze $R_0 := 0$ für $D = \{0\}$, $R_0 := \infty$ für $D = \mathbb{R}$.)

Beweis. Für $x \in \mathbb{R}$ sei $A_n(x) := \sum_{k=0}^{n} a_k x^k$. Nehmen wir nun $D \neq \{0\}$ an, und sei $0 \neq x_0 \in D$. Dann konvergiert die Folge $(A_n(x_0))_{n\in\mathbb{N}_0}$, also ist $\big(a_k x_0^k\big)_{k\in\mathbb{N}_0}$ eine Null-Folge (siehe Lemma 28.3), also inbesondere nach oben und unten beschränkt, d.h., es gibt ein $C \in \mathbb{R}_{>0}$ mit $|a_k x_0^k| \leqslant C$ für alle $k \in \mathbb{N}_0$. Sei $x \in \mathbb{R}$ beliebig mit $0 < |x| < r := |x_0|$. Dann gilt

$$|a_k x^k| = \left|a_k\, x_0^k\, \frac{x^k}{x_0^k}\right| = \left|a_k x_0^k\right| q^k \qquad \text{mit} \qquad 0 < q := \left|\frac{x}{x_0}\right| < 1.$$

Damit folgt $\sum_{k=0}^{n} \left|a_k x^k\right| \leqslant \sum_{k=0}^{n} \left|a_k x_0^k\right| q^k \leqslant C \sum_{k=0}^{n} q^k \leqslant C \sum_{k=0}^{\infty} q^k = \frac{C}{1-q}$, wobei wir Beispiel 28.2 (mit $c = q$) benutzen. Damit ist $f(x) = \sum_{k=0}^{\infty} a_k x^k$ absolut konvergent (siehe Anmerkung direkt im Anschluss an Definition 28.8), also auch selbst konvergent (Satz 28.10)

und damit $x \in D$. Dies zeigt, dass $(-r, r) \subseteq D$ gilt. Jetzt gibt es 2 Fälle:
1. Fall: D ist nicht nach oben beschränkt. Zu jedem $x \in \mathbb{R}$ gibt es also ein $r \in D$ mit $|x| < r$. Wie oben folgt $x \in (-r, r) \subseteq D$. In diesem Fall ist also $D = \mathbb{R}$.
2. Fall: D ist nach oben beschränkt; sei $R_0 := \sup(D) \in \mathbb{R}$. Es ist stets $0 \in D$, also $R_0 \geqslant 0$. Ist $R_0 = 0$, so $D = \{0\}$. Sei nun $R_0 > 0$. Für $x \in (-R_0, R_0)$ ist $|x|$ keine obere Schranke für D, also gibt es ein $r \in D$ mit $|x| < r$. Wie oben folgt wieder $x \in (-r, r) \subseteq D$, also $(-R_0, R_0) \subseteq D$. Sei nun $x \in D$; dann ist natürlich $x \leqslant R_0$. Wäre $x < -R_0$, so $R_0 < r := -x = |x|$ und es gibt ein $x_1 \in \mathbb{R}$ mit $R_0 < x_1 < r$; es folgt wieder $x_1 \in (-r, r) \subseteq D$, Widerspruch zu $R_0 = \sup(D)$. □

Beispiel 29.2. (a) Sei $a_n := 1$ für alle $n \in \mathbb{N}_0$. Für $x \in \mathbb{R}$ erhalten wir also die Potenzreihe $f(x) = \sum_{n=0}^{\infty} x^n$. Für $x = \pm 1$ ist diese divergent, also $R_0 \leqslant 1$. Für $|x| < 1$ ist $f(x)$ nach Beispiel 28.2 konvergent, mit Grenzwert $1/(1-x)$. Also ist hier $R_0 = 1$.

(b) Sei $a_0 := 0$ und $a_n := \frac{1}{n}$ für $n \in \mathbb{N}$. Für $x \in \mathbb{R}$ erhalten wir also die Potenzreihe

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = \sum_{n=1}^{\infty} \frac{x^n}{n} = x + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \dots .$$

Was ist der Konvergenzradius R_0 ? Hier ist wieder $f(1)$ divergent, also $R_0 \leqslant 1$. Sei nun $x \in \mathbb{R}$ mit $|x| < 1$. Mit Beispiel 28.2 sehen wir $\sum_{k=1}^{n} \frac{|x|^k}{k} \leqslant \sum_{k=1}^{n} |x|^k \leqslant 1/(1-|x|)$ für alle $n \in \mathbb{N}$, also ist $R_0 = 1$. (Übrigens ist hier nach Beispiel 28.15(b) auch $f(-1)$ konvergent, aber eben nicht absolut konvergent.) — Wir werden später noch sehen, was genau diese Funktion f ist.

(c) Sei $a_0 := 0$ und $a_n := \frac{2^n}{n^2}$ für alle $n \in \mathbb{N}$. Für $x \in \mathbb{R}$ erhalten wir also die Potenzreihe

$$f(x) = \sum_{n=0}^{\infty} a_n x^n = \sum_{n=1}^{\infty} \frac{(2x)^n}{n^2} = 2x + \frac{4x^2}{4} + \frac{8x^3}{9} + \frac{16x^4}{16} + \frac{32x^5}{25} + \dots .$$

Die Folge $(a_n)_{n \in \mathbb{N}_0}$ ist bestimmt divergent gegen $+\infty$ (siehe Übungen). Aber für $x = \frac{1}{2}$ ist $f(x) = \sum_{n=1}^{\infty} \frac{1}{n^2}$ dennoch konvergent (siehe Beispiel 28.9), also $R_0 \geqslant \frac{1}{2}$. Man kann zeigen, dass für $x > \frac{1}{2}$ die Folge $(a_n x^n)_{n \in \mathbb{N}}$ keine Null-Folge ist (siehe Übungen), also die Reihe $f(x)$ divergent ist. Folglich ist hier $R_0 = \frac{1}{2}$ und $f(x)$ absolut konvergent auf ganz $D = [-\frac{1}{2}, \frac{1}{2}]$.

Bemerkung 29.3. Sei $(a_n)_{n \in \mathbb{N}_0}$ eine Folge und $f\colon D \to \mathbb{R}$ wie oben definiert, also $f(x) = \sum_{n=0}^{\infty} a_n x^n$ für $x \in D$. Zur näherungsweisen Berechnung von $f(x)$ können wir die Partialsummen $F_n(x) := \sum_{k=0}^{n} a_k x^k$ verwenden, für hinreichend große $n \in \mathbb{N}_0$. Wir benötigen dann typischerweise auch eine Abschätzung für den Fehler $|f(x) - F_n(x)|$.
Zur Illustration betrachten wir die Potenzreihe in Beispiel 29.2(b), also $f(x) = \sum_{n=1}^{\infty} \frac{x^n}{n}$ für $x \in (-1, 1)$. Für $m, n \in \mathbb{N}$ mit $n > m$ ist dann

$$\begin{aligned} |F_n(x) - F_m(x)| &= \Big| \sum_{k=m+1}^{n} \frac{x^k}{k} \Big| \leqslant \sum_{k=m+1}^{n} \frac{|x|^k}{k} \leqslant \frac{|x|^{m+1}}{m+1} \sum_{k=m+1}^{n} \frac{m+1}{k} |x|^{k-m-1} \\ &\leqslant \frac{|x|^{m+1}}{m+1} \sum_{k=m+1}^{n} |x|^{k-m-1} \leqslant \frac{|x|^{m+1}}{m+1} \sum_{k=0}^{\infty} |x|^k = \frac{|x|^{m+1}}{(1-|x|)(m+1)}, \end{aligned}$$

wobei die letzte Gleichheit wieder mit Beispiel 28.2 folgt. Sei nun $d_n(x) := |f(x) - F_n(x)|$ für $n \in \mathbb{N}$. Da $(F_n(x))_{n\in\mathbb{N}}$ gegen $f(x)$ konvergiert, ist $(d_n(x))_{n\in\mathbb{N}}$ eine Null-Folge. Sei $n > m \geqslant 1$, mit m fest. Dann ist $d_m(x) = |f(x)-F_n(x)+F_n(x)-F_m(x)| \leqslant |f(x)-F_n(x)|+|F_n(x)-F_m(x)| \leqslant d_n(x) + |F_n(x) - F_m(x)| \leqslant d_n(x) + \frac{|x|^{m+1}}{(1-|x|)(m+1)}$. Mit Beispiel 27.10 folgt

$$\big|f(x) - F_m(x)\big| = d_m(x) \leqslant \frac{|x|^{m+1}}{(1-|x|)(m+1)} \qquad \text{für alle } m \in \mathbb{N} \text{ und } x \in (-1,1).$$

Geben wir uns also $x \in (-1,1)$ und die gewünschte Anzahl N der korrekt zu berechnenden Nachkommastellen in der Dezimalentwicklung von $f(x)$ vor, so wählen wir einfach $m \in \mathbb{N}$ so, dass $\frac{|x|^{m+1}}{(1-|x|)(m+1)} \leqslant 10^{-(N+1)}$ gilt; für jedes solche m ist dann $|f(x) - F_m(x)| \leqslant 10^{-(N+1)}$.

In Sage lässt sich dies zum Beispiel auf sehr einfache Weise programmieren:

```
sage: def freihe(x,nachkomma):        # f wie oben, nachkomma = Anzahl
....:   m=1                           #     korrekter Nachkommastellen
....:   while abs(x)^(m+1)*10^(nachkomma+1)>(m+1)*(1-abs(x)): m+=1
....:   return sum(x^(k+1)/(k+1) for k in range(m))
sage: dec_exp(freihe(1/2,16))         # x=1/2, 16 korrekte Nachkommastellen
'0.6931471805599453'                  # kennen Sie diese Zahl?
```

(Am Ende benutzen wir wieder die Funktion `dec_exp`; siehe Tabelle 2, S. 23.)

In Anlehnung an die Reihenentwicklung von e in Beispiel 28.6 setzen wir nun:

$$e_k(x) := \frac{x^k}{k!} \quad \text{und} \quad E_n(x) := \sum_{k=0}^{n} e_k(x) = \sum_{k=0}^{n} \frac{x^k}{k!} \qquad \text{für alle } k,n \in \mathbb{N}_0 \text{ und } x \in \mathbb{R}.$$

Die nähere Betrachtung der zugehörigen Potenzreihe wird es uns am Ende erlauben, so etwas wie e^x für $x \in \mathbb{R}$ zu definieren.

Lemma 29.4. *Für jedes $x \in \mathbb{R}$ ist $(e_k(x))_{k\in\mathbb{N}_0}$ eine Null-Folge. Sind $m,n \in \mathbb{N}_0$ mit $n > m \geqslant 2|x| - 2$, so gilt $|E_n(x) - E_m(x)| \leqslant 2|e_{m+1}(x)|$.*

Beweis. Sei $x \in \mathbb{R}$ fest und $m \in \mathbb{N}_0$ so, dass $m + 2 \geqslant 2|x|$ gilt. Für alle $k \geqslant m+1$ gilt dann

$$\frac{|x|^k}{k!} = \frac{|x|^{m+1}}{(m+1)!} \cdot \underbrace{\frac{|x|}{m+2} \cdot \frac{|x|}{m+3} \cdot \ldots \cdot \frac{|x|}{k-1} \cdot \frac{|x|}{k}}_{k-m-1 \text{ Faktoren}} = |e_{m+1}(x)| \cdot \frac{|x|}{m+2} \cdot \frac{|x|}{m+3} \cdot \ldots \cdot \frac{|x|}{k-1} \cdot \frac{|x|}{k}$$

$$\leqslant |e_{m+1}(x)| \cdot \frac{|x|}{m+2} \cdot \frac{|x|}{m+2} \cdot \frac{|x|}{m+2} \cdot \ldots \cdot \frac{|x|}{m+2} \leqslant \frac{|e_{m+1}(x)|}{2^{k-m-1}} = \frac{2^{m+1}|e_{m+1}(x)|}{2^k}.$$

Da $(1/2^k)_{k\in\mathbb{N}_0}$ eine Null-Folge ist, folgt aus der obigen Abschätzung, dass auch $(e_k(x))_{k\in\mathbb{N}_0}$ eine Null-Folge ist. Sei nun $n \in \mathbb{N}_0$ mit $n > m$. Mit $c := |e_{m+1}(x)|$ folgt dann:

$$|E_n(x) - E_m(x)| = \Big|\sum_{k=m+1}^{n} \frac{x^k}{k!}\Big| \leqslant \sum_{k=m+1}^{n} \frac{|x|^k}{k!} \leqslant \sum_{k=m+1}^{n} \frac{c}{2^{k-m-1}} \leqslant c \cdot \underbrace{\sum_{k=0}^{n-m-1} \frac{1}{2^k}}_{\leqslant 2} \leqslant 2c. \qquad \square$$

Satz 29.5 (***Exponential-Funktion***). *Für jedes* $x \in \mathbb{R}$ *konvergiert die Folge* $(E_n(x))_{n\in\mathbb{N}_0}$. *Wir erhalten also eine Funktion* $\exp\colon \mathbb{R} \to \mathbb{R}$ *mit*

$$\exp(x) = \lim_{n\to\infty} \big(E_n(x)\big) = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Diese Reihe konvergiert absolut; die obige Potenzreihe hat Konvergenzradius $R_0 = \infty$.

Beweis. Sei $x \in \mathbb{R}$ beliebig und $r := |x| \geqslant 0$. Nach Lemma 29.4 gibt es ein $n_0 \in \mathbb{N}_0$ mit $2|e_{m+1}(r)| \leqslant 1$ und $|E_n(r) - E_m(r)| \leqslant 2|e_{m+1}(r)| \leqslant 1$ für alle $n > m \geqslant n_0$. Für $n \geqslant n_0$ folgt:

$$E_n(r) = E_n(r) - E_{n_0}(r) + E_{n_0}(r) \leqslant |E_n(r) - E_{n_0}(r)| + E_{n_0}(r) \leqslant 1 + E_{n_0}(r).$$

Also ist die Folge $(E_n(r))_{n\in\mathbb{N}_0}$ nach oben beschränkt. Wegen $E_{n+1}(r) = E_n(r) + \frac{r^{n+1}}{(n+1)!} \geqslant E_n(r)$ für alle $n \in \mathbb{N}_0$ ist die Folge also nach dem Monotonie-Prinzip konvergent. D.h., die unendliche Reihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ ist absolut konvergent und damit auch konvergent. □

Die Funktion $\exp(x)$ ist fundamental für die Mathematik selbst und kommt gleichzeitig in unzähligen Anwendungen in allen möglichen Wissenschaften vor, vor allem bei der Beschreibung von Wachstumsprozessen; denken Sie etwa an Zinzeszinsrechnungen oder das bedrohliche "exponentielle Wachstum" zu Corona-Krisenzeiten. (Mehr dazu später in Beispiel 38.1.)

Beispiel 29.6. Für $x \in \mathbb{R}$ und $m \in \mathbb{N}_0$ mit $m \geqslant 2|x| - 2$ gilt $\big|\exp(x) - E_m(x)\big| \leqslant \dfrac{2|x|^{m+1}}{(m+1)!}$. (Dies folgt aus der Abschätzung für $|E_n(x) - E_m(x)|$ in Lemma 29.4, mit dem gleichen Argument wie in Bemerkung 29.3.) Mit Sage erhalten wir:

```
sage: def expreihe(x,nachkomma):
....:   m=0
....:   while m<2*abs(x)-2: m+=1
....:   while 2*abs(x)^(m+1)*10^(nachkomma+1)>factorial(m+1): m+=1
....:   return sum(x^k/factorial(k) for k in range(m+1))
sage: dec_exp(expreihe(1,64),64)        # e=exp(1) korrekt auf 64 Stellen
'2.7182818284590452353602874713526624977572470936999595749669676277'
sage: dec_exp(expreihe(-1,16))          # x=-1
'0.3678794411714423'
```

Satz 29.7 (***Funktionalgleichung***). *Für alle* $x, y \in \mathbb{R}$ *gilt* $\exp(x+y) = \exp(x) \cdot \exp(y)$.

Beweis. Seien $x, y \in \mathbb{R}$ fest. Für $n \in \mathbb{N}_0$ sei $x_n := \frac{x^n}{n!}$ und $y_n := \frac{y^n}{n!}$. Wir betrachten die beiden konvergenten Reihen $\exp(x) = \sum_{n=0}^{\infty} x_n$ und $\exp(y) = \sum_{n=0}^{\infty} y_n$; wie oben bemerkt, sind diese absolut konvergent. Mit dem Cauchy–Produkt für Reihen erhalten wir eine konvergente Reihe $\exp(x) \cdot \exp(y) = \sum_{n=0}^{\infty} w_n$, wobei $w_n = \sum_{k=0}^{n} x_n y_{n-k}$ für alle $n \in \mathbb{N}_0$ gilt. Nun ist

$$w_n = \sum_{k=0}^{n} \frac{x^k}{k!} \frac{y^{n-k}}{(n-k)!} = \sum_{k=0}^{n} \frac{1}{k!(n-k)!} x^k y^{n-k} = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \frac{1}{n!}(x+y)^n,$$

wobei wir den Binomischen Lehrsatz verwenden. Also ist $\sum_{n=0}^{\infty} w_n = \exp(x+y)$. □

Folgerung 29.8. *Sei* $x \in \mathbb{R}$. *Dann gilt* $\exp(x) > 0$*; ist* $x \geqslant 0$*, so gilt auch* $\exp(x) \geqslant 1 + x$. *Sind* $x, y \in \mathbb{R}$ *mit* $x < y$*, so gilt* $\exp(x) < \exp(y)$*, d.h., die Funktion* $\exp\colon \mathbb{R} \to \mathbb{R}$ *ist streng monoton wachsend und damit injektiv.*

Beweis. Sei zuerst $x \geqslant 0$. Dann ist $\exp(x) \geqslant E_n(x)$ für alle $n \in \mathbb{N}_0$. Für $n = 1$ ergibt dies $\exp(x) \geqslant 1 + x > 0$. Sei nun $x < 0$. Mit der Funktionalgleichung folgt $1 = \exp(0) = \exp(x - x) = \exp(x) \cdot \exp(-x)$ also $\exp(x) = \exp(-x)^{-1} > 0$. Seien nun $x, y \in \mathbb{R}$ mit $x < y$. Dann ist $y - x > 0$ und mit der Funktionalgleichung folgt $\exp(y) = \exp(x + (y - x)) = \exp(x) \cdot \exp(y - x)$. Nun ist $\exp(y - x) \geqslant 1 + (y - x) > 1$, also $\exp(y) > \exp(x)$. □

Ist $n \in \mathbb{N}$, so schreibe $n = 1 + 1 + \ldots + 1$ (mit n Summanden). Mit der obigen Funktionalgleichung folgt dann $\exp(n) = \exp(1)^n = e^n$. Die Funktion exp interpoliert also die Potenzen e^n für natürliche Exponenten $n \in \mathbb{N}$ und dehnt diese auf beliebige reelle Exponenten aus; daher der Name "Exponential-Funktion" und die Schreibweise $e^x := \exp(x)$ für $x \in \mathbb{R}$.

30. *Konvergenz in* $\mathbb{C}$*, und die Euler-Gleichung* $e^{x\boldsymbol{i}} = \cos(x) + \sin(x)\boldsymbol{i}$

Wir wollen nun auch $\exp(z)$ für $z \in \mathbb{C}$ definieren. Zunächst einige allgemeine Vorbereitungen.

Völlig analog zu Definition 27.1 können wir Grenzwerte von Folgen $(z_n)_{n\in\mathbb{N}_0}$ mit komplexen Zahlen $z_n \in \mathbb{C}$ definieren. Die Bedingung $(*)$ in Definition 27.1 wird exakt genauso formuliert, wobei wir den komplexen Absolutbetrag $|z| = \sqrt{z\bar{z}} \in \mathbb{R}$ (für $z \in \mathbb{C}$, siehe Bemerkung 11.1, Kapitel II) anstelle des reellen Absolutbetrags $|x|$ (für $x \in \mathbb{R}$) verwenden. Für diesen gilt ebenfalls die nützliche Dreiecksungleichung $|z + z'| \leqslant |z| + |z'|$ für alle $z, z' \in \mathbb{C}$. (Denn: Sei $z = a + bi$ mit $a, b \in \mathbb{R}$. Dann ist $|z| = \sqrt{a^2 + b^2}$ die Euklidische Norm des Spaltenvektors $\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$; die Dreiecksungleichung folgt also aus der Cauchy–Schwarz–Ungleichung in Satz 19.2, Kapitel IV.) Ist $z = a + bi \in \mathbb{C}$ mit $a, b \in \mathbb{R}$, so gilt $|a| = \sqrt{a^2} \leqslant \sqrt{a^2 + b^2} = |z|$ und analog $|b| \leqslant |z|$, d.h., die reellen Absolutbeträge des Realteils von $z \in \mathbb{C}$ und des Imaginärteils von $z \in \mathbb{C}$ sind stets kleiner oder gleich $|z|$; dies wird im Folgenden ständig benutzt werden. Beachte auch: Für $x \in \mathbb{R} \subseteq \mathbb{C}$ stimmen der Absolutbetrag in $\mathbb{R}$ und der komplexe Absolutbetrag überein; außerdem ist $|xi| = |x|$.

Lemma 30.1. *Sei* $(z_n)_{n\in\mathbb{N}_0}$ *eine Folge komplexer Zahlen. Für* $n \in \mathbb{N}$ *schreibe* $z_n = a_n + b_n i$ *mit* $a_n, b_n \in \mathbb{R}$. *Genau dann ist* $(z_n)_{n\in\mathbb{N}_0}$ *konvergent in* $\mathbb{C}$*, wenn sowohl* $(a_n)_{n\in\mathbb{N}_0}$ *als auch* $(b_n)_{n\in\mathbb{N}_0}$ *in* $\mathbb{R}$ *konvergieren. In diesem Fall gilt* $\lim\limits_{n\to\infty}(z_n) = \big(\lim\limits_{n\to\infty}(a_n)\big) + \big(\lim\limits_{n\to\infty}(b_n)\big)i$.

Beweis. Sei zuerst $(z_n)_{n\in\mathbb{N}_0}$ in $\mathbb{C}$ konvergent, mit Grenzwert $\gamma = \alpha + \beta i \in \mathbb{C}$, wobei $\alpha, \beta \in \mathbb{R}$. Sei $\varepsilon > 0$. Dann gibt es ein $n_0 \in \mathbb{N}_0$ mit $|z_n - \gamma| < \varepsilon$ für alle $n \geqslant n_0$. Nun ist $a_n - \alpha \in \mathbb{R}$ der Realteil von $z_n - \gamma \in \mathbb{C}$. Also folgt mit obiger Bemerkung $|a_n - \alpha| \leqslant |z_n - \gamma| < \varepsilon$ für alle

$n \geqslant n_0$. Analog ist $b_n - \beta$ der Imaginärteil von $z_n - \gamma$, also auch $|b_n - \beta| < \varepsilon$ für alle $n \geqslant n_0$. Also konvergiert $(a_n)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$ gegen α und $(b_n)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$ gegen β. — Sei jetzt umgekehrt vorausgesetzt, dass $(a_n)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$ konvergiert, mit Grenzwert $\alpha \in \mathbb{R}$, und $(b_n)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$ konvergiert, mit Grenzwert $\beta \in \mathbb{R}$. Wir wollen zeigen, dass dann $(z_n)_{n\in\mathbb{N}_0}$ in $\mathbb{C}$ konvergiert, mit Grenzwert $\gamma := \alpha + \beta i \in \mathbb{C}$. Sei $\varepsilon > 0$. Dann gibt es $n_1, n_2 \in \mathbb{N}_0$ mit $|a_n - \alpha| < \varepsilon/2$ für alle $n \geqslant n_1$ und $|b_n - \beta| < \varepsilon/2$ für alle $n \geqslant n_2$. Für $n \geqslant n_0 := \max\{n_1, n_2\}$ folgt dann

$$|z_n - \gamma| = |(a_n - \alpha) + (b_n - \beta)i| \leqslant |a_n - \alpha| + |(b_n - \beta)i| = |a_n - \alpha| + |b_n - \beta| < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

wobei wir die Dreiecksungleichung in $\mathbb{C}$ verwenden. Also konvergiert $(z_n)_{n\in\mathbb{N}_0}$ gegen γ. □

Folgerung 30.2. *Sei $(z_n)_{n\in\mathbb{N}_0}$ eine konvergente Folge komplexer Zahlen. Dann konvergiert die komplex-konjugierte Folge $(\bar{z}_n)_{n\in\mathbb{N}_0}$ in $\mathbb{C}$, mit $\lim_{n\to\infty}(\bar{z}_n) = \overline{\lim_{n\to\infty}(z_n)}$.*

Beweis. Sei $\gamma := \lim_{n\to\infty}(z_n) \in \mathbb{C}$ und schreibe $\gamma = \alpha + \beta i$ mit $\alpha, \beta \in \mathbb{R}$. Für $n \in \mathbb{N}_0$ schreibe $z_n = a_n + b_n i$ mit $a_n, b_n \in \mathbb{R}$. Nach Lemma 30.1 konvergiert $(a_n)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$ gegen α, und $(b_n)_{n\in\mathbb{N}_0}$ konvergiert in $\mathbb{R}$ gegen β. Nun ist $\bar{z}_n = a_n - b_n i$ für alle $n \in \mathbb{N}_0$. Also folgt wiederum aus Lemma 30.1, dass $(\bar{z}_n)_{n\in\mathbb{N}_0}$ in $\mathbb{C}$ konvergiert, mit Grenzwert $\alpha - \beta i = \bar{\gamma}$. □

Sei $(z_n)_{n\in\mathbb{N}_0}$ eine Folge komplexer Zahlen. Für $n \in \mathbb{N}_0$ sei $Z_n := z_0 + z_1 + \ldots + z_n \in \mathbb{C}$ die zugehörige Partialsumme. Völlig analog zu Definition 28.1 bezeichnen wir mit dem Symbol $\sum_{n=0}^{\infty} z_n$ die Folge $(Z_n)_{n\in\mathbb{N}_0}$ (egal ob diese konvergiert oder nicht); wenn $(Z_n)_{n\in\mathbb{N}_0}$ konvergiert, so bezeichnen wir wiederum den Grenzwert auch mit $\sum_{n=0}^{\infty} z_n$.
Für $n \in \mathbb{N}_0$ sei außerdem $Z_n^* := |z_0| + |z_1| + \ldots + |z_n| \in \mathbb{R}$ die Partialsumme der komplexen Absolutbeträge. Dann ist also $(Z_n^*)_{n\in\mathbb{N}_0}$ eine Folge reeller Zahlen. Wir sagen, dass die unendliche Reihe $\sum_{n=0}^{\infty} z_n$ ***absolut konvergiert***, wenn die Folge $(Z_n^*)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$ konvergiert. Insbesondere ist in diesem Fall die Folge $(Z_n^*)_{n\in\mathbb{N}_0}$ nach oben beschränkt.

Satz 30.3. *Sei $(z_n)_{n\in\mathbb{N}_0}$ eine Folge komplexer Zahlen. Es gebe eine Konstante $C \in \mathbb{R}$ mit $Z_n^* := |z_0| + |z_1| + \ldots + |z_n| \leqslant C$ für alle $n \in \mathbb{N}_0$. Dann konvergiert die Folge $(Z_n^*)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$ und die unendliche Reihe $\sum_{n=0}^{\infty} z_n$ konvergiert in $\mathbb{C}$.*

Beweis. Es gilt $Z_{n+1}^* = Z_n^* + |z_{n+1}| \geqslant Z_n^*$ für alle $n \in \mathbb{N}_0$. Wegen $Z_n^* \leqslant C$ folgt mit dem Monotonie-Prinzip, dass die Folge $(Z_n^*)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$ konvergiert. Für $n \in \mathbb{N}_0$ schreibe nun $z_n = a_n + b_n i$ mit $a_n, b_n \in \mathbb{R}$. Sei $A_n^* := |a_0| + |a_1| + \ldots + |a_n| \in \mathbb{R}$. Nun ist $|a_n| \leqslant |z_n|$; also folgt $A_n^* = |a_0| + |a_1| + \ldots + |a_n| \leqslant |z_0| + |z_1| + \ldots + |z_n| = Z_n^* \leqslant C$ für alle $n \in \mathbb{N}_0$. Wegen $A_{n+1}^* = A_n^* + |a_{n+1}| \geqslant A_n^*$ konvergiert die Folge $(A_n^*)_{n\in\mathbb{N}_0}$ also nach dem Monotonie-Prinzip, d.h., die unendliche Reihe $\sum_{n=0}^{\infty} a_n$ ist absolut konvergent in $\mathbb{R}$. Nach Satz 28.10(b) konvergiert sie auch im gewöhnlichen Sinne, es existiert also $\alpha := \sum_{n=0}^{\infty} a_n \in \mathbb{R}$, d.h., die Folge $(A_n)_{n\in\mathbb{N}_0}$ der Partialsummen $A_n := a_0 + a_1 + \ldots + a_n$ konvergiert in $\mathbb{R}$, mit Grenzwert α. Völlig analog zeigt man, dass auch $\beta := \sum_{n=0}^{\infty} b_n \in \mathbb{R}$ existiert, d.h., die Folge $(B_n)_{n\in\mathbb{N}_0}$ der Partialsummen $B_n := b_0 + b_1 + \ldots + b_n$ konvergiert in $\mathbb{R}$, mit Grenzwert β. Nach Lemma 30.1(a) konvergiert dann auch die Folge $(A_n + B_n i)_{n\in\mathbb{N}_0}$ in $\mathbb{C}$. Wegen $A_n + B_n i = z_0 + z_1 + \ldots + z_n = Z_n$ für alle $n \in \mathbb{N}_0$ bedeutet dies genau, dass die unendliche Reile $\sum_{n=0}^{\infty} z_n$ in $\mathbb{C}$ konvergiert. □

Bemerkung 30.4. Sei $(z_n)_{n\in\mathbb{N}_0}$ eine Folge komplexer Zahlen. Analog zum Fall reeller Potenzreihen sei $D \subseteq \mathbb{C}$ die Menge aller $z \in \mathbb{C}$ so dass die "***Potenzreihe***" $f(z) := \sum_{n=0}^{\infty} z_n z^n$ konvergiert. Wir erhalten also wieder eine Funktion $f\colon D \to \mathbb{C}$ mit $f(z) := \sum_{n=0}^{\infty} z_n z^n \in \mathbb{C}$ für $z \in D$. Es gilt dann folgende analoge Version von Satz 29.1 (mit einem ähnlichen Beweis): Es ist entweder $D = \mathbb{C}$ oder $D = \{0\}$ oder es gibt ein $R_0 \in \mathbb{R}$, $R_0 > 0$, so dass $K^\circ(R_0) \subseteq D \subseteq K(R_0)$, wobei $K(R_0) := \{u \in \mathbb{C} \mid |u| \leqslant R_0\}$ und $K^\circ(R_0) := \{u \in \mathbb{C} \mid |u| < R_0\}$. (Das Intervall $[-R_0, R_0] \subseteq \mathbb{R}$ in Satz 29.1 wird also durch den Kreis $K(R_0)$ mit Radius R_0 in der komplexen Zahlenebene ersetzt.)

Satz 30.5 (***Exponential-Funktion in*** $\mathbb{C}$)**.** *Für jedes $z \in \mathbb{C}$ konvergiert die unendliche Reihe $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ und wir setzen $\exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!} \in \mathbb{C}$. Wie zuvor gilt:*

(a) $\displaystyle \Big| \exp(z) - \sum_{k=0}^{m} \frac{z^k}{k!} \Big| \leqslant \frac{2|z|^{m+1}}{(m+1)!}$ *für jedes $m \in \mathbb{N}_0$ mit $m \geqslant 2|z| - 2$,*

(b) $\exp(z + z') = \exp(z) \cdot \exp(z')$ *für alle $z, z' \in \mathbb{C}$ (Funktionalgleichung).*

Beweis. Sei $z \in \mathbb{C}$ fest. Für $n \in \mathbb{N}_0$ setze $z_n := \frac{z^n}{n!}$ und $Z_n^* := |z_0| + |z_1| + \ldots + |z_n| = \sum_{k=0}^{n} \frac{|z|^k}{k!} \in \mathbb{R}$. Sei $m \in \mathbb{N}_0$ fest mit $m \geqslant 2|z| - 2$. Genauso wie in Lemma 29.4 zeigt man

$$\Big| \sum_{k=m+1}^{n} \frac{z^k}{k!} \Big| \leqslant \sum_{k=m+1}^{n} \frac{|z|^k}{k!} \leqslant C := 2\frac{|z|^{m+1}}{(m+1)!} \qquad \text{für alle } n > m.$$

Nun ist $Z_n^* \leqslant Z_m^*$ für $0 \leqslant n < m$. Für $n > m$ folgt mit obiger Abschätzung:

$$Z_n^* = \sum_{k=0}^{n} \frac{|z|^k}{k!} = \sum_{k=0}^{m} \frac{|z|^k}{k!} + \sum_{k=m+1}^{n} \frac{|z|^k}{k!} \leqslant \sum_{k=0}^{m} \frac{|z|^k}{k!} + C = Z_m^* + C.$$

Also gilt $Z_n^* \leqslant Z_m^* + C \in \mathbb{R}$ für alle $n \in \mathbb{N}_0$. Damit existiert $\exp(z) = \sum_{n=0}^{\infty} z_n \in \mathbb{C}$; siehe Satz 30.3. Nun folgt (a) genauso wie in Beispiel 29.6. Die Funktionalgleichung in (b) folgt genauso wie in Satz 29.7; die Aussage über das Cauchy-Produkt in Satz 28.13 gilt auch für absolut konvergente Reihen in $\mathbb{C}$, mit fast wörtlich dem gleichen Beweis. □

Bemerkung 30.6. Sei $z \in \mathbb{C}$. Dann gilt $\exp(\bar{z}) = \overline{\exp(z)}$. Denn: Für $n \in \mathbb{N}_0$ sei $z_n = \frac{z^n}{n!}$ und $Z_n = \sum_{k=0}^{n} z_k$. Dann ist $\bar{Z}_n = \sum_{k=0}^{n} \bar{z}_k = \sum_{k=0}^{n} \frac{\bar{z}^k}{k!}$, also $\exp(\bar{z}) = \sum_{n=0}^{\infty} \bar{z}_n$. Mit Folgerung 30.2 erhalten wir $\exp(\bar{z}) = \overline{\exp(z)}$.

Ab hier Woche 6 Nachdem $\exp(x)$ auf $\mathbb{C}$ fortgesetzt ist, erhalten wir nun auch eine "analytische" Definition für $\sin(x)$ und $\cos(x)$ (im Gegensatz zur "anschaulichen" Definition wie in §11, Kapitel II), aus der sich sofort einige Eigenschaften dieser Funktionen herleiten lassen.

Definition 30.7. Sei $x \in \mathbb{R}$. Dann betrachte $\exp(xi) \in \mathbb{C}$ und schreibe $\exp(xi) = a + bi \in \mathbb{C}$ mit $a, b \in \mathbb{R}$. Wir setzen $\cos(x) := a$ und $\sin(x) := b$. Damit erhalten wir Funktionen $\sin\colon \mathbb{R} \to \mathbb{R}$ und $\cos\colon \mathbb{R} \to \mathbb{R}$. Es gilt $\exp(xi) = \cos(x) + \sin(x)i$ für alle $x \in \mathbb{R}$.

Ebenso wie $\exp(x)$ spielen $\sin(x)$ und $\cos(x)$ eine fundamentale Rolle sowohl in der Mathematik selbst als auch in vielen Anwendungen, typischerweise bei der Beschreibung von periodischen Wellenbewegungen oder Schwingungen (mehr dazu in Kapitel VIII??).

Satz 30.8. *Für* $x \in \mathbb{R}$ *seien* $\sin(x)$ *und* $\cos(x)$ *wie oben definiert. Dann gilt:*
(a) $\sin(0) = 0$ *und* $\cos(0) = 1$*;*
(b) $\sin(x)^2 + \cos(x)^2 = 1$, $\sin(-x) = -\sin(x)$ *und* $\cos(-x) = \cos(x)$ *für alle* $x \in \mathbb{R}$*;*
(c) $\sin(x + y) = \sin(x)\cos(y) + \cos(x)\sin(y)$ *und* $\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y)$ *für alle* $x, y \in \mathbb{R}$*. (**Additionstheoreme**; siehe auch §11, Kapitel II.)*

Beweis. (a) Es gilt $1 = \exp(0) = \cos(0) + \sin(0)i$, also $\cos(0) = 1$ und $\sin(0) = 0$.
(b) Sei $z = xi \in \mathbb{C}$. Dann ist $\bar{z} = -xi$, also $\overline{\exp(xi)} = \exp(-xi)$; siehe Bemerkung 30.6. Damit folgt $\cos(x) - \sin(x)i = \overline{\cos(x) + \sin(x)i} = \overline{\exp(xi)} = \exp(-xi) = \cos(-x) + \sin(-x)i$, also $\cos(-x) = \cos(x)$ und $\sin(-x) = -\sin(x)$. Außerdem ergibt sich $1 = \exp(0) = \exp(xi) \cdot \exp(-xi) = \exp(xi) \cdot \overline{\exp(xi)} = |\exp(xi)|^2 = \cos(x)^2 + \sin(x)^2$.
(c) Es gilt $\exp\big((x+y)i\big) = \cos(x+y) + \sin(x+y)i$. Andererseits ist die linke Seite auch gleich $\exp(xi) \cdot \exp(yi) = \big(\cos(x) + \sin(x)i\big)\big(\cos(y) + \sin(y)i\big)$. Ausmultiplizieren und Vergleich von Real- und Imaginärteil ergibt die gewünschten Formeln. □

Satz 30.9 (***Reihenentwicklung für*** sin)**.** *Für alle* $x \in \mathbb{R}$ *gilt*

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \ldots,$$

und diese unendliche Reihe ist absolut konvergent. Es gilt

$$\Big|\sin(x) - \sum_{k=0}^{m} (-1)^k \frac{x^{2k+1}}{(2k+1)!}\Big| \leqslant \frac{2|x|^{2m+3}}{(2m+3)!} \qquad \text{für jedes } m \in \mathbb{N}_0 \text{ mit } m \geqslant |x| - 2.$$

Beweis. Sei $x \in \mathbb{R}$ und $z := xi$; dann ist $|z| = |x|$. Sei $m \in \mathbb{N}_0$ mit $m \geqslant |x| - 2$. Dann ist $2m + 2 \geqslant 2|x| - 2$ und mit Satz 30.5(a) folgt

$$\Big|\exp(xi) - \sum_{l=0}^{2m+2} \frac{(xi)^l}{l!}\Big| \leqslant \frac{2|x|^{2m+3}}{(2m+3)!}.$$

Sei $b(x) \in \mathbb{R}$ der Imaginärteil von $\sum_{l=0}^{2m+2} \frac{(xi)^l}{l!}$. Dann ist $\sin(x) - b(x)$ der Imaginärteil von $\exp(xi) - \sum_{l=0}^{2m+2} \frac{(xi)^l}{l!}$, also folgt

$$\big|\sin(x) - b(x)\big| \leqslant \Big|\exp(xi) - \sum_{l=0}^{2m+2} \frac{(xi)^l}{l!}\Big| \leqslant \frac{2|x|^{2m+3}}{(2m+3)!}.$$

Nun ist $i^l = \pm 1$ falls l gerade. Für $b(x)$ brauchen wir also nur ungerade l zu betrachten.

Sei $l = 2k+1$ mit $k \in \mathbb{N}_0$; dann ist $(xi)^l = (xi)^{2k+1} = i^{2k+1}x^{2k+1} = (-1)^k x^{2k+1} i$. Dies ergibt $b(x) = \sum_{k=0}^{m} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$ und damit die gewünschte Abschätzung. Daraus folgt dann auch die Konvergenz der Reihe, und mit Satz 29.1 die absolute Konvergenz. $\square$

Satz 30.10 (***Reihenentwicklung für*** cos)**.** *Für alle* $x \in \mathbb{R}$ *gilt*

$$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \ldots,$$

und diese unendliche Reihe ist absolut konvergent. Es gilt

$$\Big| \cos(x) - \sum_{k=0}^{m} \frac{x^{2k}}{(2k)!} \Big| \leqslant \frac{2|x|^{2m+2}}{(2m+2)!} \qquad \text{für jedes } m \in \mathbb{N}_0 \text{ mit } m \geqslant |x| - 3/2.$$

Beweis. Analog zum obigen Beweis. $\square$

Die obigen Abschätzungen können noch etwas verbessert werden (siehe §35), aber damit kann man jedenfalls die Werte $\sin(x)$ und $\cos(x)$ für alle $x \in \mathbb{R}$ sehr gut näherungsweise bestimmen. Was jetzt noch fehlt, ist der Zusammenhang mit der Kreiszahl $\pi \approx 3{,}1415\ldots$; zum Beispiel sollte $\sin(\pi) = \cos(\pi/2) = 0$ gelten, sowie $\sin(x + 2\pi) = \sin(x)$ und $\cos(x + 2\pi) = \cos(x)$! Der folgende Hilfssatz ist eine erste Vorbereitung dazu. (Wir betrachten das Intervall $[0, 2]$ weil $0 \leqslant \pi/2 \leqslant 2$.)

Lemma 30.11. (a) *Es gilt* $\cos(2) < 0$ *und* $\sin(x) > 0$ *für* $0 < x \leqslant 2$.
(b) *Für* $0 \leqslant x < y \leqslant 2$ *ist* $\cos(2) \leqslant \cos(y) < \cos(x) \leqslant 1$.

Beweis. (a) Mit Sage erhalten wir folgende Annäherung an den Wert $\cos(2)$:

```
sage: def cosreihe(x,nachkomma):
....:   m=0
....:   while m<abs(x)-3/2: m+=1
....:   while 2*abs(x)^(2*m+2)*10^(nachkomma+1)>factorial(2*m+2): m+=1
....:   return sum((-1)^k*x^(2*k)/factorial(2*k) for k in range(m+1))
sage: dec_exp(cosreihe(2,16))
'-0.4161468365471424'   # cos(2)<0 wegen 16 korrekten Nachkommastellen
```

Sei $q(x) := x - \frac{x^3}{3!}$. Nach Satz 30.9 gilt dann $|\sin(x) - q(x)| \leqslant 2|x|^5/5!$ für $1 \geqslant |x| - 2$, also $|x| \leqslant 3$. Für $0 < x \leqslant 2$ ist $q(x) = x(1 - x^2/6) \geqslant x(1 - 4/6) = x/3$. Andererseits $2|x|^5/5! = x(x^4)/60 \leqslant 16x/60 < x/3$, also $2|x|^5/5! < x/3$. Also folgt $\sin(x) > 0$.

(b) Setze $u := (x+y)/2$ und $v := (y-x)/2$. Dann gilt $0 < u, v < 2$; außerdem ist $x = u - v$, $y = u + v$ und mit dem Additionstheorem für cos folgt

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

wobei wir auch Satz 30.8(b) verwenden. Also folgt $\cos(y) - \cos(x) = \cos(u+v) - \cos(u-v) = -2\sin(u)\sin(v)$. Wegen $0 < u, v < 2$ ist $\sin(u) > 0$ und $\sin(v) > 0$; siehe (a). Also folgt

$\cos(y) - \cos(x) < 0$, d.h., $\cos(y) < \cos(x)$. Schließlich: Für $y = 2$ folgt $\cos(2) < \cos(x)$ und für $x = 0$ auch $\cos(y) < \cos(0) = 1$. □

Durchläuft $x \in \mathbb{R}$ das Intervall $[0, 2]$ von links nach rechts, so zeigt Lemma 30.11, dass $\cos(x)$ immer kleiner wird, von $\cos(0) = 1$ am Anfang bis $\cos(2) < 0$ am Ende des Intervalls. Es scheint anschaulich klar zu sein, dass irgendwo zwischen 0 und 2 der Graph der Funktion $y = \cos(x)$ die x-Achse schneiden muss, es also ein $\xi_0 \in (0, 2)$ gibt mit $\cos(\xi_0) = 0$. Dieses ξ_0 sollte dann genau $\pi/2$ sein ... (Mehr dazu im nächsten Kapitel; siehe dort Definition 31.9.)

Satz 30.12. *Nehmen wir an, es gibt tatsächlich* $\xi_0 \in (0, 2)$ *mit* $\cos(\xi_0) = 0$. *Dann gilt:*
(a) $\sin(\xi_0) = 1$, $\cos(2\xi_0) = -1$ *und* $\cos(x) = \sin(x + \xi_0)$ *für alle* $x \in \mathbb{R}$.
(b) $\cos\big((2n + 1)\xi_0\big) = 0$ *und* $\sin\big(2n\xi_0\big) = 0$ *für alle* $n \in \mathbb{Z}$.
(c) $\cos(x + 2\xi_0) = -\cos(x)$ *und* $\sin(x + 2\xi_0) = -\sin(x)$ *für alle* $x \in \mathbb{R}$.
(d) $\cos(x + 4\xi_0) = \cos(x)$ *und* $\sin(x + 4\xi_0) = \sin(x)$ *für alle* $x \in \mathbb{R}$.
(e) $\sin$ *ist streng monoton wachsend auf* $[-\xi_0, \xi_0]$; $\cos$ *ist streng monoton fallend auf* $[0, 2\xi_0]$.

Beweis. (a) Aus $\sin(\xi_0)^2 + \cos(\xi_0)^2 = 1$ folgt $\sin(\xi_0)^2 = 1$, also $\sin(\xi_0) = \pm 1$ und schließlich $\sin(\xi_0) = 1$; siehe Lemma 30.11(a). Mit dem Additionstheorem für cos folgt $\cos(2\xi_0) = \cos(\xi_0)^2 - \sin(\xi_0)^2 = 0 - 1 = -1$. Mit dem Additionstheorem für sin folgt $\sin(x + \xi_0) = \sin(x)\cos(\xi_0) + \cos(x)\sin(\xi_0) = \cos(x)$ für alle $x \in \mathbb{R}$.
(b) Zuerst zeigen wir die Aussage für $\sin(2n\xi_0)$. Wegen $\sin(-x) = -\sin(x)$ für alle $x \in \mathbb{R}$ genügt es, nur $n \in \mathbb{N}_0$ zu betrachten. Dazu verwenden wir vollständige Induktion nach n. Für $n = 0$ ist $\sin(0) = 0$. Sei nun $n \geqslant 0$ und angenommen, dass die Aussage bereits für n gilt. Mit dem Additionstheorem für sin folgt dann

$$\sin\big(2(n + 1)\xi_0\big) = \sin(2n\xi_0 + 2\xi_0) = \sin(2n\xi_0)\cos(2\xi_0) + \cos(2n\xi_0)\sin(2\xi_0).$$

Die rechte Seite ist gleich 0. Denn nach Induktion ist $\sin(2n\xi_0) = 0$; aus (a) (mit $x = \xi_0$) folgt auch $\sin(2\xi_0) = \sin(\xi_0 + \xi_0) = \cos(\xi_0) = 0$. Nun folgt die Aussage auch für $\cos\big((2n+1)\xi_0\big)$. Denn mit (a) ist dies gleich $\cos\big(2n\xi_0 + \xi_0\big) = \sin\big(2n\xi_0 + 2\xi_0\big) = \sin\big(2(n + 1)\xi_0\big) = 0$.
(c), (d) Mit (a), (b) und dem Additionstheorem für sin folgt $\sin(x+2\xi_0) = \sin(x)\cos(2\xi_0) + \cos(x)\sin(2\xi_0) = -\sin(x)$ für $x \in \mathbb{R}$. Mit $y := x+2\xi_0$ folgt $\sin(x+4\xi_0) = \sin(y+2\xi_0) = -\sin(y) = -\sin(x+2\xi_0) = -(-\sin(x)) = \sin(x)$. Die Aussagen für cos folgen mit (a).
(e) Nach Lemma 30.11 ist cos streng monoton fallend auf $[0, \xi_0]$. Wegen $\cos(x) = \cos(-x)$ für alle $x \in \mathbb{R}$ ist dann cos streng monoton wachsend auf $[-\xi_0, 0]$. Nach (c) ist $\cos(x+2\xi_0) = -\cos(x)$ für alle $x \in \mathbb{R}$; folglich ist cos streng monoton fallend auf $[\xi_0, 2\xi_0]$. Insgesamt ist also cos streng monoton fallend auf $[0, 2\xi_0]$. Mit (a) folgt die analoge Aussage für sin. □

Kapitel VII: Differential- und Integralrechnung

Wir verlassen nun den Bereich der Folgen und betrachten Funktionen $f\colon D \to \mathbb{R}$, wobei typischerweise D ein Intervall oder ganz $\mathbb{R}$ ist. In dieser Allgemeinheit gibt es ziemlich exotische oder ungewöhnliche Beispiele, etwa die Funktion $f_0\colon \mathbb{R} \to \mathbb{R}$ mit

$$f_0(x) := \begin{cases} 1 & \text{falls } x \in \mathbb{Q}, \\ 0 & \text{falls } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

Da zwischen zwei beliebigen reellen Zahlen stets eine rationale Zahl liegt, ist es praktisch unmöglich, sich den Graph von $y = f_0(x)$ in der (x, y)-Ebene anschaulich vorzustellen. Man braucht also Methoden, um solche exotischen Funktionen von denen zu unterscheiden, die interessante oder sinnvolle Eigenschaften haben, mit denen man weiterarbeiten kann und die in Anwendungen auch tatsächlich vorkommmen. Um dies präzise zu formulieren, hat sich wiederum der Grenzwertbegriff als entscheidend herausgestellt.

Wir führen gleich an dieser Stelle einige nützliche Bezeichnungen ein: Für $a \in \mathbb{R}$ sei

$$\mathbb{R}_{\geqslant a} := \{x \in \mathbb{R} \mid x \geqslant a\}, \qquad \mathbb{R}_{>a} := \{x \in \mathbb{R} \mid x > a\}, \qquad \text{auch } [a, +\infty) \text{ bzw. } (a, +\infty);$$
$$\mathbb{R}_{\leqslant a} := \{x \in \mathbb{R} \mid x \leqslant a\}, \qquad \mathbb{R}_{<a} := \{x \in \mathbb{R} \mid x < a\}, \qquad \text{auch } (-\infty, a] \text{ bzw. } (-\infty, a).$$

Wir bezeichnen diese Teilmengen von $\mathbb{R}$ ebenfalls als Intervalle.

31. *Stetige Funktionen*

Sei $f\colon D \to \mathbb{R}$ eine Funktion, mit $D \subseteq \mathbb{R}$. Ist $a \in D$, so interessiert man sich für das Verhalten der Funktionswerte $f(x)$ für $x \in D$ “in der Nähe” von a. Dies führt zu folgender Definition:

Definition 31.1. Gegeben seien eine Funktion $f\colon D \to \mathbb{R}$, und ein $a \in D$. Dann heißt f ***stetig*** im Punkt a, wenn für jede konvergente Folge $(x_n)_{n\in\mathbb{N}_0}$ mit Grenzwert a und $x_n \in D$ für alle $n \in \mathbb{N}_0$ auch die Folge der Funktionswerte $(f(x_n))_{n\in\mathbb{N}_0}$ konvergiert, und zwar immer mit Grenzwert $f(a)$. Ist f stetig in jedem $a \in D$, so sagen wir kurz, dass f ***stetig auf*** D ist.

[Bemerkung: Wir erwähnen noch eine äquivalente Version der Definition der Stetigkeit, die nicht auf Folgen Bezug nimmt. Sei also $f\colon D \to \mathbb{R}$ und $a \in D$. Genau dann ist f stetig in a, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt mit $|f(x) - f(a)| < \varepsilon$ für alle $x \in D$ mit $|x - a| < \delta$ (“ε-δ-***Kriterium” für Stetigkeit***). Anders formuliert: $\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D: \ |x - a| < \delta \ \Rightarrow\ |f(x) - f(a)| < \varepsilon$. (Siehe zum Beispiel §11, Satz 3 im 1. Band des Buches von Forster für einen Beweis dieser Äquivalenz. Wir werden dies hier nicht benötigen.)]

Beispiel 31.2. (a) Sei $D = \mathbb{R}$, $c \in \mathbb{R}$ fest und $f(x) := c$ für alle $x \in \mathbb{R}$. Dann ist f stetig auf $\mathbb{R}$. Denn ist $a \in \mathbb{R}$ und $(x_n)_{n\in\mathbb{N}_0}$ konvergent mit $\lim\limits_{n\to\infty}(x_n) = a$, so ist $f(x_n) = c$ für alle $n \in \mathbb{N}_0$, also $(f(x_n))_{n\in\mathbb{N}_0}$ automatisch konvergent, mit Grenzert $c = f(a)$. Genauso sieht man sofort, dass die Funktion $f\colon \mathbb{R} \to \mathbb{R}$ mit $f(x) := x$ für alle $x \in \mathbb{R}$ stetig auf ganz $\mathbb{R}$ ist.

(b) Sei $D = \mathbb{R} \setminus \{0\}$ und $f(x) = 1/x$ für alle $x \in D$. Dann ist f stetig auf D. Denn sei $0 \neq a \in \mathbb{R}$ und $(x_n)_{n\in\mathbb{N}_0}$ eine konvergente Folge mit $\lim_{n\to\infty}(x_n) = a$ und $x_n \neq 0$ für $n \in \mathbb{N}_0$. Nach Satz 27.8 ist die Folge der Funktionswerte $(1/x_n)_{n\in\mathbb{N}_0}$ konvergent, mit Grenzwert $1/a$.

(c) Sei $D := \mathbb{R}_{\geqslant 0}$ und $f\colon D \to \mathbb{R}$ definiert durch $f(x) := \sqrt{x}$ für $x \in D$. Dann ist f stetig auf D. (Siehe Aufgabe 4.6 im Buch von Glosauer, sowie die Lösung auf S. 372.)

(d) Sei $f\colon \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) := 1$ falls $x \geqslant 0$, und $f(x) := -1$ falls $x < 0$. Der Graph der Funktion macht also einen Sprung von -1 nach 1 an der Stelle $a = 0$. Behauptung: f ist nicht stetig in $a = 0$. Dazu betrachte einerseits die Null-Folge $(1/n)_{n\in\mathbb{N}}$. Dann ist $f(1/n) = 1$ für alle $n \in \mathbb{N}$, die Folge der Funktionswerte konvergiert also, mit Grenzwert 1. Betrachte andererseits die Null-Folge $(-1/n)_{n\in\mathbb{N}}$. Hier konvergiert die Folge der Funktionswerte ebenfalls, aber nun mit Grenzwert $-1 \neq 1$. Also ist f nicht stetig in $a = 0$. Auf ähnliche Weise sieht man, dass obiges f_0 in keinem Punkt $a \in \mathbb{R}$ stetig ist.

Sei $D \subseteq \mathbb{R}$ und seien $f, g\colon D \to \mathbb{R}$ Funktionen. Wir definieren die Summe $f + g\colon D \to \mathbb{R}$ durch $(f + g)(x) := f(x) + g(x)$ für alle $x \in D$; wir definieren das Produkt $f \cdot g\colon D \to \mathbb{R}$ durch $(f \cdot g)(x) := f(x)g(x)$ für alle $x \in D$. Ist $g(x) \neq 0$ für alle $x \in D$, so definieren den Quotienten $f/g\colon D \to \mathbb{R}$ durch $(f/g)(x) := f(x)/g(x)$ für alle $x \in D$.

Satz 31.3. *Mit obigen Bezeichnungen sei* $a \in D$. *Sind* f *und* g *stetig in* a, *so sind auch* $f + g$, $f \cdot g$ *und* f/g *stetig in* a.

Beweis. Dies folgt sofort aus den analogen Grenzwertregeln für Folgen in Satz 27.8. □

Beispiel 31.4. (a) Sei $D = \mathbb{R}$ und $f\colon \mathbb{R} \to \mathbb{R}$ eine ***Polynomfunktion***. Es gibt also $n \in \mathbb{N}_0$ und $a_0, a_1, \ldots, a_n \in \mathbb{R}$ mit $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ für alle $x \in \mathbb{R}$. Beachte dass f aufgebaut ist, mit Hilfe von Summen und Produkten, aus konstanten Funktionen und der Funktion $p\colon \mathbb{R} \to \mathbb{R}$ mit $p(x) = x$ für alle $x \in \mathbb{R}$. Weil diese beiden Typen von Funktionen stetig sind (siehe Beispiel 31.2(a)), folgt sofort aus Satz 31.3 (ohne jede weitere Rechnung), dass auch f auf $\mathbb{R}$ stetig ist.

(b) Sei $g\colon \mathbb{R} \to \mathbb{R}$ eine weitere Polynomfunktion. Es gibt also $m \in \mathbb{N}_0$ und $b_0, b_1, \ldots, b_m \in \mathbb{R}$ mit $g(x) = b_0 + b_1x + b_2x^2 + \cdots + b_mx^m$ für alle $x \in \mathbb{R}$. Nehmen wir an, es sei $m \geqslant 1$ und $b_m \neq 0$. Dann hat g höchstens m Nullstellen (siehe Folgerung 9.2, Kapitel II); sei $D' := \{x \in \mathbb{R} \mid g(x) \neq 0\}$. Mit (a) und Satz 31.3 folgt, dass $f/g\colon D' \to \mathbb{R}$ stetig ist; ein solcher Quotient von Polynomfunktionen heißt eine ***rationale Funktion***.

Sind zum Beispiel $f(x) = 2x^3 - 7x + 1$ und $g(x) = x^2 - 1$, so ist $D' = \mathbb{R} \setminus \{\pm 1\}$ und wir erhalten die rationale Funktion $f/g\colon D' \to \mathbb{R}$ mit $(f/g)(x) = \dfrac{2x^3 - 7x + 1}{x^2 - 1}$ für alle $x \in D'$.

Satz 31.5. *Die Funktionen* exp, sin *und* cos *sind stetig auf ganz* $\mathbb{R}$.

Beweis. Wir beginnen mit exp. Betrachte zuerst $a = 0$. Nach Beispiel 29.6 (mit $m = 0$) ist $|\exp(x) - 1| \leqslant 2|x|$ für alle $x \in \mathbb{R}$ mit $0 = m \geqslant 2|x| - 2$, also $|x| \leqslant 1$. Sei $(x_n)_{n\in\mathbb{N}_0}$ eine Null-Folge. Dann gibt es ein $n_0 \in \mathbb{N}_0$ mit $|x_n| \leqslant 1$ für alle $n \geqslant n_0$. Nun ist $0 \leqslant |\exp(x_n) - 1| \leqslant 2|x_n|$ für alle $n \geqslant n_0$. Also folgt $\lim\limits_{n\to\infty}(\exp(x_n)) = 1$. Sei jetzt $a \in \mathbb{R}$ beliebig und $(x_n)_{n\in\mathbb{N}_0}$ eine konvergente Folge mit Grenzwert a; schreibe $h_n := x_n - a$ für $n \in \mathbb{N}_0$; dann ist $(h_n)_{n\in\mathbb{N}_0}$ eine Null-Folge. Mit der Funktionalgleichung in Satz 29.7 folgt $\exp(x_n)-\exp(a) = \exp(a+h_n)-\exp(a) = \exp(a)\exp(h_n)-\exp(a) = \exp(a)\big(\exp(h_n)-1\big)$ für alle $n \in \mathbb{N}_0$. Wir haben gerade zuvor gesehen, dass $(\exp(h_n))_{n\in\mathbb{N}_0}$ gegen 1 konvergiert; also ist $\lim\limits_{n\to\infty}(\exp(x_n)) = \exp(a)$ und damit exp stetig in a.
Nun zu sin und cos, zuerst wieder für $a = 0$. Nach Satz 30.10 (mit $m = 0$) ist $|\cos(x)-1| \leqslant |x|^2$ für $x \in \mathbb{R}$ mit $0 = m \geqslant |x|-3/2$, also $|x| \leqslant 3/2$. Genau wie oben folgt die Stetigkeit in $a = 0$; beachte $\cos(0) = 1$. Nach Satz 30.9 (mit $m = 0$) gilt $|\sin(x) - x| \leqslant |x|^3/3$ für alle $x \in \mathbb{R}$ mit $0 = m \geqslant |x| - 2$, also $|x| \leqslant 2$. Genau wie oben folgt analog die Stetigkeit in $a = 0$; beachte $\sin(0) = 0$. Sei jetzt $a \in \mathbb{R}$ beliebig und $(x_n)_{n\in\mathbb{N}_0}$ eine konvergente Folge mit Grenzwert a; schreibe wieder $h_n := x_n - a$ für $n \in \mathbb{N}_0$. Mit dem Additionstheorem für cos folgt

$$\begin{aligned}\cos(x_n) - \cos(a) &= \cos(a + h_n) - \cos(a) = \cos(a)\cos(h_n) - \sin(a)\sin(h_n) - \cos(a)\\ &= \cos(a)\big(\cos(h_n) - 1\big) - \sin(a)\sin(h_n).\end{aligned}$$

Wie oben folgt, dass die Folge $(\cos(x_n))_{n\in\mathbb{N}_0}$ gegen $\cos(a)$ konvergiert. Also ist cos auch stetig in a. Die Stetigkeit von sin folgt analog, mit dem Additionstheorem für sin. □

Bemerkung 31.6. Seien $f\colon D \to \mathbb{R}$ und $g\colon E \to \mathbb{R}$ Funktionen. Gilt $f(D) \subseteq E$, so können wir die Hintereinanderausführung $g \circ f\colon D \to \mathbb{R}$ bilden, mit $(g \circ f)(x) := g(f(x))$ für alle $x \in D$. Ist f stetig in $a \in D$ und g stetig in $f(a) \in E$, so ist $g \circ f$ auch stetig in a.
Sei zum Beispiel $f = \cos\colon [0, \xi_0] \to \mathbb{R}$ (mit $\xi_0 \in (0, 2)$ wie in Satz 30.12) und $g\colon \mathbb{R} \setminus \{0\} \to \mathbb{R}$ gegeben durch $g(x) = 1/x$ für $x \in \mathbb{R} \setminus \{0\}$. Nach Beispiel 31.2(b) ist g stetig auf $\mathbb{R} \setminus \{0\}$. Nun ist $\cos(\xi_0) = 0$ und $\cos(x) > 0$ für $0 < x < \xi_0$. Setzen wir $D := [0, \xi_0)$, so gilt $\cos(D) \subseteq \mathbb{R}\setminus\{0\}$ und wir können $g\circ f\colon D \to \mathbb{R}$ bilden, mit $(g\circ f)(x) = 1/\cos(x)$ für alle $x \in D$. Nach Satz 31.5 ist cos stetig in jedem $a \in \mathbb{R}$. Also ist $g \circ f$ auch stetig in jedem $a \in D$.

Wir kommen nun zu zwei zentralen Eigenschaften von stetigen Funktionen. Wie bereits Beispiel 31.2(d) zeigt, sollte der Graph einer stetigen Funktion keine Lücken oder Sprünge haben. Der folgende Satz macht diese Aussage präziser.

Satz 31.7 (***Zwischenwertsatz***). *Sei $f\colon D \to \mathbb{R}$ stetig, wobei $D = [a, b]$ ein abgeschlossenes Intervall ist (mit $a, b \in \mathbb{R}$, $a < b$). Sei $c \in \mathbb{R}$ mit $f(a) < c < f(b)$ oder $f(a) > c > f(b)$. Dann gibt es ein $x \in (a, b)$ mit $f(x) = c$.*

Beweis. Sei $f(a) < c < f(b)$. (Das Argument für $f(a) > c > f(b)$ ist analog.) Wir benutzen die sogenannte "***Intervallhalbierungs-Methode***". Dazu definieren wir rekursiv abgeschlossene Intervalle $I_n = [a_n, b_n] \subseteq \mathbb{R}$ für alle $n \in \mathbb{N}_0$, mit $a_n < b_n$, $f(a_n) \leqslant c < f(b_n)$ und $\ell(I_n) = (b-a)/2^n$. Für $n = 0$ sei $I_0 := [a, b]$. Sei nun $n \geqslant 0$ und $I_n = [a_n, b_n]$ bereits definiert. Wir teilen I_n in zwei gleiche Teilintervalle auf, also $I_n = I' \cup I''$ mit $I' = [a_n, d]$ und $I'' = [d, b_n]$, wobei $d := (a_n + b_n)/2$. Ist $f(d) \leqslant c$, so setze $I_{n+1} = [a_{n+1}, b_{n+1}]$ mit $a_{n+1} := d$ und $b_{n+1} := b_n$; ist $f(d) > c$, so setze $I_{n+1} = [a_{n+1}, b_{n+1}]$ mit $a_{n+1} := a_n$ und $b_{n+1} := d$. Nach Konstruktion ist $I_0 \supseteq I_1 \supseteq I_2 \supseteq \ldots$; wegen $\ell(I_n) = (b-a)/2^n$ liegt also eine Intervallschachtelung vor. Nach Satz 10.9 (Kapitel II) gibt es genau ein $x \in \mathbb{R}$ mit $x \in I_n$ für alle $n \in \mathbb{N}_0$. Nach Beispiel 27.14 konvergieren $(a_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$, beide mit Grenzwert x. Wegen der Stetigkeit von f in x konvergieren $(f(a_n))_{n\in\mathbb{N}_0}$ und $(f(b_n))_{n\in\mathbb{N}_0}$, beide mit Grenzwert $f(x)$. Wegen $f(a_n) \leqslant c$ für alle $n \in \mathbb{N}_0$ ist $f(x) \leqslant c$; wegen $f(b_n) > c$ für alle $n \in \mathbb{N}_0$ ist $f(x) \geqslant c$ (siehe Satz 27.8(4)). Also gilt $f(x) = c$. □

Tabelle 3. Sage–Programm zur näherungsweisen Zwischenwertbestimmung

```
def f0(f,a,b,c,eps):          # Betrachte Fall f(a)<c<f(b)
  n=0; eps1=(b-a)/eps         # n = Anzahl Halbierungen
  while 2^n<eps1:
    d=(a+b)/2                 # Halbierungs-Schritt
    if f(d)<=c:               # neues Intervall [d,b]
      a=d
    else:                     # neues Intervall [a,d]
      b=d
    n+=1
  return d
```

Bemerkung 31.8. Ist $f\colon [a, b] \to \mathbb{R}$ eine konkret gegebene stetige Funktion, die sich numerisch gut berechnen lässt (zum Beispiel eine Polynomfunktion), so liefert der obige Beweis sogar einen Algorithmus, um mit beliebig vorgegebener Präzision $\varepsilon > 0$ ein $x \in (a, b)$ mit $f(x) = c$ näherungsweise zu bestimmen. Dieser konvergiert sehr gut, weil bei der wiederholten Halbierung die Intervalle sehr schnell sehr klein werden. Sei etwa $f(x) = x^2 - 2$. Dann ist $f(0) = -2$ und $f(2) = 2$, also gibt es ein $x_0 \in (0, 2)$ mit $f(x_0) = 0$. Mit dem einfachen Sage-Programm in Tabelle 3 ergibt sich eine Annäherung $x_0 \approx \sqrt{2}$.

```
sage: def f(x): return x^2-2
sage: dec_exp(f0(f,0,2,0,10^-8))    # a=0, b=2, c=0, eps=10^-8
'1.4142135614529252
sage: dec_exp(f0(f,0,2,0,10^-17))   # a=0, b=2, c=0, eps=10^-17
'1.4142135623730950'                # sqrt(2) = 1.414213562373095048...
```

Ab hier Woche 7

Definition 31.9 (***Analytische Definition von*** π)**.** Da die Funktion cos streng monoton fallend und stetig auf $[0, 2]$ ist (siehe Lemma 30.11 und Satz 31.5), mit $\cos(0) = 1$ und

$\cos(2) < 0$, gibt es nach dem Zwischenwertsatz genau ein $\xi_0 \in (0, 2)$ mit $f(\xi_0) = 0$ (wie in Satz 30.12). Wir setzen dann $\boxed{\pi := 2\xi_0}$. Definieren wir mit Hilfe der Sage Funktion `cosreihe` (siehe Beweis von Lemma 30.11) eine Funktion `p(x)`, um die Werte von $-\cos(x)$ bis auf 32 Nachkommastellen genau zu bestimmen, so erhalten wir auch eine Annäherung an π:

```
sage: def p(x): return -cosreihe(x,32)
sage: dec_exp(2*f0(p,0,2,0,10^-33),32)
'3.14159265358979323846264338327950'   # alle Nachkommastellen korrekt
```

In §34 wird gezeigt, dass das so definierte π wirklich der Flächeninhalt eines Kreises mit Radius 1 ist; siehe auch Beispiel 31.13 weiter unten.

Satz 31.10 (***Extremwertsatz von Weierstraß***). *Sei $f\colon [a, b] \to \mathbb{R}$ stetig. Dann ist die Menge $S := \{f(x) \mid x \in [a, b]\}$ nach oben und nach unten beschränkt. Es gibt $x_1, x_2 \in [a, b]$ mit $f(x_1) \leqslant f(x) \leqslant f(x_2)$ für alle $x \in [a, b]$, also $f(x_1) = \inf(S)$ und $f(x_2) = \sup(S)$.*

Beweis. Annahme, S wäre nicht nach oben beschränkt. Analog zum obigen Beweis liefert die Intervallhalbierungs-Methode eine Folge von Intervallen $I_0 = [a, b] \supseteq I_1 \supseteq I_2 \supseteq \ldots$ mit $\ell(I_n) = (b - a)/2^n$ und so, dass $\{f(x) \mid x \in I_n\}$ nicht nach oben beschränkt ist für alle $n \in \mathbb{N}_0$. Schreiben wir $I_n = [a_n, b_n]$ wie oben, so konvergieren $(a_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$ gegen ein $x \in [a, b]$. In jedem I_n gibt es ein $c_n \in I_n$ mit $f(c_n) \geqslant n$. Nun ist $a_n \leqslant c_n \leqslant b_n$, also auch $x = \lim\limits_{n\to\infty}(c_n)$. Wegen der Stetigkeit von f folgt $\lim\limits_{n\to\infty}(f(c_n)) = f(x) < \infty$; Widerspruch, da $f(c_n) \to +\infty$ für $n \to \infty$. Also ist S nach oben beschränkt; sei $M := \sup(S)$. Annahme, es wäre $f(x) < M$ für alle $x \in [a, b]$. Dann betrachte $g(x) := 1/(M - f(x))$ für $x \in [a, b]$. Nun ist g auch stetig, also nach oben beschränkt; sei $K \in \mathbb{R}$ mit $g(x) \leqslant K$ für alle $x \in [a, b]$. Es folgt $0 < 1/(M - f(x)) \leqslant K$, also $f(x) \leqslant M' := M - 1/K < M$ für alle $x \in [a, b]$, Widerspruch zu $M = \sup(S)$. Also gibt es ein $x_2 \in [a, b]$ mit $f(x_2) = M$.
Analog findet man $x_1 \in [a, b]$ mit $f(x_1) = \inf(S)$. □

Allein mit Hilfe des Zwischenwertsatzes ist es oft möglich, "globale" Aussagen über den Gesamt-Wertebereich einer (stetigen) Funktion zu treffen.

Beispiel 31.11. (a) Es gilt $\exp(\mathbb{R}) = \mathbb{R}_{>0}$. Denn: Nach Folgerung 29.8 ist $\exp(\mathbb{R}) \subseteq \mathbb{R}_{>0}$. Sei nun $c \in \mathbb{R}_{>0}$ beliebig. Wir müssen zeigen, dass es ein $x \in \mathbb{R}$ gibt mit $\exp(x) = c$. Ist $c = 1$, so gilt $\exp(0) = 1 = c$. Sei nun $c > 1$ und betrachte das Intervall $I = [0, c]$. Nach Satz 31.5 ist f stetig auf I. Nach Folgerung 29.8 ist $\exp(c) \geqslant 1 + c > c > 1 = \exp(0)$. Nach Satz 31.7 gibt es also ein $x \in (0, c)$ mit $f(x) = c$. Sei nun $0 < c < 1$. Dann ist $1/c > 1$ also gibt es ein $x \in \mathbb{R}$ mit $\exp(x) = 1/c$. Nun ist mit der Funktionalgleichung $1 = \exp(0) = \exp(x - x) = \exp(x)\exp(-x)$, also $\exp(-x) = 1/\exp(x) = c$.

(b) Es gilt $\sin(\mathbb{R}) = \cos(\mathbb{R}) = [-1, 1]$. Denn: Wegen $\cos(x)^2 + \sin(x)^2 = 1$ für alle $x \in \mathbb{R}$ ist $\cos(\mathbb{R}) \subseteq [-1, 1]$ und $\sin(\mathbb{R}) \subseteq [-1, 1]$. Nun gilt $\cos(0) = 1$ und $\cos(\pi) = -1$; siehe Satz 30.8(a) und Satz 30.12(a). Für $-1 < c < 1$ gibt es also wieder ein $x \in (0, \pi)$ mit $\cos(x) = c$. Die Aussage für sin folgt dann mit Satz 30.12(a).

Definition 31.12. Nach Lemma 29.8 ist $\exp\colon \mathbb{R} \to \mathbb{R}$ streng monoton wachsend, also injektiv. Nach obigem Beispiel gilt außerdem $\exp(\mathbb{R}) = \mathbb{R}_{>0}$. Also ist exp eine Bijektion von $\mathbb{R}$ auf $\mathbb{R}_{>0}$. Analog folgt mit Satz 30.12 und obigem Beispiel, dass sin eine Bijektion von $[-\frac{\pi}{2}, \frac{\pi}{2}]$ auf $[-1, 1]$ ist, sowie cos eine Bijektion von $[0, \pi]$ auf $[-1, 1]$. Es existieren also jeweils die Umkehrfunktionen. Wir bezeichnen:

$\log := \exp^{-1}\colon \mathbb{R}_{>0} \to \mathbb{R}$, "***natürlicher Logarithmus***",
$\arcsin := \sin^{-1}\colon [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$, "***Arcus-Sinus***",
$\arccos := \cos^{-1}\colon [-1, 1] \to [0, \pi]$, "***Arcus-Cosinus***".

Man kann dann auch zeigen, dass diese ebenfalls stetig sind; siehe zum Beispiel §12, Satz 1 im 1. Band des Buches von Forster.

Beispiel 31.13. In der reellen Ebene sei $K_1 := \{(a, b) \in \mathbb{R} \times \mathbb{R} \mid a^2 + b^2 = 1\}$ der "***Einheitskreis***". Behauptung: $\gamma\colon [0, 2\pi) \to K_1,\ t \mapsto (\cos(t), \sin(t))$, ist eine Bijektion.

Dazu: Für $t \in \mathbb{R}$ gilt $\sin(t)^2 + \cos(t)^2 = 1$ und damit $\gamma(t) \in K_1$; also ist $\gamma([0, 2\pi)) \subseteq K_1$. Sei umgekehrt $(a, b) \in K_1$ beliebig. Ist $b \geqslant 0$, so folgt leicht mit den obigen Regeln, dass es genau ein $t \in [0, \pi]$ gibt mit $a = \cos(t)$. Dann ist $b^2 = 1 - a^2 = 1 - \cos(t)^2 = \sin(t)^2$, also $b = \pm\sin(t)$; wegen $b \geqslant 0$ und $\sin(t) \geqslant 0$ für $t \in [0, \pi]$ folgt automatisch $b = \sin(t)$. Für $t \in [\pi, 2\pi]$ gilt $\sin(t) \leqslant 0$, und $\cos\colon [\pi, 2\pi] \to [-1, 1]$ ist ebenfalls bijektiv. Ist $b < 0$, so folgt damit analog, dass es genau ein $t \in (\pi, 2\pi)$ gibt mit $a = \cos(t)$ und $b = \sin(t)$.

Durchläuft t das Intervall $[0, 2\pi)$, so durchläuft $(\cos(t), \sin(t))$ alle Punkte von K_1, und zwar gegen den Uhrzeigersinn beginnend mit $(1, 0)$ (für $t = 0$), dann z.B. $(0, 1)$ (für $t = \pi/2$), danach $(-1, 0)$ (für $t = \pi$) und so weiter; für $t = 2\pi$ erhalten wir wieder $(1, 0)$. — Für $(a, b) \in K_1$ heißt das eindeutige $t \in [0, 2\pi)$ mit $\gamma(t) = (a, b)$ der ***Winkel im Bogenmaß*** (oder auch ***Radiant***) zwischen der x-Achse (also der Geraden durch $(0, 0)$ und $(1, 0)$), und der Geraden durch $(0, 0)$ und (a, b); man kann dieses t tatsächlich als die Länge des Kreisbogens zwischen $(1, 0) \in K_1$ und $(a, b) \in K_1$ interpretieren, wie wir in §35 noch sehen werden.

Definition 31.14. Sei $f\colon D \to \mathbb{R}$. Analog zur bestimmten Divergenz für Folgen (siehe Definition 27.16) führen wir folgende Bezeichnungen ein:

(a) Ist $c \in \mathbb{R}$ und D nicht nach oben beschränkt, so schreiben wir $\lim\limits_{x \to +\infty} f(x) = c$, falls für jede bestimmt gegen $+\infty$ divergente Folge $(x_n)_{n \in \mathbb{N}_0}$ mit $x_n \in D$ für alle $n \in \mathbb{N}_0$ die Folge der Werte $(f(x_n))_{n \in \mathbb{N}_0}$ konvergiert mit Grenzwert c. (Analog wird $\lim\limits_{x \to -\infty} f(x) = c$ definiert.)

(b) Ist D nicht nach oben beschränkt, so schreiben wir $\lim\limits_{x\to+\infty} f(x) = +\infty$, falls für jede bestimmt gegen $+\infty$ divergente Folge $(x_n)_{n\in\mathbb{N}_0}$ mit $x_n \in D$ für alle $n \in \mathbb{N}_0$ die Folge der Werte $(f(x_n))_{n\in\mathbb{N}_0}$ bestimmt divergent gegen $+\infty$ ist. (Analog werden $\lim\limits_{x\to\pm\infty} f(x) = \pm\infty$ für alle Kombinationen der Vorzeichen definiert.)

Beispiel 31.15. (a) Betrachte die Funktion $\exp\colon \mathbb{R} \to \mathbb{R}$. Dann ist

$$\lim_{x\to+\infty} \exp(x) = +\infty \qquad \text{und} \qquad \lim_{x\to-\infty} \exp(x) = 0.$$

Denn: Sei $(x_n)_{n\in\mathbb{N}_0}$ bestimmt divergent gegen $+\infty$; dann ist $\exp(x_n) \geqslant 1 + x_n$ für alle $n \in \mathbb{N}_0$ und $x_n \geqslant 0$ (siehe Folgerung 29.8), also die Folge $(\exp(x_n))_{n\in\mathbb{N}_0}$ auch bestimmt divergent gegen $+\infty$. Also gilt $\lim\limits_{x\to+\infty} \exp(x) = +\infty$. Sei nun $(x_n)_{n\in\mathbb{N}_0}$ bestimmt divergent gegen $-\infty$. Dann ist $(-x_n)_{n\in\mathbb{N}_0}$ bestimmt divergent gegen $+\infty$, also auch $(\exp(-x_n))_{n\in\mathbb{N}_0}$ bestimmt divergent gegen $+\infty$ (wie wir gerade gesehen haben). Nach Lemma 27.17 ist dann $(1/\exp(-x_n))_{n\in\mathbb{N}_0}$ eine Null-Folge. Aus der Funktionalgleichung folgt aber $\exp(x_n)\exp(-x_n) = 1$, also ist auch $(\exp(x_n))_{n\in\mathbb{N}_0}$ eine Null-Folge und damit $\lim\limits_{x\to-\infty} \exp(x) = 0$.

(b) Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine Polynomfunktion, also $f(x) = a_d x^d + a_{d-1}x^{d-1} + \ldots + a_1 x + a_0$ für alle $x \in \mathbb{R}$, wobei $d \geqslant 1$ und $a_0, a_1, \ldots, a_d \in \mathbb{R}$. Nehmen wir $a_d = 1$ an. Dann ist

$$\lim_{x\to+\infty} f(x) = +\infty \qquad \text{und} \qquad \lim_{x\to-\infty} f(x) = \begin{cases} +\infty & \text{falls } d \text{ gerade,} \\ -\infty & \text{falls } d \text{ ungerade.} \end{cases}$$

Denn: Für $x \neq 0$ ist $f(x) = x^d g(x)$ mit $g(x) := 1 + \frac{a_{d-1}}{x} + \frac{a_{d-2}}{x^2} + \ldots + \frac{a_0}{x^d}$. Für $x \geqslant C := \max\{1, 2d|a_1|, \ldots, 2d|a_{d-1}|\}$ gilt dann $g(x) \geqslant 1/2$, also $f(x) \geqslant x^d/2 \geqslant x/2$. Sei nun $(x_n)_{n\in\mathbb{N}_0}$ eine bestimmt gegen $+\infty$ divergente Folge. Dann gibt es ein $n_0 \in \mathbb{N}_0$ mit $x_n \geqslant C$ für alle $n \geqslant n_0$, also $f(x_n) \geqslant x_n/2$ für $n \geqslant n_0$. Damit ist $\lim\limits_{x\to+\infty} f(x) = +\infty$. Um die Behauptung für $x \to -\infty$ zu zeigen, setze $h(x) := x^d - a_{d-1}x^{d-1} + \ldots + (-1)^{d-1}a_1 x + (-1)^d a_0$ für alle $x \in \mathbb{R}$; dann gilt $f(-x) = (-1)^d h(x)$. Damit folgt $\lim\limits_{x\to-\infty} f(x) = \lim\limits_{x\to+\infty} f(-x) = \lim\limits_{x\to+\infty} (-1)^d h(x)$, und die rechte Seite ist gleich $+\infty$ (falls d gerade) bzw. gleich $-\infty$ (falls d ungerade).

Damit erhält man den folgenden Spezialfall des ***Fundamentalsatzes der Algebra***.

Folgerung 31.16. *Sei $f\colon \mathbb{R} \to \mathbb{R}$ eine Polynomfunktion wie in Beispiel 31.15(b). Ist d ungerade, so gibt es ein $x \in \mathbb{R}$ mit $f(x) = 0$.*

Beweis. Wegen $\lim\limits_{x\to+\infty} f(x) = +\infty$ gibt es ein $b > 0$ mit $f(b) \geqslant 1$. Wegen $\lim\limits_{x\to-\infty} f(x) = -\infty$ gibt es ein $a < 0$ mit $f(a) \leqslant -1$. Die Behauptung folgt also aus dem Zwischenwertsatz. $\square$

32. ***Differenzierbare Funktionen***

Im letzten Abschnitt haben wir gesehen, dass der technisch klingende Begriff der Stetigkeit genau richtig ist, um nützliche Aussagen über Funktionen und ihre Werte herleiten zu können.

Umso mehr trifft dies auf den Begriff der Differenzierbarkeit zu, den wir jetzt einführen wollen. Zuerst noch die folgende allgemeine Definition.

Definition 32.1. Sei $f\colon D \to \mathbb{R}$ und $a \in \mathbb{R}$ (also nicht unbedingt $a \in D$).
Sei $c \in \mathbb{R}$. Wir schreiben $\lim_{x\to a} f(x) = c$ (oder auch $f(x) \to c$ für $x \to a$), wenn für jede konvergente Folge $(x_n)_{n\in\mathbb{N}_0}$ mit Grenzwert a und $x_n \in D$ für alle $n \in \mathbb{N}_0$ auch die Folge der Funktionswerte $(f(x_n))_{n\in\mathbb{N}_0}$ konvergiert, und zwar immer mit $\lim_{n\to\infty}(f(x_n)) = c$. In diesem Fall heißt c der ***Grenzwert*** von $f(x)$ für $x \to a$.
Ist $a \in D$, so gilt insbesondere: $\quad f$ stetig in $a \quad\Leftrightarrow\quad \lim_{x\to a} f(x) = f(a)$.
Außerdem schreiben wir $\lim_{x\to a} f(x) = +\infty$, wenn für jede konvergente Folge $(x_n)_{n\in\mathbb{N}_0}$ mit Grenzwert a und $x_n \in D$ für alle $n \in \mathbb{N}_0$ die Folge der Werte $(f(x_n))_{n\in\mathbb{N}_0}$ bestimmt divergent gegen $+\infty$ ist. (Analog wird $\lim_{x\to a} f(x) = -\infty$ definiert.)

Beachte: Damit diese Definitionen wirklich sinnvoll sind (vor allem falls $a \notin D$), müssen wir voraussetzen, dass es mindestens eine konvergente Folge $(x_n)_{n\in\mathbb{N}_0}$ gibt mit Grenzwert a und $a \neq x_n \in D$ für alle $n \in \mathbb{N}_0$; ein solcher Punkt a heißt ***Häufungspunkt*** von D. (Zum Beispiel ist jede reelle Zahl in einem Intervall ein Häufungspunkt dieses Intervalls; ist $D = (a, b]$ mit $a < b$, so ist $a \notin D$ und a ein Häufungspunkt von D; aber $a = 2$ ist kein Häufungspunkt von $D = [0, 1]$, nicht einmal von $D = [0, 1] \cup \{2\}$.) Wir werden dies hier nicht besonders thematisieren; in allen konkreten Beispielen wird a stets ein Häufungspunkt von D sein.

Beispiel 32.2. (a) Sei $D := \mathbb{R} \setminus \{0\}$ und $f\colon D \to \mathbb{R}$ definiert durch $f(x) := \frac{\sin(x)}{x}$ für alle $x \in D$. Für $a = 0$ sind Zähler und Nenner gleich 0, also $f(0)$ nicht definiert — aber wir können natürlich trotzdem versuchen uns anzuschauen, was "in der Nähe" von $a = 0$ passiert; beachte, dass $a = 0$ ein Häufungspunkt von D ist. (Es gibt zum Beispiel die Null-Folge mit $x_0 := 1$ und $x_n := 1/n$ für alle $n \in \mathbb{N}$.) Sei $(x_n)_{n\in\mathbb{N}_0}$ eine beliebige Null-Folge mit $x_n \in D$, also $x_n \neq 0$ für alle $n \in \mathbb{N}_0$. Wie im Beweis von Satz 31.5 gilt $|\sin(x) - x| \leqslant |x|^3/3$ für $x \in \mathbb{R}$ mit $|x| \leqslant 2$. Sei $n_0 \in \mathbb{N}_0$ mit $|x_n| \leqslant 2$ für alle $n \geqslant n_0$. Für $n \geqslant n_0$ gilt dann

$$0 \leqslant \big|f(x_n) - 1\big| = \left|\frac{\sin(x_n) - x_n}{x_n}\right| = \frac{1}{|x_n|}\big|\sin(x_n) - x_n\big| \leqslant \frac{1}{|x_n|}\frac{|x_n|^3}{3} = \frac{|x_n|^2}{3}.$$

Die rechte Seite ist eine Null-Folge, also folgt $\lim_{n\to\infty} f(x_n) = 1$ und damit $\lim_{x\to 0} f(x) = 1$. Noch eine Bemerkung: Nach Satz 31.3 (sowie Beispiel 31.2(a), Satz 31.5) ist f stetig auf D. Definieren wir nun $f(0) := 1$, so erhalten wir eine stetige Funktion $f\colon \mathbb{R} \to \mathbb{R}$. In dieser Situation sagt man auch, dass wir eine ***stetige Fortsetzung*** von f auf ganz $\mathbb{R}$ gefunden haben.

(b) Sei $D := \mathbb{R} \setminus \{0\}$ und diesmal $f(x) := \frac{\exp(x)-1}{x}$ für alle $x \in D$. Völlig analog zeigt man auch hier, dass $\lim_{x\to 0} f(x) = 1$ gilt, mit Hilfe der Abschätzung in Beispiel 29.6 (mit $m = 1$).

Nun zum Begriff der Differenzierbarkeit. Gegeben sei eine Funktion $f\colon D \to \mathbb{R}$. Wir möchten f in der Nähe eines gegebenen $a \in D$ durch eine einfachere Funktion annähern; im einfachsten

Fall durch eine Gerade. Dazu wählen wir einen weiteren Punkt $b \in D$ und betrachten die Gerade durch die Punkte $(a, f(a))$ und $(b, f(b))$; diese Gerade, auch ***Sekante*** genannt, ist gegeben durch die Gleichung

$$g(x) := f(a) + \Delta(a,b)(x-a) \quad \text{für alle } x \in \mathbb{R}, \quad \text{wobei} \quad \Delta(a,b) := \frac{f(b)-f(a)}{b-a}.$$

(Hier ist $\Delta(a,b)$ die Steigung der Geraden.) Die Annäherung wird umso besser werden, je näher $b \in D$ am gegebenen $a \in D$ liegt. Dies führt auf die Idee, den Grenzwert von $\Delta(a,b)$ für $b \to a$ zu betrachten. Existiert $\Delta_a := \lim_{b\to a} \Delta(a,b)$, so erhalten wir mit dieser "optimalen" Steigung Δ_a die ***Tangente*** an f durch den Punkt $(a, f(a))$:

$$t(x) := f(a) + \Delta_a(x-a) \qquad \text{für alle } x \in \mathbb{R}.$$

Diese letztlich genial einfache und gleichzeitig überaus weitreichende Idee geht auf Newton und Leibniz am Ende des 17. Jahrhunderts zurück. Etwas formaler:

Definition 32.3. Sei $f\colon D \to \mathbb{R}$ und $a \in D$. Definiere $D_a := D \setminus \{a\}$ und, analog zu oben:

$$F_a\colon D_a \to \mathbb{R} \qquad \text{mit} \qquad F_a(x) := \frac{f(x)-f(a)}{x-a} \qquad \text{für alle } x \in D_a.$$

(Beachte, dass a <u>nicht</u> im Definitionsbereich D_a von F_a ist.) Die Funktion f heißt ***differenzierbar*** in a, wenn der Grenzwert $\Delta_a := \lim_{x\to a} F_a(x)$ existiert[1]; dann heißt $f'(a) := \Delta_a$ die ***Ableitung*** von f in a. Ist f differenzierbar für alle $a \in D$, so sagen wir kurz, dass f differenzierbar auf D ist; wir erhalten dann eine Funktion $f'\colon D \to \mathbb{R}$.

Beispiel 32.4. (a) Sei $f\colon \mathbb{R} \to \mathbb{R}$ selbst eine Gerade, also $f(x) = cx + d$ für alle $x \in \mathbb{R}$, wobei $c, d \in \mathbb{R}$ fest sind. Für $a \in \mathbb{R}$ und $a \neq x \in \mathbb{R}$ ist dann

$$F_a(x) = \frac{f(x)-f(a)}{x-a} = \frac{(cx+d)-(ca+d)}{x-a} = \frac{cx-ca}{x-a} = c.$$

Also ist f differenzierbar in a, mit Ableitung $f'(a) = c$. (Hier ist f gleich der Tangente an f.) Insbesondere ist $f'(a) = 0$ falls $c = 0$, d.h., wenn f konstant ist.

(b) Sei $f(x) = x^2$ für alle $x \in \mathbb{R}$. Sei $a \in \mathbb{R}$ fest. Für $a \neq x \in \mathbb{R}$ ist dann

$$F_a(x) = \frac{f(x)-f(a)}{x-a} = \frac{x^2-a^2}{x-a} = \frac{(x-a)(x+a)}{x-a} = x + a \to 2a \qquad \text{für } x \to a.$$

Also ist f differenzierbar in a, mit Ableitung $f'(a) = 2a$.
Sei $g(x) := \sqrt{x}$ für $x \in \mathbb{R}_{>0}$. Sei $a > 0$ fest; für $a \neq x > 0$ ist dann

$$G_a(x) = \frac{\sqrt{x}-\sqrt{a}}{x-a} = \frac{\sqrt{x}-\sqrt{a}}{(\sqrt{x}-\sqrt{a})(\sqrt{x}+\sqrt{a})} = \frac{1}{\sqrt{x}+\sqrt{a}} \to \frac{1}{2\sqrt{a}} \qquad \text{für } x \to a.$$

Also ist g differenzierbar in a, mit Ableitung $g'(a) = \frac{1}{2\sqrt{a}}$. Beachte: $g(x)$ ist auch für $a = 0$ definiert, aber nicht differenzierbar in $a = 0$.

[1] Wir nehmen hier stillschweigend an, dass a ein Häufungspunkt von D ist, und damit auch von D_a; also können wir den Grenzwert von $F_a(x)$ für $x \to a$ sinnvoll bilden; siehe dazu noch einmal Definition 32.1.

(c) Sei $f\colon \mathbb{R} \to \mathbb{R}$ definiert durch $f(x) := |x|$ (Absolutbetrag) für $x \in \mathbb{R}$. Für $a > 0$ ist f differenzierbar mit $f'(a) = 1$ (siehe (a)); für $a < 0$ ist f ebenfalls differenzierbar mit $f'(a) = -1$ (wiederum nach (a)). Sei nun $a = 0 = f(a)$. Betrachten wir die Null-Folge $(1/n)_{n\in\mathbb{N}}$, so ist $F_0(1/n) = \frac{1/n}{1/n} = 1$ für alle $n \in \mathbb{N}_0$; also konvergiert $F_0(1/n)$ mit Grenzwert 1. Für die Null-Folge $(-1/n)_{n\in\mathbb{N}}$ ist $F_0(1/n) = \frac{1/n}{-1/n} = -1$, also konvergiert $F_0(-1/n)$ mit Grenzwert -1. Folglich ist f nicht differenzierbar in $a = 0$.

Satz 32.5. *Sei* $f\colon D \to \mathbb{R}$ *differenzierbar in* $a \in D$. *Dann ist* f *auch stetig in* a

Beweis. Sei $(x_n)_{n\in\mathbb{N}_0}$ eine beliebige konvergente Folge mit Grenzwert a und $x_n \in D$ für alle $n \in \mathbb{N}_0$. Sei $S := \{n \in \mathbb{N}_0 \mid x_n \neq a\}$. Ist $|S| < \infty$, so gibt es ein $n_0 \in \mathbb{N}_0$ mit $x_n = a$ für alle $n \geqslant n_0$. In diesem Fall ist auch $f(x_n) = f(a)$ für alle $n \geqslant n_0$, also konvergiert $(f(x_n))_{n\in\mathbb{N}_0}$ mit Grenzwert $f(a)$. Sei nun $|S| = \infty$. Dann ist S abzählbar und $S = \{n_k \mid k \in \mathbb{N}_0\}$ mit $0 \leqslant n_0 < n_1 < n_2 < \dots$. Die Folge $(x_{n_k})_{k\in\mathbb{N}_0}$ konvergiert dann auch mit Grenzwert a. Nun gilt $f(x_{n_k}) = f(a) + F_a(x_{n_k})(x_{n_k} - a)$ für alle $k \in \mathbb{N}_0$. Mit den Regeln in Satz 27.8 folgt, dass $(f(x_{n_k}))_{k\in\mathbb{N}_0}$ konvergiert, mit Grenzwert $f(a) + f'(a)(a-a) = f(a)$. Da $f(x_n) = f(a)$ für alle $n \in \mathbb{N}_0 \setminus S$ gilt, konvergiert dann auch die ganze Folge $(f(x_n))_{n\in\mathbb{N}_0}$ gegen $f(a)$. □

Die Umkehrung von Satz 32.5 gilt nicht. Denn zum Beispiel ist die Funktion $f(x) = |x|$ nach Beispiel 32.4(c) nicht differenzierbar in $a = 0$, aber f stetig auf ganz $\mathbb{R}$, wie man sofort sieht.

Satz 32.6. *Die Funktionen* $\exp$, $\sin$ *und* $\cos$ *sind differenzierbar auf ganz* $\mathbb{R}$, *mit* $\exp'(x) = \exp(x)$, $\sin'(x) = \cos(x)$ *und* $\cos'(x) = -\sin(x)$ *für alle* $x \in \mathbb{R}$.

Beweis. Sei $a \in \mathbb{R}$ beliebig und $(x_n)_{n\in\mathbb{N}_0}$ eine konvergente Folge mit Grenzwert a und $x_n \neq a$ für alle $n \in \mathbb{N}_0$. Setze $h_n := x_n - a \neq 0$ für alle $n \in \mathbb{N}_0$; dann ist $(h_n)_{n\in\mathbb{N}_0}$ eine Null-Folge mit $h_n \neq 0$ für alle $n \in \mathbb{N}_0$. Mit der Funktionalgleichung für exp folgt

$$\frac{\exp(x_n) - \exp(a)}{x_n - a} = \frac{\exp(a+h_n) - \exp(a)}{h_n} = \frac{\exp(a)\exp(h_n) - \exp(a)}{h_n} = \exp(a)\frac{\exp(h_n) - 1}{h_n}.$$

Nach Beispiel 32.2(b) konvergiert die Folge $\big((\exp(h_n) - 1)/h_n\big)_{n\in\mathbb{N}_0}$ gegen 1. Damit konvergiert die linke Seite gegen $\exp(a)$, also ist $\exp'(a) = \exp(a)$. Das Argument für die Ableitungen von sin und cos ist völlig analog, mit Hilfe der Additionstheoreme anstelle der Funktionalgleichung (siehe auch Beweis von Satz 31.5). □

Satz 32.7 (Ableitungsregeln, 1. Teil)**.** *Seien* $f, g \in D \to \mathbb{R}$ *differenzierbar in* $a \in D$. *Dann sind auch* $f + g$, $f \cdot g$ *und* f/g *differenzierbar in* a, *mit*

$$(f+g)'(a) = f'(a) + g'(a), \qquad \textit{(Summenregel)},$$

$$(f \cdot g)'(a) = f'(a)g(a) + f(a)g'(a) \qquad \textit{(Produktregel)},$$

$$(f/g)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2} \qquad \textit{falls } g(a) \neq 0 \qquad \textit{(Quotientenregel)}.$$

Beweis. Siehe zum Beispiel die Abschnitte §5.2.1, §5.2.4, §5.2.7 im Buch von Glosauer. □

Beispiel 32.8. (a) Sei $n \in \mathbb{N}$ fest und $p_n(x) := x^n$ für alle $x \in \mathbb{R}$. Wir zeigen mit vollständiger Induktion nach $n \in \mathbb{N}$, dass p_n differenzierbar auf $\mathbb{R}$ ist mit $p_n'(a) = na^{n-1}$ für alle $a \in \mathbb{R}$. Für $n = 1$ gilt dies nach Beispiel 32.4(a). Sei nun $n > 1$ und die Behauptung bereits gezeigt für p_{n-1}. Nun ist $p_n = p_1 \cdot p_{n-1}$, also mit der Produktregel und der Induktionsvoraussetzung

$$p_n'(a) = p_1'(a) \cdot p_{n-1}(a) + p_1(a) \cdot p_{n-1}'(a) = 1 \cdot a^{n-1} + a \cdot ((n-1)a^{n-2}) = na^{n-1}$$

für alle $a \in \mathbb{R}$. Mit den weiteren Regeln im obigen Satz folgt, dass jede Polynomfunktion auf ganz $\mathbb{R}$ differenzierbar ist. Ebenso folgt, dass jede rationale Funktion in ihrem Definitionsbereich differenzierbar ist.

(b) Sei $f(x) := \frac{\sqrt{x}}{1+x^2}$ für $x \in \mathbb{R}_{\geqslant 0}$. Sei $g(x) := \sqrt{x}$, $h(x) := 1 + x^2$. Mit (a) und Beispiel 32.4(b) gilt $g'(a) = 1/(2\sqrt{a})$ und $h'(a) = 2a$ für $a > 0$. Mit der Quotientenregel folgt

$$f'(a) = \frac{g'(a)h(a) - g(a)h'(a)}{g(a)^2} = \frac{\frac{1}{2\sqrt{a}}(1+a^2) - 2a\sqrt{a}}{(1+a^2)^2} = \frac{1-3a^2}{2\sqrt{a}(1+a^2)^2}.$$

(c) Sei $g\colon D \to \mathbb{R}$ differenzierbar; es gelte $g(x) \neq 0$ für alle $x \in D$. Dann erhalten wir die Funktion $f(x) := 1/g(x)$ für $x \in D$. Als Spezialfall der Quotientenregel ergibt sich $f'(a) = -\frac{g'(a)}{g(a)^2}$ falls $g(a) \neq 0$. Ist zum Beispiel $g(x) := x^n$ mit $n \in \mathbb{N}$, so erhalten wir

$$f'(a) = -\frac{na^{n-1}}{a^{2n}} = -\frac{n}{a^{n+1}} \qquad \text{für alle } a \neq 0.$$

Es folgen erste Anwendungen der Differentialrechnung. Diese zeigen vor allem, wie man aus Eigenschaften der Ableitung einer Funktion auf die Funktion selbst zurückschließen kann.

Definition 32.9. Sei $f\colon D \to \mathbb{R}$ und $a \in D$. Wir sagen, dass f in a ein ***lokales Maximum*** hat, wenn es ein $\varepsilon > 0$ gibt mit $(a-\varepsilon, a+\varepsilon) \subseteq D$ und $f(x) \leqslant f(a)$ für alle $x \in (a-\varepsilon, a+\varepsilon)$. Analog sagen wir, dass f in a ein ***lokales Minimum*** hat, wenn es ein $\varepsilon > 0$ gibt mit $(a-\varepsilon, a+\varepsilon) \subseteq D$ und $f(x) \geqslant f(a)$ für alle $x \in (a-\varepsilon, a+\varepsilon)$.
Hat f in a ein lokales Maximum oder ein lokales Minimum, so sagen wir, dass f in a ein ***lokales Extremum*** besitzt.

Ab hier Woche 8

Im Allgemeinen kann es schwierig sein, die lokalen Extrema einer Funktion zu finden. Der folgende Satz führt dieses Problem auf die Bestimmung von Nullstellen der Ableitung zurück.

Satz 32.10. *Die Funktion* $f\colon D \to \mathbb{R}$ *besitze in* $a \in D$ *ein lokales Extremum und sei in* a *differenzierbar. Dann gilt* $f'(a) = 0$.

Beweis. Nehmen wir, es liegt ein lokales Maximum in a vor. (Der Beweis für ein lokales Minimum ist analog.) Sei $\varepsilon > 0$ mit $(a-\varepsilon, a+\varepsilon) \subseteq D$ und $f(x) \leqslant f(a)$ für alle $x \in (a-\varepsilon, a+\varepsilon)$. Setze $x_n := a + \varepsilon/(n+2)$ für $n \in \mathbb{N}_0$. Dann ist $(x_n)_{n\in\mathbb{N}_0}$ konvergent mit Grenzwert a; außerdem ist $x_n \in (a-\varepsilon, a+\varepsilon)$ für alle $n \in \mathbb{N}_0$. Da f in a differenzierbar ist, folgt $\lim_{n\to\infty} F_a(x_n) = f'(a)$. Da f in a ein lokales Maximum hat, gilt $f(x_n) \leqslant f(a)$ für alle $n \in \mathbb{N}_0$. Wegen $x_n > a$ ist damit $F_a(x_n) = (f(x_n) - f(a))/(x_n - a) \leqslant 0$ für alle $n \in \mathbb{N}_0$, also auch $f'(a) = \lim_{n\to\infty} F_a(x_n) \leqslant 0$. Betachten wir dann die Folge $(y_n)_{n\in\mathbb{N}_0}$ mit $y_n := a - \varepsilon/(n+2)$ für $n \in \mathbb{N}_0$, so folgt analog auch $f'(a) \geqslant 0$, also schließlich $f'(a) = 0$. □

Ist f differenzierbar in $a \in D$ mit $f'(a) = 0$, so heißt a ein ***kritischer Punkt*** von f. Beachte: Die Umkehrung des obigen Satzes gilt nicht. Zum Beispiel ist $a = 0$ ein kritischer Punkt der Funktion $f(x) = x^3$, aber in $a = 0$ liegt kein lokales Extremum vor.

Satz 32.11 (***Mittelwertsatz der Differentialrechnung***)**.** *Sei* $f\colon [a,b] \to \mathbb{R}$ *stetig und* f *differenzierbar auf* (a,b). *Dann gibt es ein* $x \in (a,b)$ *mit* $f(b) = f(a) + f'(x)(b-a)$.

Beweis. Zuerst betrachtet man den Fall $f(a) = f(b)$. Dann wird also ein $x \in (a,b)$ gesucht mit $f'(x) = 0$. Ist $f(x) = f(a)$ für alle $x \in [a,b]$, so ist $f'(x) = 0$ für alle $x \in (a,b)$ also die Behauptung klar. Andernfalls gibt es $x_1, x_2 \in [a,b]$ mit $f(x_1) \leqslant f(x) \leqslant f(x_2)$ für alle $x \in [a,b]$ (siehe Satz 31.10), wobei $f(x_1) < f(a) = f(b)$ oder $f(x_2) > f(a) = f(b)$. Im ersten Fall ist $x_1 \in (a,b)$ und f hat ein lokales Minimum in x_1; also $f'(x_1) = 0$ nach Satz 32.10. Im zweiten Fall folgt analog $x_2 \in (a,b)$ und $f'(x_2) = 0$. Ist $f(a) \neq f(b)$, so wendet man das vorherige Argument an auf die Funktion $g\colon [a,b] \to \mathbb{R}$ mit

$$g(x) := f(x)(b-a) - \big(f(b) - f(a)\big)(x-a) \qquad \text{für } x \in [a,b].$$

Beachte: Die Funktion g ist ebenfalls stetig auf $[a,b]$ und differenzierbar auf (a,b); außerdem gilt $g(b) = f(b)(b-a) - (f(b) - f(a))(b-a) = f(a)(b-a) = g(a)$. □

Folgerung 32.12. *Sei* $f\colon [a,b] \to \mathbb{R}$ *stetig und* f *differenzierbar auf* (a,b).

(a) *Gilt* $f'(x) \geqslant 0$ *für alle* $x \in (a,b)$, *so ist* f *monoton wachsend.*

(b) *Gilt* $f'(x) \leqslant 0$ *für alle* $x \in (a,b)$, *so ist* f *monoton fallend.*

Gilt $f'(x) > 0$ *für alle* $x \in (a,b)$ *(bzw.* $f'(x) < 0$ *für alle* $x \in (a,b)$*), so ist* f *sogar streng monoton wachsend (bzw. fallend).*

Beweis. (a) Seien $x_1, x_2 \in [a,b]$ mit $x_1 < x_2$. Dann ist f stetig auf $[x_1, x_2]$ und differenzierbar auf (x_1, x_2). Nach dem Mittelwertsatz gibt es ein $x \in (x_1, x_2)$ mit $f(x_2) = f(x_1) + f'(x)(x_2 - x_1)$. Wegen $f'(x) \geqslant 0$ und $x_1 < x_2$ folgt $f(x_2) \geqslant f(x_1)$. Ist $f'(x) > 0$, so folgt $f(x_2) > f(x_1)$. Der Beweis von (b) ist völlig analog. $\square$

Folgerung 32.13. *Seien* $f, g \colon [a,b] \to \mathbb{R}$ *stetig und differenzierbar auf* (a,b). *Gilt* $f'(x) = g'(x)$ *für alle* $x \in (a,b)$, *so gibt es eine Konstante* $C \in \mathbb{R}$ *mit* $g = f + C$. *Inbesondere: Ist* $f'(x) = 0$ *für alle* $x \in (a,b)$, *so ist* f *konstant.*

Beweis. Sei $h := g - f \colon [a,b] \to \mathbb{R}$. Dann ist h stetig und h differenzierbar mit $h'(x) = 0$ für alle $x \in (a,b)$. Nach Folgerung 32.12 ist h sowohl monoton wachsend als auch monoton fallend, also konstant. $\square$

Beispiel 32.14. Für $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ ist $\cos(x) > 0$; wir definieren die ***Tangens-Funktion*** durch $\tan(x) := \frac{\sin(x)}{\cos(x)}$. Mit der Quotientenregel folgt, dass tan differenzierbar ist mit

$$\tan'(a) = \frac{\sin'(a)\cos(a) - \sin(a)\cos'(a)}{\cos(a)^2} = \frac{\sin(a)^2 + \cos(a)^2}{\cos(a)^2} = \frac{1}{\cos(a)^2} \quad \text{für } a \in \left(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right).$$

Wegen $1 = \cos(a)^2 + \sin(a)^2$ können wir dies auch als $\tan'(a) = 1 + \tan(a)^2$ schreiben; insbesondere ist $\tan'(a) > 0$. Nach Folgerung 32.12 ist tan streng monoton wachsend auf $(-\frac{\pi}{2}, \frac{\pi}{2})$. Für $x \in (0, \frac{\pi}{2})$ ist $\sin(x) > 0$; für $x \in (-\frac{\pi}{2}, 0)$ ist $\sin(x) < 0$; außerdem ist $\sin(\pm\frac{\pi}{2}) = \pm 1$ und $\cos(\pm\frac{\pi}{2}) = 0$. Daraus folgt $\lim\limits_{x \to \pm\pi/2} \tan(x) = \pm\infty$. Also ist $\tan \colon (-\frac{\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ bijektiv und wir erhalten die stetige Umkehrfunktion $\arctan \colon \mathbb{R} \to (-\frac{\pi}{2}, \frac{\pi}{2})$ mit $\lim\limits_{x \to \pm\infty} \arctan(x) = \pm\frac{\pi}{2}$.

Beispiel 32.15. Sei $f(x) := x^5 - 5x^2 + 1$ für $x \in \mathbb{R}$. Welches sind die Nullstellen und die lokalen Extrema von f ? Dazu bestimmen wir die kritischen Punkte von f sowie die Bereiche von $\mathbb{R}$, auf denen f monoton wachsend bzw. fallend ist (Stichwort "***Kurvendiskussion***").
Es gilt $f'(x) = 5x^4 - 10x = 5x(x^3 - 2)$. Für $g(x) := x^3 - 2$ ist $g'(x) = 3x^2 > 0$ für $x \neq 0$, damit g streng monoton wachsend auf $\mathbb{R}_{<0}$ und auf $\mathbb{R}_{>0}$, also auf ganz $\mathbb{R}$. Daher hat g nur eine reelle Nullstelle, nämlich $\sqrt[3]{2}$. Folglich hat f genau 2 kritische Stellen, 0 und $\sqrt[3]{2}$.

Für $x < 0$ ist auch $x^3 - 2 < 0$, also $f'(x) > 0$ und damit f streng monoton wachsend. Wegen $\lim\limits_{x \to -\infty} f(x) = -\infty$ (siehe Beispiel 31.15(b)) und $f(0) = 1 > 0$ gibt es im Bereich $(-\infty, 0)$ genau eine reelle Nullstelle. Wegen $f(-1) = -5 < 0$ ist diese im Intervall $(-1, 0)$ zu finden. Analog ist f streng monoton wachsend für $x > \sqrt[3]{2}$. Wegen $f(\sqrt[3]{2}) = -3\sqrt[3]{2}^2 + 1 < 0$ und $\lim\limits_{x \to +\infty} f(x) = +\infty$ gibt es im Bereich $(\sqrt[3]{2}, +\infty)$ genau eine reelle Nullstelle. Wegen $f(3/2) = -85/32$ und $f(2) = 13 > 0$ ist diese im Intervall $(3/2, 2)$ zu finden. Schließlich ist f streng monoton fallend für $0 < x < \sqrt[3]{2}$. Wegen $f(0) > 0$ und $f(\sqrt[3]{2}) < 0$ gibt es im Bereich $(0, \sqrt[3]{2})$ genau eine weitere reelle Nullstelle. Wegen $f(1) = -3$ ist diese im Intervall $(0, 1)$.
Jetzt sehen wir auch, dass an der Stelle 0 ein lokales Maximum von f vorliegt, und an der

Stelle $\sqrt[3]{2}$ ein lokales Minimum. Beachte: Wegen $\lim\limits_{x\to\pm\infty} f(x) = \pm\infty$ sind dies keine globalen Extrema. Zum Beispiel mit Sage erhält man ein anschauliches Bild des Graphen von f mit

```
sage: plot(x^5-5*x^2+1,-1.5,2)}
```

33. *Berechnung von Ableitungen und Stammfunktionen*

In diesem Abschnitt behandeln wir einige weitere, und sehr nützliche Regeln zum Berechnen von Ableitungen; insbesondere werden wir am Ende sämtliche Formeln in Tabelle 4 nachvollziehen können. Außerdem betrachten wir das Umkehrproblem, d.h., die Frage, wie man zu einer gegebenen Funktion f eine differenzierbare Funktion F finden kann mit $F' = f$.

Satz 33.1 (***Kettenregel***)**.** *Seien* $f\colon D \to \mathbb{R}$ *und* $g\colon E \to \mathbb{R}$ *mit* $f(D) \subseteq E$*; wir können also die Hintereinanderausführung* $g\circ f\colon D \to \mathbb{R}$ *bilden. Sei* $a \in D$*. Ist* f *differenzierbar in* a *und* g *differenzierbar in* $f(a) \in E$*, so ist* $g \circ f$ *differenzierbar in* a *mit* $(g \circ f)'(a) = f'(a)g'\big(f(a)\big)$*.*

Beweis. Siehe zum Beispiel §5.2.5 im Buch von Glosauer. □

Beispiel 33.2. (a) Sei $c \in \mathbb{R}$ fest und $f(x) := \exp(cx)$ für $x \in \mathbb{R}$. Dann ist $f = \exp\circ g$ mit $g(x) := cx$ für alle $x \in \mathbb{R}$. Mit der Kettenregel und Beispiel 32.4(a) folgt also $f'(x) = g'(x)\exp\big(g(x)\big) = c\exp(cx) = cf(x)$ für alle $x \in \mathbb{R}$.

(b) Sei $f\colon [-1, 1] \to \mathbb{R}$ definiert durch $f(x) := \sqrt{1 - x^2}$ für $x \in [-1, 1]$. Dann ist $f(x) = \sqrt{g(x)}$ mit $g(x) := 1 - x^2$ und $g'(x) = -2x$. Mit der Kettenregel und Beispiel 32.4(b) folgt $f'(x) = g'(x)/(2\sqrt{g(x)}) = -x/\sqrt{1 - x^2}$ für $x \in (-1, 1)$.

(c) Sei $f(x) := \log(|x|)$ für $x \in \mathbb{R} \setminus \{0\}$. Mit der Kettenregel und Beispiel 32.4(c) folgt hier sofort $f'(x) = 1/x$ für alle $x \in \mathbb{R} \setminus \{0\}$.

Satz 33.3 (***Differenzierbarkeit der Umkehrfunktion***)**.** *Sei* $f\colon D \to \mathbb{R}$ *injektiv und* $E := f(D) \subseteq \mathbb{R}$*; dann können wir die Umkehrfunktion* $f^{-1}\colon E \to \mathbb{R}$ *bilden. Ist* f *differenzierbar in* $a \in D$ *mit* $f'(a) \neq 0$*, und* f^{-1} *stetig in* $b := f(a) \in E$*, so ist* f^{-1} *differenzierbar in* b *mit*

$$\big(f^{-1}\big)'(b) = \frac{1}{f'(a)} = \frac{1}{f'\big(f^{-1}(b)\big)}.$$

Beweis. Siehe zum Beispiel §15, Satz 3 im 1. Band des Buches von Forster. □

Beispiel 33.4. Nach Definition 31.12 ist $\log = \exp^{-1}\colon \mathbb{R}_{>0} \to \mathbb{R}$. Wir haben dort auch bemerkt, dass log stetig auf $\mathbb{R}_{>0}$ ist. Also folgt nun

$$\log'(b) = \frac{1}{\exp'(\log(b))} = \frac{1}{\exp(\log(b))} = \frac{1}{b} \qquad \text{für alle } b > 0.$$

Analog folgt für alle $b \in (-1, 1)$:

$$\arcsin'(b) = \frac{1}{\sin'(\arcsin(b))} = \frac{1}{\cos(\arcsin(b))} = \frac{1}{\sqrt{1-\sin(\arcsin(b))^2}} = \frac{1}{\sqrt{1-b^2}},$$

sowie auch die Formel für arccos′ in Tabelle 4. Die Formel für arctan′ ergibt sich sofort aus der Formel für tan′ in Beispiel 32.14. Für die restlichen Formeln in Tabelle 4 siehe Übungen. Die ***Hyperbel-Funktionen*** sinh, cosh und tanh sind definiert durch

$$\sinh(x) := \tfrac{1}{2}\big(\exp(x) - \exp(-x)\big), \qquad \cosh(x) := \tfrac{1}{2}\big(\exp(x) + \exp(-x)\big),$$
$$\tanh(x) := \frac{\sinh(x)}{\cosh(x)} = \frac{\exp(2x)-1}{\exp(2x)+1} \qquad \text{für alle } x \in \mathbb{R}.$$

Die Funktion sinh ist eine Bijektion von $\mathbb{R}$ nach $\mathbb{R}$, die Funktion cosh eine Bijektion von $\mathbb{R}_{\geqslant 0}$ nach $\mathbb{R}_{\geqslant 1}$ und tanh eine Bijektion von $\mathbb{R}$ nach $(-1, 1)$; die Umkehrfunktionen werden mit $\operatorname{Arsinh}\colon \mathbb{R} \to \mathbb{R}$, $\operatorname{Arcosh}\colon \mathbb{R}_{\geqslant 1} \to \mathbb{R}$ und $\operatorname{Artanh}\colon (-1,1) \to \mathbb{R}$ bezeichnet. (Siehe Übungen.)

Tabelle 4. Ableitungsformeln

$$\begin{aligned}
f(x) &= c && (x \in \mathbb{R}, c \in \mathbb{R} \text{ fest}) && \Rightarrow && f'(x) = 0,\\
f(x) &= x^n && (x \in \mathbb{R}, n \in \mathbb{N} \text{ fest}) && \Rightarrow && f'(x) = nx^{n-1},\\
f(x) &= \frac{1}{x^n} && (x \neq 0, n \in \mathbb{N} \text{ fest}) && \Rightarrow && f'(x) = \frac{-n}{x^{n+1}},\\
f(x) &= \sqrt[n]{x} && (x > 0, n \geqslant 2 \text{ fest}) && \Rightarrow && f'(x) = \frac{1}{nx}\sqrt[n]{x}\\
f(x) &= x^\alpha && (x > 0, \alpha \in \mathbb{R} \text{ fest}) && \Rightarrow && f'(x) = \alpha x^{\alpha-1}
\end{aligned}$$

$$\exp'(x) = \exp(x), \qquad \sin'(x) = \cos(x), \qquad \cos'(x) = -\sin(x) \qquad (x \in \mathbb{R}),$$

$$\exp_a'(x) = \log(a)\exp_a(x) \quad (x \in \mathbb{R}, a > 0 \text{ fest}), \qquad \log'(x) = \frac{1}{x} \quad (x > 0),$$

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}} \quad (|x| < 1), \qquad \arccos'(x) = \frac{-1}{\sqrt{1-x^2}} \quad (|x| < 1),$$

$$\tan'(x) = \frac{1}{\cos(x)^2} \quad (|x| < \tfrac{\pi}{2}), \qquad \arctan'(x) = \frac{1}{x^2+1} \quad (x \in \mathbb{R}),$$

$$\sinh'(x) = \cosh(x) \quad (x \in \mathbb{R}), \qquad \operatorname{Arsinh}'(x) = \frac{1}{\sqrt{x^2+1}} \quad (x \in \mathbb{R}),$$

$$\cosh'(x) = \sinh(x) \quad (x \in \mathbb{R}), \qquad \operatorname{Arcosh}'(x) = \frac{1}{\sqrt{x^2-1}} \quad (x > 1),$$

$$\tanh'(x) = \frac{1}{\cosh(x)^2} \quad (x \in \mathbb{R}), \qquad \operatorname{Artanh}'(x) = \frac{1}{1-x^2} \quad (|x| < 1).$$

Definition 33.5 (***Allgemeine Potenzfunktion***)**.** Sei $a \in \mathbb{R}$, $a > 0$. Dann definieren wir $\exp_a\colon \mathbb{R} \to \mathbb{R}_{>0}$ durch $\exp_a(x) := \exp\big(\log(a)x\big)$ für alle $x \in \mathbb{R}$. Wir schreiben auch kurz $a^x := \exp_a(x)$ für $x \in \mathbb{R}$. Aus der Funktionalgleichung für exp ergeben sich sofort die folgenden Regeln, für alle $x, y \in \mathbb{R}$ und $m, n \in \mathbb{N}$ mit $m \geqslant 2$:

$$a^{x+y} = a^x \cdot a^y, \quad (a^x)^y = a^{xy}, \quad a^n = a{\cdot}a{\cdot}\cdots a \ \ (n \text{ Faktoren}), \quad a^{n/m} = \sqrt[m]{a^n}.$$

Ist $\alpha \in \mathbb{R}$ fest, so erhalten wir analog eine Funktion $\mathrm{pow}_\alpha \colon \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ mit $\mathrm{pow}_\alpha(x) := \exp\big(\alpha \log(x)\big)$ für $x \in \mathbb{R}_{>0}$. Es gilt dann $x^\alpha = \exp_x(\alpha) = \mathrm{pow}_\alpha(x)$ für $x > 0$.
Für $n \in \mathbb{N}$, $n \geqslant 2$, ist zum Beispiel $\mathrm{pow}_{1/n}(x) = x^{1/n} = \sqrt[n]{x}$ für $x > 0$.

Mit der Kettenregel und der Formel $\exp' = \exp$ folgt, dass $\exp_a$ differenzierbar ist auf $\mathbb{R}$, mit $\exp_a'(x) = \log(a)\exp_a(x)$ für alle $x \in \mathbb{R}$. Ebenso ist pow_α differenzierbar auf $\mathbb{R}_{>0}$, mit

$$\mathrm{pow}_\alpha'(x) = \alpha \log'(x) \exp\big(\alpha \log(x)\big) = \alpha x^{-1} x^\alpha = \alpha x^{\alpha-1} \qquad \text{für } x > 0.$$

Beispiel 33.6. Es gilt $\lim\limits_{n\to\infty} \dfrac{\log(n)}{n} = 0$ und $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$.
Dazu: Nach obigen Regeln ist $n = \sqrt{n}^2 = \exp_{\sqrt{n}}(2) = \exp\big(2\log(\sqrt{n})\big)$, also $\log(n) = 2\log(\sqrt{n})$. Für $n \geqslant 1$ ist auch $\sqrt{n} \geqslant 1$ und damit $\exp(\sqrt{n}-1) \geqslant \sqrt{n}$; also $2(\sqrt{n}-1) \geqslant 2\log(\sqrt{n}) = \log(n)$. Damit erhalten wir $0 \leqslant \frac{\log(n)}{n} \leqslant 2\frac{\sqrt{n}-1}{n} = 2\big(\frac{1}{\sqrt{n}} - \frac{1}{n}\big)$. Die rechte Seite ist eine Nullfolge, also folgt $\lim\limits_{n\to\infty} \frac{\log(n)}{n} = 0$. Schließlich erhalten wir $\sqrt[n]{n} = n^{1/n} = \exp\big(\log(n)/n\big) \to \exp(0) = 1$ für $n \to \infty$ (mit der Stetigkeit von exp).

Mit Hilfe der obigen Regeln und Tabelle 4 ist es in den meisten Fällen relativ leicht möglich, die Ableitungen von Funktionen zu bestimmen, die durch konkrete Formeln gegeben sind. In Sage gibt es dazu die Funktion `diff`; zum Beispiel:

```
sage: diff(sin(x^3)^4*sqrt(1-x^2),x)
12*sqrt(-x^2 + 1)*x^2*cos(x^3)*sin(x^3)^3 - x*sin(x^3)^4/sqrt(-x^2 + 1)
```

Wesentlich komplizierter ist das Umkehrproblem, auf das wir jetzt eingehen.

Definition 33.7. Gegeben sei $f \colon D \to \mathbb{R}$. Eine Funktion $F \colon D \to \mathbb{R}$ heißt ***Stammfunktion*** von f (englisch: “antiderivative”), wenn F differenzierbar ist mit $F'(x) = f(x)$ für alle $x \in D$. Nach Folgerung 32.13 ist F (falls F existiert) “fast” eindeutig bestimmt; genauer: Ist D ein Intervall und $G \colon D \to \mathbb{R}$ eine weitere Stammfunktion von f, so gibt es eine Konstante $C \in \mathbb{R}$ mit $G = F + C$. Das Auffinden einer Stammfunktion wird auch als “***Integration***” bezeichnet; eine Erklärung für diese Bezeichnung erfolgt im nächsten Abschnitt.

Beispiel 33.8. (a) Mit Tabelle 4 kann man für einige Funktionen leicht eine Stammfunktion angeben. Ist zum Beispiel $n \in \mathbb{Z}$, $n \neq -1$, und $f(x) = x^n$ für alle $x \in \mathbb{R}$, so ist eine Stammfunktion gegeben durch $F(x) := \frac{x^{n+1}}{n+1}$ für alle $x \in \mathbb{R}$. Ist $f(x) = \frac{1}{x}$ für alle $x \in \mathbb{R} \setminus \{0\}$, so ist eine Stammfunktion gegeben durch $F(x) := \log(|x|)$ für alle $x \in \mathbb{R} \setminus \{0\}$. Das zweite Beispiel zeigt allerdings bereits, dass Stammfunktionen im Allgemeinen “komplizierter” zu werden scheinen als die gegebene Funktion selbst. Wenn wir log noch nicht definiert hätten (als Umkehrfunktion der nicht ganz so elementaren Funktion exp), so hätten wir hier keine Chance, eine Stammfunktion der so einfach aussehenden Funktion $f(x) = \frac{1}{x}$ zu finden!

(b) Sei nun $f(x) := \sqrt{1-x^2}$ für $x \in [-1,1]$. Aus Tabelle 4 ist nicht direkt ein Kandidat für eine Stammfunktion ersichtlich. Aus der Tabelle könnte man aber zumindest auf die Idee kommen, dass arcsin eine Rolle spielen sollte. In Sage gibt es die Funktion `integral`, die Stammfunktionen bestimmt. In diesem Fall erhalten wir:

```
sage: integral((1-x^2)^(1/2),x)
1/2*sqrt(-x^2 + 1)*x + 1/2*arcsin(x)
```

Setzen wir $F(x) := \frac{1}{2}x\sqrt{1-x^2} + \frac{1}{2}\arcsin(x)$ für $x \in [-1,1]$, so können wir natürlich leicht nachprüfen, dass F differenzierbar ist mit $F'(x) = f(x)$ für alle $x \in [-1,1]$.

Die wirkliche Frage ist also: Wie hat Sage das obige F gefunden? Dazu gibt es einen ziemlich ausgeklügelten Algorithmus; siehe Chapter 12 in
K. O. Geddes, S. R. Czapor and G. Labahn, Algorithms for Computer Algebra, Kluwer Academic Publications, 1992 (oder https://en.wikipedia.org/wiki/Risch_algorithm).

Wir behandeln hier nur einige Spezialfälle und elementare Teilaspekte des Algorithmus.

Beispiel 33.9 (Umkehrung der Produktregel, oder "***partielle Integration***")**.** Sei $f: D \to \mathbb{R}$. Seien $u, v: D \to \mathbb{R}$ differenzierbar, so dass die Werte von f durch die Formel $f(x) = u(x)v'(x)$ für alle $x \in D$ gegeben sind. Ist G eine Stammfunktion von $u' \cdot v$, so ist $F := u \cdot v - G$ eine Stammfunktion von f (wie sofort aus der Produktregel folgt). — Drei Beispiele:

(a) Sei $f(x) := x\sin(x)$ für $x \in \mathbb{R}$. Dann gilt $f = u \cdot v'$ mit $u(x) = x$ und $v(x) = -\cos(x)$. Nun ist $u'(x)v(x) = -\cos(x)$, also ist $G(x) := -\sin(x)$ eine Stammfunktion von $u' \cdot v$. Folglich ist $F(x) = u(x)v(x) - G(x) = -x\cos(x) + \sin(x)$ eine Stammfunktion von f.

(b) Sei $f(x) := \log(x)$ für $x \in \mathbb{R}_{>0}$. Dann gilt $f = u \cdot v'$ mit $u(x) = \log(x)$ und $v(x) = x$. Nun ist $u'(x) = 1/x$ und $u'(x)v(x) = x/x = 1$; also ist $G(x) := x$ eine Stammfunktion von $u' \cdot v$. Folglich ist $F(x) = u(x)v(x) - G(x) = x\log(x) - x$ eine Stammfunktion von $\log(x)$.

(c) Sei $f(x) := \sqrt{1-x^2}$. Dann gilt $f(x) = u \cdot v'$ mit $u(x) := \sqrt{1-x^2}$ und $v(x) = x$. Nun ist $u'(x) = -x/\sqrt{1-x^2}$, also

$$g(x) := u'(x)v(x) = \frac{-x^2}{\sqrt{1-x^2}} = \frac{1-1-x^2}{\sqrt{1-x^2}} = \sqrt{1-x^2} - \frac{1}{\sqrt{1-x^2}}.$$

Nehmen wir nun an, F_1 sei eine Stammfunktion von f. Mit Tabelle 4 sehen wir, dass $G(x) := F_1(x) - \arcsin(x)$ eine Stammfunktion von g ist. Also ist auch $F_2(x) := u(x)v(x) - G(x) = x\sqrt{1-x^2} - F_1(x) + \arcsin(x)$ eine Stammfunktion von f. Nun gibt es eine Konstante $C \in \mathbb{R}$ mit $F_2 = F_1 + C$. Also folgt $F_1(x) + C = x\sqrt{1-x^2} - F_1(x) + \arcsin(x)$ und damit $F_1(x) = \frac{1}{2}x\sqrt{1-x^2} + \frac{1}{2}\arcsin(x) - \frac{C}{2}$. Nachdem wir diesen Kandidaten gefunden haben, können wir auch direkt nachprüfen, dass durch $F(x) := \frac{1}{2}x\sqrt{1-x^2} + \frac{1}{2}\arcsin(x)$ tatsächlich eine Stammfunktion von f gegeben ist, wie bereits vorher von Sage gefunden.

Beispiel 33.10 (Kettenregel "rückwärts")**.** Sei $f\colon D \to \mathbb{R}$ so, dass die Werte durch eine Formel der Form $f(x) = \varphi'(x)g(\varphi(x))$ für $x \in D$ gegeben sind, wobei $\varphi\colon D \to \mathbb{R}$ und $g\colon E \to \mathbb{R}$ differenzierbar sind mit $\varphi(D) \subseteq E$. Ist G eine Stammfunktion von g, so ist $F := G \circ \varphi$ eine Stammfunktion von f (wie sofort aus der Kettenregel folgt). — Zwei Beispiele:

(a) Sei $f(x) = x\sin(x^2)$ für $x \in \mathbb{R}$. Mit $\varphi(x) := x^2$ und $g(x) := \sin(x)/2$ sind wir genau in der obigen Situation. Nun ist $G(x) = -\cos(x)/2$ eine Stammfunktion von g; also ist durch $F(x) := -\cos(x^2)/2$ eine Stammfunktion von f gegeben.

(b) Seien $a, b \in \mathbb{R}$ mit $a \neq 0$ und $f(x) = g(ax+b)$ für $x \in D$, wobei $g\colon E \to \mathbb{R}$ differenzierbar ist mit $ax + b \in E$ für alle $x \in D$. Ist G eine Stammfunktion von g, so ist durch $F(x) := a^{-1}G(ax + b)$ eine Stammfunktion von f gegeben.

Bemerkung 33.11. Die obigen Verfahren (und Verfeinerungen davon) funkionieren in vielen Fällen, aber nicht immer: Es gibt Funktionen, von denen man zeigen kann, dass sie eine Stammfunktion besitzen, aber auch, dass es keine geschlossene Formel für diese Stammfunktionen geben kann. Nach Beispiel 12.15 im oben genannten Buch von Geddes et al. ist $f(x) = \exp(-x^2)$ eine solche Funktion. Die Stammfunktionen solcher Funktionen sind dann sozusagen "neue" Funktionen, die man — je nach Interesse oder Wichtigkeit — zum Sortiment der üblicherweise bekannten Funktionen (wie exp, sin, cos, ...) hinzufügen muss.

Bemerkung 33.12. Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge reeller Zahlen und $f\colon D \to \mathbb{R}$ die zugehörige Potenzreihe, also $f(x) = \sum_{n=0}^{\infty} a_n x^n$ für alle $x \in D$. Sei R_0 der Konvergenzradius von f; siehe Satz 29.1. Dann ist f differenzierbar auf $(-R_0, R_0)$ und es gilt

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \qquad \text{für alle } x \in (-R_0, R_0).$$

Außerdem besitzt f eine Stammfunktion, gegeben durch

$$F(x) := \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} \qquad \text{für alle } x \in (-R_0, R_0).$$

Ableitung und Integration von Potenzreihen dürfen also einfach "gliedweise" ausgeführt werden. (Für einen Beweis siehe zum Beispiel Anhang A im Buch von Haggarty.) Damit erhält man auch sehr leicht die Formeln in Satz 32.6.

Ab hier Woche 9

Wir zeigen jetzt noch, dass sich zu einer ***rationalen Funktion*** stets eine Stammfunktion finden lässt, und zwar mit einem expliziten Algorithmus. Das Ergebnis ist eine Kombination von rationalen Funktionen und zusätzlichen Termen, in denen log und arctan vorkommen (wie bereits durch Tabelle 4 und Beispiel 33.8(a) nahegelegt). Die Grundidee besteht darin, eine rationale Funktion in einfachere Bestandteile zu zerlegen, für die man dann direkt Stammfunktionen angeben kann. Diese Technik geht auf Leibniz und Bernoulli Anfang des 18. Jahrhunderts zurück; siehe `https://de.wikipedia.org/wiki/Partialbruchzerlegung`.

Satz 33.13 (***Partialbruchzerlegung***). *Gegeben seien Polynomfunktionen* $f, g\colon \mathbb{R} \to \mathbb{R}$, *so dass* g *nicht die Null-Funktion und normiert ist; sei* $D := \{x \in \mathbb{R} \mid g(x) \neq 0\}$. *Dann gibt es eine Polynom-Funktion* $\tilde{f}(x)$, *sowie Konstanten* $x_i, p_i, q_i, a_{ij}, b_{ij}, c_{ij} \in \mathbb{R}$ *mit*

$$h(x) := \frac{f(x)}{g(x)} = \tilde{f}(x) + \sum_{i=1}^{r} \sum_{j=1}^{r_i} \frac{a_{ij}}{(x-x_i)^j} + \sum_{i=1}^{s} \sum_{j=1}^{s_i} \frac{b_{ij}x + c_{ij}}{(x^2 + p_i x + q_i)^j} \qquad \text{für alle } x \in D,$$

wobei p_i, q_i *so sind, dass die Polynomfunktionen* $\psi_i(x) := x^2 + p_i x + q_i$ *keine reelle Nullstelle haben (also* $4q_i - p_i^2 > 0$*). Die Zahlen* $r, s, r_i, s_i \in \mathbb{N}_0$ *sind hierbei durch* g *bestimmt; es gilt*

$$g(x) = \prod_{i=1}^{r} (x - x_i)^{r_i} \cdot \prod_{i=1}^{s} (x^2 + p_i x + q_i)^{s_i} \qquad \text{für alle } x \in \mathbb{R}.$$

Anstelle eines formalen Beweises betrachten wir das Beispiel mit $f(x) := x^6$ und $g(x) := x^5 + x^4 + x^3 - x^2 - x - 1$. Zuerst teilen wir f durch g mit Rest, also $f(x) = (x-1)g(x) + 2x^3 - 1$; außerdem finden wir die Faktorisierung $g(x) = (x-1)(x^2+x+1)^2$. Damit ist

$$h(x) = \tfrac{f(x)}{g(x)} = \tfrac{(x-1)g(x)+2x^3-1}{g(x)} = x - 1 + \tfrac{2x^3-1}{(x-1)(x^2+x+1)^2}.$$

Gemäß der Formel im obigen Satz machen wir nun einen Ansatz wie folgt:

$$\begin{aligned}\tfrac{2x^3-1}{(x-1)(x^2+x+1)^2} &= \tfrac{a_{11}}{x-1} + \tfrac{b_{11}x+c_{11}}{x^2+x+1} + \tfrac{b_{12}x+c_{12}}{(x^2+x+1)^2} = \tfrac{a_{11}(x^2+x+1)^2 + (b_{11}x+c_{11})(x-1)(x^2+x+1) + (b_{12}x+c_{12})(x-1)}{(x-1)(x^2+x+1)^2} \\ &= \tfrac{(a_{11}+b_{11})x^4 + (2a_{11}+c_{11})x^3 + (3a_{11}+b_{12})x^2 + (2a_{11}-b_{11}-b_{12}+c_{12})x + a_{11}-c_{11}-c_{12}}{(x-1)(x^2+x+1)^2}.\end{aligned}$$

Durch Koeffizientenvergleich erhält man ein LGS mit den Unbekannten $a_{11}, b_{11}, c_{11}, b_{12}, c_{12}$; dieses hat die eindeutige Lösung $a_{11} = 1/9$, $b_{11} = -1/9$, $c_{11} = 16/9$, $b_{12} = -1/3$, $c_{12} = -2/3$. In Sage wird dies durch die Funktion `partial_fraction_decomposition` geleistet:

```
sage: h=x^6/((x-1)*(x^2+x+1)^2)
sage: h.partial_fraction_decomposition()
[-1/9*(x-16)/(x^2+x+1), -1/3*(x+2)/(x^2+x+1)^2, x, 1/9/(x-1), -1]
```

Die Funktion $x - 1$ hat als Stammfunktion $\frac{1}{2}x^2 - x$; der Bruch $\frac{1}{x-1}$ hat als Stammfunktion $\log(x-1)$. Wir müssen also "nur" noch Stammfunktionen für $\frac{b_{11}x+c_{11}}{x^2+x+1}$ und $\frac{b_{12}x+c_{12}}{(x^2+x+1)^2}$ finden.

Beispiel 33.14. Seien $p, q \in \mathbb{R}$ so, dass die Polynomfunktion $\psi(x) = x^2 + px + q$ keine reelle Nullstelle hat; also $\psi(x) > 0$ für $x \in \mathbb{R}$. Sei $d \in \mathbb{N}$. Gesucht sind Stammfunktionen von

$$f_d(x) := \frac{1}{(x^2+px+q)^d} \qquad \text{und} \qquad g_d(x) := \frac{x}{(x^2+px+q)^d} \qquad (x \in \mathbb{R}).$$

Betrachte zuerst $f_1(x)$: Mit quadratischer Ergänzung erhalten wir $\psi(x) = (x + p/2)^2 + d/4$, wobei $d := 4q - p^2 > 0$. Setzen wir $\varphi(x) := \sqrt{d}^{-1}(2x+p)$ für $x \in \mathbb{R}$, so gilt

$$f_1(x) = \frac{1}{\psi(x)} = \frac{4}{(2x+p)^2+d} = \frac{4}{d} \cdot \frac{1}{1 + \left(\sqrt{d}^{-1}(2x+p)\right)^2} = \frac{4}{d} \cdot \frac{1}{1+\varphi(x)^2}.$$

Wegen $\varphi'(x) = 2\sqrt{d}^{-1}$ and $\arctan'(x) = \frac{1}{1+x^2}$ ist eine Stammfunktion von f_1 gegeben durch

$$\boxed{F_1(x) := \frac{2}{\sqrt{d}} \cdot \arctan\big(\varphi(x)\big) = \frac{2}{\sqrt{4q-p^2}} \cdot \arctan\Big(\frac{2x+p}{\sqrt{4q-p^2}}\Big).}$$

Für $d \geqslant 2$ definieren wir $F_d\colon \mathbb{R} \to \mathbb{R}$ rekursiv durch:

$$\boxed{F_d(x) := \frac{1}{(4q-p^2)(d-1)} \cdot \Big(\frac{2x+p}{(x^2+px+q)^{d-1}} + (4d-6)F_{d-1}(x) \Big).}$$

Mit vollständiger Induktion nach d, und unter Benutzung der Quotienten- und der Kettenregel prüft man leicht nach, dass F_d eine Stammfunktion von f_d ist.
Schließlich zu $g_d(x)$ für $d \geqslant 1$: Mit $\psi(x) = x^2 + px + q$ und $\psi'(x) = 2x + p$ erhalten wir

$$g_d(x) = \frac{x}{\psi(x)^d} = \frac{1}{2} \cdot \frac{2x+p}{\psi(x)^d} - \frac{p}{2} \cdot \frac{1}{\psi(x)^d} = \frac{1}{2} \cdot \frac{\psi'(x)}{\psi(x)^d} - \frac{p}{2} \cdot \frac{1}{\psi(x)^d}.$$

Mit der Kettenregel sowie den Formeln $\log'(x) = 1/x$ und $(1/x^{d-1})' = -(d-1)/x^d$ für $d \geqslant 2$ sind Stammfunktionen von g_1 bzw. g_d für $d \geqslant 2$ gegegen durch

$$\boxed{G_1(x) := \frac{1}{2} \cdot \log\big(\psi(x)\big) - \frac{p}{2} \cdot F_1(x) = \frac{1}{2} \cdot \log\big(x^2 + px + q\big) - \frac{p}{2} \cdot F_1(x),}$$

$$\boxed{G_d(x) := -\frac{1}{2d-2} \cdot \frac{1}{(x^2+px+q)^{d-1}} - \frac{p}{2} F_d(x) \qquad \text{für } d \geqslant 2.}$$

Sind $b, c \in \mathbb{R}$, so ist eine Stammfunktion von $\dfrac{bx+c}{(x^2+px+q)^d}$ gegegen durch $cF_d + bG_d$.

Damit können wir sofort eine Stammfunktion für obiges $h(x)$ hinschreiben. Mit Sage:

```
sage: integral(x^6/((x-1)*(x^2+x+1)^2),x)
1/2*x^2 + sqrt(3)*arctan(1/3*sqrt(3)*(2*x + 1)) - x
     - 1/3*x/(x^2 + x + 1) - 1/18*log(x^2 + x + 1) + 1/9*log(x - 1)
```

Für umfangreiche Tabellen mit Stammfunktionen siehe auch das mittlerweile klassische "Taschenbuch der Mathematik" von Bronstein–Semedjajev; heute verfügbar in der Ausgabe von E. Zeidler (siehe Literatur S. iii).

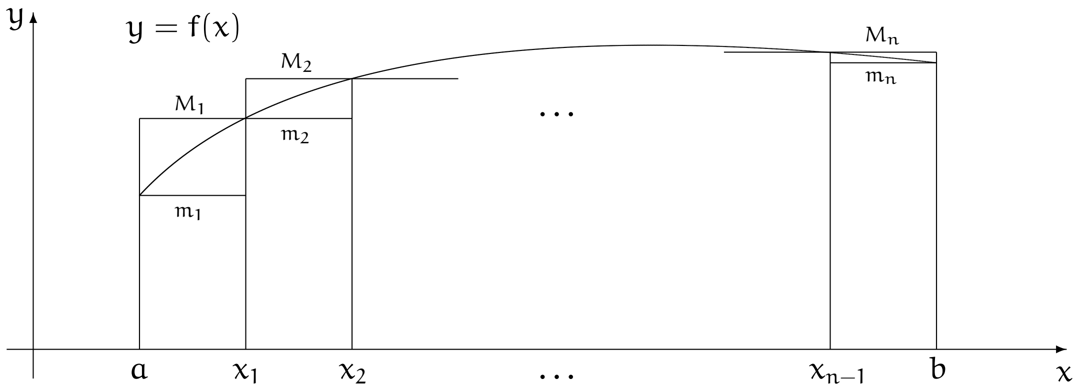
34. *Das Riemann–Integral; die Kreiszahl* π

Gegeben sei eine Funktion $f\colon [a,b] \to \mathbb{R}$, wobei $a, b \in \mathbb{R}$ mit $a < b$. Die Theorie der Integrale wurde erfunden um den Flächeninhalt zwischen der x-Achse und dem Graph der Funktion $y = f(x)$ zu berechnen. Hier ist die grundlegende Idee: Sei $P \subseteq [a,b] \subseteq \mathbb{R}$ eine endliche Teilmenge mit $a, b \in P$, auch als "Zerlegung" von $[a,b]$ bezeichnet. Sei $n := |P| - 1 \in \mathbb{N}$ und bilde die Liste $\underline{P}$ der $n+1$ Elemente von P, aufsteigend geordnet, also

$$\underline{P}: \quad a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b.$$

Schauen wir uns nun die Werte von f auf den Teilintervallen $[x_{k-1}, x_k]$ für $k = 1, \ldots, n$ an. Nehmen wir an, dass f nach oben und nach unten beschränkt ist auf dem Intervall $[a,b]$, d.h., es gibt $m, M \in \mathbb{R}$ mit $m \leqslant f(x) \leqslant M$ für alle $x \in [a,b]$. Dies gilt dann auch für jedes der Teilintervalle $[x_{k-1}, x_k]$. Also können wir definieren:

$$m_k := \inf\{f(x) \mid x \in [x_{k-1}, x_k]\} \quad \text{und} \quad M_k := \sup\{f(x) \mid x \in [x_{k-1}, x_k]\}.$$



Wir definieren die ***Obersumme*** und ***Untersumme*** von f bezüglich P durch

$$U_f(P) = \sum_{k=1}^{n} M_k(x_k - x_{k-1}) \qquad \text{und} \qquad L_f(P) = \sum_{k=1}^{n} m_k(x_k - x_{k-1}).$$

($\mathfrak{U}$ für "upper sum", und $\mathfrak{L}$ für "lower sum".) Es gilt $m \leqslant m_k \leqslant M_k \leqslant M$ für alle k, also folgt

$$m(b-a) \leqslant L_f(P) \leqslant U_f(P) \leqslant M(b-a) \qquad \text{für jede Zerlegung } P \text{ von } [a,b].$$

Insbesondere ist $m(b-a)$ eine untere Schranke für die Menge

$$U(f,[a,b]) = \{U_f(P) \mid P \text{ Zerlegung von } [a,b] \text{ wie oben}\} \subseteq \mathbb{R};$$

analog ist $M(b-a)$ eine obere Schranke für die Menge

$$L(f,[a,b]) = \{L_f(P) \mid P \text{ Zerlegung von } [a,b] \text{ wie oben}\} \subseteq \mathbb{R}.$$

Da $\mathbb{R}$ ein vollständiger Körper ist, existieren also

$$\boxed{\mathfrak{L} := \sup(L(f,[a,b])) \qquad \text{und} \qquad \mathfrak{U} := \inf(U(f,[a,b])).}$$

Lemma 34.1. *Mit den obigen Bezeichnungen gilt* $\mathfrak{L} \leqslant \mathfrak{U}$. *Genauer:*
(a) *Sind* P, P' *Zerlegungen von* $[a,b]$ *mit* $P \subseteq P'$, *so gilt* $L_f(P) \leqslant L_f(P')$ *und* $U_f(P) \geqslant U_f(P')$.
(b) *Sind* P, P' *beliebige Zerlegungen von* $[a,b]$, *so gilt stets* $L_f(P) \leqslant U_f(P')$.

Beweis. (a) Nehmen wir zuerst an, dass P' ein Element mehr als P enthält, also $P' \setminus P = \{y\}$ mit $y \in (a,b)$. Ist $\underline{P}$: $a = x_0 < x_1 < x_2 < \ldots < x_n = b$, so gibt es ein $l \in \{1,\ldots,n\}$ mit $x_{l-1} < y < x_l$. Dann ist $\underline{P}'$: $a = x_0 < x_1 < \ldots < x_{l-1} < y < x_l < \cdots < x_n = b$. Sei

$$m_l' = \inf\{f(x) \mid x \in [x_{l-1}, y]\}, \qquad m_l'' = \inf\{f(x) \mid x \in [y, x_l]\},$$
$$M_l' = \sup\{f(x) \mid x \in [x_{l-1}, y]\}, \qquad M_l'' = \sup\{f(x) \mid x \in [y, x_l]\}.$$

Wegen $m_l \leqslant f(x) \leqslant M_l$ für alle $x \in [x_{l-1}, x_l]$ folgt $m_l \leqslant m_l', m_l''$ and $M_l \geqslant M_l', M_l''$. Also:

$$\begin{aligned} L_f(P') &= \textstyle\sum_{k=1}^{l-1} m_k(x_k - x_{k-1}) + m_l'(y - x_{l-1}) + m_l''(x_l - y) + \sum_{k=l+1}^{n} m_k(x_k - x_{k-1}) \\ &\geqslant \textstyle\sum_{k=1}^{l-1} m_k(x_k - x_{k-1}) + m_l(y - x_{l-1}) + m_l(x_l - y) + \sum_{k=l+1}^{n} m_k(x_k - x_{k-1}) \\ &\geqslant \textstyle\sum_{k=1}^{l-1} m_k(x_k - x_{k-1}) + m_l(x_l - x_{l-1}) + \sum_{k=l+1}^{n} m_k(x_k - x_{k-1}) = L_f(P), \end{aligned}$$

und analog $U_f(P') \leqslant U_f(P)$. Ist $m := |P' \setminus P| \geqslant 2$, so betrachte $P = P_0 \subseteq P_1 \subseteq \ldots \subseteq P_m = P'$, wobei jeweils P_i ein Element mehr als P_{i-1} enthält. In m Schritten folgt dann (a).

(b) Sei $P'' := P \cup P'$. Mit (a) folgt $L_f(P) \leqslant L_f(P'') \leqslant U_f(P'') \leqslant U_f(P')$.

Jetzt zu $\mathfrak{L} \leqslant \mathfrak{U}$: Ist P' fest, so ist $U_f(P')$ nach (b) eine obere Schranke für $L(f, [a, b])$, also $\mathfrak{L} \leqslant U_f(P')$. Aber dann ist $\mathfrak{L}$ eine untere Schranke für $U(f, [a, b]]$, also schließlich $\mathfrak{L} \leqslant \mathfrak{U}$. $\square$

Definition 34.2. Sei $f\colon [a, b] \to \mathbb{R}$ nach oben und unten beschränkt. Die Funktion f heißt ***Riemann-integrierbar*** (oder einfach integrierbar), wenn $\mathfrak{L} = \mathfrak{U}$ gilt. Der gemeinsame Wert $\mathfrak{L} = \mathfrak{U}$ wird mit $\int_a^b f(x)dx$ bezeichnet und heißt das ***Integral*** von f über $[a, b]$.
Als Konvention setzen wir $\int_b^a f(x)dx := -\int_a^b f(x)dx$; ist $a = b$, dann $\int_a^a f(x)dx := 0$.

Beispiel 34.3. (a) Sei $C \in \mathbb{R}$ fest und $f(x) := C$ für $x \in [a, b]$. Dann ist f integrierbar und es gilt $\int_a^b f(x)dx = C(b - a)$ (was auch der Anschauung entspricht).

Denn: Für eine beliebige Zerlegung P wie oben gilt $m_k = M_k = C$. Also ist $L_f(P) = \sum_{k=1}^n m_k(x_k - x_{k-1}) = C\sum_{k=1}^n (x_k - x_{k-1}) = C(x_n - x_0) = C(b - a)$. Analog folgt $U_f(P) = C(b - a)$. Also sind das obige Infimum $\mathfrak{L}$ und Supremum $\mathfrak{U}$ beide gleich $C(b - a)$.

(b) Sei $f(x) := x$ für $x \in \mathbb{R}$. Dann ist f integrierbar auf jedem Intervall $[0, b]$ mit $b \in \mathbb{R}$, $b \geqslant 0$. Dazu: Für $n \in \mathbb{N}$ betrachten wir die Zerlegung $P_n : 0 = x_0 < x_1 < \ldots < x_n = b$ mit $x_k = bk/n$. Dann ist $m_k = b(k-1)/n$ und $M_k = bk/n$. Dies ergibt

$$U_f(P_n) = \textstyle\sum_{k=1}^n \frac{bk}{n}\big(\frac{bk}{n} - \frac{b(k-1)}{n}\big) = \frac{b^2}{n^2}(1 + 2 + \cdots + n) = \frac{b^2}{n^2} \cdot \frac{1}{2}n(n+1) = \frac{1}{2}b^2(1 + \frac{1}{n});$$

analog $L_f(P_n) = \frac{1}{2}b^2(1 - \frac{1}{n})$. Also gilt $\frac{1}{2}b^2\big(1 - \frac{1}{n}\big) = L_f(P_n) \leqslant \mathfrak{L} \leqslant \mathfrak{U} \leqslant U_f(P_n) = \frac{1}{2}b^2\big(1 + \frac{1}{n}\big)$ mit Lemma 34.1. Weil dies für alle $n \in \mathbb{N}$ gilt, folgt $\mathfrak{L} = \mathfrak{U}$ und damit $\int_0^b x dx = \frac{1}{2}b^2$. (Dies entspricht auch der Anschauung, dass $\int_0^b x dx$ die Hälfte der Fläche des Quadrats mit Seitenlänge b sein sollte.)

(c) Wir definieren $f\colon [0, 1] \to \mathbb{R}$ durch $f(x) = \begin{cases} 1 & \text{falls } x \text{ rational,} \\ 0 & \text{falls } x \text{ irrational.} \end{cases}$
Dann ist f nicht R-integrierbar. Denn: Sei P eine beliebige Zerlegung von $[0, 1]$. Jedes Intervall $[x_{k-1}, x_k]$ enthält sowohl eine rationale als auch eine irrationale Zahl, also ist $m_k = 0$ und $M_k = 1$. Damit folgt $L_f(P) = 0$ und $U_f(P) = 1$ für alle P, also $\mathfrak{L} = 0$ and $\mathfrak{U} = 1$.

Beispiel 34.4. Seien $f, g\colon [a, b] \to \mathbb{R}$ nach oben und unten beschränkt. Nehmen wir an, dass f, g integrierbar sind und $f(x) \leqslant g(x)$ für alle $x \in [a, b]$ gilt. Dann ist $\int_a^b f(x)dx \leqslant \int_a^b g(x)dx$.

Denn: Sei P eine Zerlegung von $[a, b]$ und $\underline{P}$: $a = x_0 < x_1 < \ldots < x_n = b$ wie oben. Für $k = 1, \ldots, n$ sei $m_k := \inf\{f(x) \mid x \in [x_{k-1}, x_k]\}$ und $m_k' := \inf\{g(x) \mid x \in [x_{k-1}, x_k]\}$. Wegen $m_k \leqslant f(x) \leqslant g(x)$ für $x \in [x_{k-1}, x_k]$ ist $m_k \leqslant m_k'$. Daraus folgt $L_f(P) \leqslant L_g(P) \leqslant \int_a^b g(x)dx$. Weil dies für alle P gilt, folgt $\int_a^b f(x)dx \leqslant \int_a^b g(x)dx$.

Satz 34.5 (Riemann–Kriterium). *Sei $f\colon [a, b] \to \mathbb{R}$ nach oben und unten beschränkt. Genau dann ist f integrierbar, wenn es zu jedem $\varepsilon > 0$ eine Zerlegung P von $[a, b]$ gibt mit $U_f(P) - L_f(P) < \varepsilon$.*

Beweis. Sei zuerst f integrierbar; sei $I := \mathfrak{L} = \mathfrak{U}$ und $\varepsilon > 0$. Dann ist $\mathfrak{L} - \varepsilon/2$ keine obere Schranke für $L(f, [a, b])$, also gibt es eine Zerlegung P' von $[a, b]$ mit $\mathfrak{L} - \varepsilon/2 < L_f(P') \leqslant I$. Analog gibt es eine Zerlegung P'' von $[a, b]$ mit $I \leqslant U_f(P'') < \mathfrak{U} + \varepsilon/2$. Folglich ist $U_f(P'') - L_f(P') < (\mathfrak{U} + \varepsilon/2) - (\mathfrak{L} - \varepsilon/2) \leqslant \varepsilon$. Sei nun $P := P' \cup P''$. Mit Lemma 34.1 folgt $L_f(P') \leqslant L_f(P)$ und $U_f(P) \leqslant U_f(P'')$, also $U_f(P) - L_f(P) \leqslant U_f(P'') - L_f(P') < \varepsilon$.

Sei umgekehrt die obige Bedingung erfüllt. Nehmen wir an, es sei $\mathfrak{L} < \mathfrak{U}$ und setze $\varepsilon := (\mathfrak{U} - \mathfrak{L})/2 > 0$. Dann gibt es eine Zerlegung P von $[a, b]$ mit $U_f(P) - L_f(P) < \varepsilon$. Aber $L_f(P) \leqslant \mathfrak{L} \leqslant \mathfrak{U} \leqslant U_f(P)$ und damit $U_f(P) - L_f(P) \geqslant \mathfrak{U} - \mathfrak{L} = 2\varepsilon > \varepsilon$, Widerspruch. Also ist $\mathfrak{L} \geqslant \mathfrak{U}$. Nach Lemma 34.1 gilt auch $\mathfrak{L} \leqslant \mathfrak{U}$ und damit $\mathfrak{L} = \mathfrak{U}$. □

Wir präsentieren nun drei Folgerungen aus dem obigen Kriterium, mit dem wir anschließend viele Integrale konkret ausrechnen können, ohne mühsam Zerlegungen von $[a, b]$ sowie die zugehörigen Ober- und Untersummen wie in den obigen Beispielen betrachten zu müssen.

Folgerung 34.6. *Sei* $f\colon [a, b] \to \mathbb{R}$ *monoton wachsend auf* $[a, b]$ *oder monoton fallend auf* $[a, b]$. *Dann ist* f *integrierbar.*

Beweis. Nehmen wir an, f sei monoton wachsend auf $[a, b]$. Sei $m := f(a)$ und $M := f(b)$. Dann ist $m \leqslant f(x) \leqslant M$ für alle $x \in [a, b]$, also f nach oben und unten beschränkt. Für $m = M$ siehe Beispiel 34.3. Sei nun $m < M$. Sei $\varepsilon > 0$ beliebig. Sei $\delta := \varepsilon/(M - m) > 0$ und wähle $n \in \mathbb{N}$ mit $n > (b - a)/\delta$. Setze $x_k := a + k(b - a)/n$ für $k = 0, 1, \ldots, n$. Dann ist $x_0 = a$, $x_n = b$, also P eine Zerlegung von $[a, b]$; es gilt $x_k - x_{k-1} = (b - a)/n < \delta$ für alle k. Da f monoton wachsend ist auch auf jedem Teilintervall $[x_{k-1}, x_k]$, gilt $m_k = f(x_{k-1})$ und $M_k = f(x_k)$. Dies ergibt:

$$\begin{aligned} U_f(P) - L_f(P) &= \textstyle\sum_{k=1}^n (M_k - m_k)(x_k - x_{k-1}) < \delta \sum_{k=1}^n (M_k - m_k) = \delta \sum_{k=1}^n \big(f(x_k) - f(x_{k-1})\big) \\ &= \delta\big(f(x_1) - f(x_0) + f(x_2) - f(x_1) + \cdots - f(x_{n-1}) + f(x_n)\big) = (M - m)\delta = \varepsilon. \end{aligned}$$

Also ist f integrierbar nach Satz 34.5. Für f monoton fallend ist das Argument analog. □

Folgerung 34.7 (***Zerlegungsregel***)**.** *Sei* $f\colon [a, b] \to \mathbb{R}$ *nach oben und unten beschränkt. Sei* $a < c < b$. *Ist* f *integrierbar auf* $[a, c]$ *und auf* $[c, b]$, *so ist* f *integrierbar auf* $[a, b]$ *und es gilt* $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$.

Beweis. Ist P eine Zerlegung von $[a, c]$ und P' eine Zerlegung von $[c, b]$, so ist $P'' := P \cup P'$ eine Zerlegung von $[a, b]$ mit $L_f(P'') = L_f(P) + L_f(P')$ und $U_f(P'') = U_f(P) + U_f(P')$. Daraus ergeben sich leicht die Aussagen; siehe zum Beispiel §7.1.10(2) im Buch von Haggarty. □

Mit den obigen beiden Folgerungen ist eine Function $f\colon [a,b] \to \mathbb{R}$ integrierbar, wenn f ***stückweise monoton*** ist, d.h., wenn wir $[a,b]$ in endlich viele Teilintervalle aufteilen können, auf denen f jeweils monoton wachsend oder fallend ist. Viele der üblichen Funktionen haben diese Eigenschaft; finden Sie ein Beispiel einer nicht stückweise monotonen Funktion!

Satz 34.8 (***Hauptsatz der Differential- und Integralrechnung, 1. Teil***).
Sei $f\colon [a,b] \to \mathbb{R}$ integrierbar. Es gebe eine Funktion $F\colon [a,b] \to \mathbb{R}$, die stetig auf $[a,b]$ ist und differenzierbar für alle $x \in (a,b)$ mit $F'(x) = f(x)$. Dann gilt $\int_a^b f(x)dx = F(b) - F(a)$.

Beweis. Sei $\varepsilon > 0$ beliebig. Nach Voraussetzung und Satz 34.5 gibt es eine Zerlegung P von $[a,b]$ mit $U_f(P) - L_f(P) < \varepsilon$. Sei $\underline{P}\colon a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$. Für jedes $k \in \{1,\ldots,n\}$ gibt es nach dem Mittelwertsatz der Differentialrechnung ein $y_k \in (x_{k-1}, x_k)$ mit $F(x_k) - F(x_{k-1}) = F'(y_k)(x_k - x_{k-1}) = f(y_k)(x_k - x_{k-1})$. Damit folgt

$$F(b) - F(a) = \textstyle\sum_{k=1}^n \big(F(x_k) - F(x_{k-1})\big) = \sum_{k=1}^n f(y_k)(x_k - x_{k-1}).$$

Wie oben seien $m_k := \inf\{f(x) \mid x \in [x_{k-1}, x_k]\}$ und $M_k := \sup\{f(x) \mid x \in [x_{k-1}, x_k]\}$. Dann ist $m_k \leqslant f(y_k) \leqslant M_k$ für alle k und damit $L_f(P) \leqslant F(b) - F(a) \leqslant U_f(P)$. Es gilt also

$$F(b){-}F(a){-}\varepsilon \leqslant U_f(P){-}\varepsilon < L_f(P) \leqslant \textstyle\int_a^b f(x)dx \leqslant U_f(P) < L_f(P){+}\varepsilon \leqslant F(b){-}F(a){+}\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt daraus $\int_a^b f(x)dx = F(b) - F(a)$. □

Der obige Satz ist von überragender praktischer Bedeutung. Im Beweis sieht man sehr gut, wie all die vorherigen Schritte im systematischen Aufbau der gesamten Theorie (Mittelwertsatz, Extremwertsatz von Weierstraß, Intervallschachtelung, ...) ineinandergreifen.

Beispiel 34.9. (a) Sei $n \in \mathbb{N}$ und $f(x) := x^n$ für $x \in \mathbb{R}$. Ist n ungerade, so ist f monton wachsend auf $\mathbb{R}$; ist n gerade, so ist f monoton wachsend auf $\mathbb{R}_{\geqslant 0}$ und monoton fallend auf $\mathbb{R}_{\leqslant 0}$. Auf jeden Fall ist f stückweise monoton und damit integrierbar. Nach Tabelle 4 ist die Funktion $F(x) := x^{n+1}/(n+1)$ eine Stammfunktion von f. Also gilt

$$\int_a^b x^n dx = F(b) - F(a) = \frac{1}{n+1}(b^{n+1} - a^{n+1}) \qquad \text{für alle } a, b \in \mathbb{R} \text{ mit } a < b.$$

(b) Sei $f(x) := \sqrt{1-x^2}$ für $x \in [0,1]$. Dann ist f monoton fallend, also integrierbar. Der Flächeninhalt zwischen der x-Achse und dem Graph von f ist genau ein Viertel des Einheitskreises, also sollte gelten: $\boxed{\pi = 4\int_0^1 \sqrt{1-x^2}\,dx.}$ Dies können wir nun wie folgt zeigen; beachte, dass wir π in Definition 31.9 als die eindeutige Nullstelle von cos im Intervall $(0,2)$ definiert haben. Nach Beispiel 33.9(c) ist eine Stammfunktion von f gegeben durch $F(x) := \frac{1}{2}x\sqrt{1-x^2} + \frac{1}{2}\arcsin(x)$. Nun ist $F(1) = \frac{1}{2}\arcsin(1) = \frac{\pi}{4}$ und $F(0) = \frac{1}{2}\arcsin(0) = 0$, also tatsächlich $4\int_0^1 f(x)dx = 4F(1) - 4F(0) = \pi$. — Damit sind schließlich die “analytische” und die “anschauliche” Definition von π in Einklang gebracht!

Beispiel 34.10. Zur Erinnerung: Es gilt $1+q+q^2+\ldots+q^n = \frac{q^{n+1}-1}{q-1}$ für alle $n \in \mathbb{N}$ und $1 \neq q \in \mathbb{R}$ (siehe Beispiel 28.2). Wir können dies benutzen, um eine Reihenentwicklung für π zu finden, wie folgt. Sei $t \in \mathbb{R}$ beliebig. Mit $q := -t^2$ und einer leichten Umformung erhalten wir aus der obigen Gleichung

$$f(t) := \frac{1}{1+t^2} = 1 - t^2 + t^4 - \ldots + (-1)^n t^{2n} + (-1)^{n+1}\frac{t^{2n+2}}{1+t^2}.$$

Die Funktion $f(t)$ ist monoton fallend auf $\mathbb{R}_{\geqslant 0}$, also integrierbar auf $[0,x]$ für jedes $x \in \mathbb{R}_{\geqslant 0}$. Mit Tabelle 4 und Satz 34.8 folgt $\arctan(x) = \int_0^x \frac{1}{1+t^2}\,dt$. (Beachte $\arctan(0) = 0$.) Außerdem ist $\int_0^x t^d\,dt = \frac{x^{d+1}}{d+1}$ für alle $d \in \mathbb{N}$. Damit ergibt sich

$$\arctan(x) = \int_0^x \frac{dt}{1+t^2} = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + \underbrace{(-1)^{n+1}\int_0^x \frac{t^{2n+2}}{1+t^2}\,dt}_{R_n(x):=}.$$

Hier gilt $|R_n(x)| = \int_0^x \frac{t^{2n+2}}{1+t^2}\,dt \leqslant \int_0^x t^{2n+2}\,dt = \frac{x^{2n+3}}{2n+3}$ und die rechte Seite konvergiert gegen 0 für $n \to \infty$, falls $0 \leqslant x \leqslant 1$. Damit erhalten wir die Potenzreihen-Entwicklung

$$\arctan(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \pm \ldots \qquad \text{für } 0 \leqslant x \leqslant 1.$$

(Analog auch für $-1 \leqslant x \leqslant 0$.) Nun ist $\cos(\pi/4) = \sin(\pi/4) = \sqrt{2}/2$, also $\tan(\pi/4) = 1$ und

$$\pi = 4\arctan(1) = 4 \cdot \left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} \pm \ldots\right) \qquad \text{(“Leibniz–Reihe”, } \sim 1670).$$

Satz 34.11 (***Hauptsatz der Differential- und Integralrechnung, 2. Teil***). *Sei $f\colon [a,b] \to \mathbb{R}$ stetig. Dann ist f integrierbar und wir erhalten eine Funktion*

$$F\colon [a,b] \to \mathbb{R} \qquad \textit{mit} \qquad F(x) := \int_a^x f(t)\,dt \qquad \textit{für alle } x \in [a,b].$$

Diese Funktion F ist stetig auf $[a,b]$ und differenzierbar für alle $x \in [a,b]$, mit $F'(x) = f(x)$.

Beweis. Siehe zum Beispiel §19 im 1. Band des Buches von Forster. □

Auch wenn der obige Satz eine reine Existenzaussage ist (er sagt nicht, wie man eine Formel für F findet), so kann man daraus dennoch einige nützliche Konsequenzen ziehen.

Folgerung 34.12. *Seien $f, g\colon [a,b] \to \mathbb{R}$ stetige Funktionen. Seien $A, B \in \mathbb{R}$. Dann ist $Af + Bg\colon [a,b] \to \mathbb{R}$ integrierbar und es gilt $\int_a^b (Af+Bg)(x)\,dx = A\int_a^b f(x)\,dx + B\int_a^b g(x)\,dx$.*

Beweis. Nach Satz 34.11 gibt es differenzierbare Funktionen $F, G\colon [a,b] \to \mathbb{R}$ mit $F'(x) = f(x)$ und $G'(x) = g(x)$ für alle $x \in [a,b]$. Dann ist auch $H := AF + BG\colon [a,b] \to \mathbb{R}$ differenzierbar mit $H'(x) = (Af+Bg)(x)$ für alle $x \in [a,b]$. Mit Satz 34.8 folgt $\int_a^b (Af+Bg)(x)\,dx = H(b) - H(a) = A\big(F(b) - F(a)\big) + B\big(G(b) - G(a)\big) = A\int_a^b g(x)\,dx + B\int_a^b g(x)\,dx$. □

Folgerung 34.13 (***Positive Definitheit***)**.** *Sei $f\colon [a,b] \to \mathbb{R}$ stetig mit $f(x) \geqslant 0$ für $x \in [a,b]$. Dann gilt $\int_a^b f(x)\,dx \geqslant 0$. Ist $\int_a^b f(x)\,dx = 0$, so folgt $f(x) = 0$ für alle $x \in [a,b]$.*

Beweis. Sei $F\colon [a,b] \to \mathbb{R}$ wie in Satz 34.11. Wegen $F'(x) = f(x) \geqslant 0$ für $x \in [a,b]$ ist F monoton wachsend (siehe Folgerung 32.12). Wegen $F(a) = 0$ folgt also $F(x) \geqslant 0$ für $x \in [a,b]$; insbesondere $\int_a^b f(x)dx = F(b) \geqslant 0$. Sei nun $F(b) = \int_a^b f(x)dx = 0$. Wegen der Monotonie folgt dann $F(x) = 0$ für alle $x \in [a,b]$, also auch $f(x) = F'(x) = 0$ für alle $x \in [a,b]$. □

Die Aussage wird im Allgemeinen falsch, wenn f nicht stetig ist; Beispiel siehe Übungen.

Ab hier Woche 10

Folgerung 34.14 (***Substitutionsregel***)**.** *Sei* $f\colon D \to \mathbb{R}$ *stetig. Sei* $\varphi\colon [a,b] \to \mathbb{R}$ *differenzierbar und* φ' *stetig, mit* $\varphi([a,b]) \subseteq D$ *und* $\varphi(a) < \varphi(b)$. *Dann gilt*

$$\int_a^b \varphi'(t)f\big(\varphi(t)\big)dt = \int_{\varphi(a)}^{\varphi(b)} f(x)dx.$$

Beweis. Da f stetig auf $[\varphi(a), \varphi(b)] \subseteq D$ ist, gibt es nach Satz 34.11 eine Stammfunktion $F\colon [\varphi(a), \varphi(b)] \to \mathbb{R}$. Nach der Kettenregel ist $F \circ \varphi\colon [a,b] \to \mathbb{R}$ eine Stammfunktion der Funktion $\varphi'(t)f\big(\varphi(t)\big)$ für $t \in [a,b]$. Die letztere Funktion ist auch stetig, also integrierbar. Damit folgt $\int_a^b \varphi'(t)f\big(\varphi(t)\big)dt = (F\circ\varphi)(b) - (F\circ\varphi)(a) = F(\varphi(b)) - F(\varphi(a))$ nach Satz 34.8; analog ist die rechte Seite gleich $\int_{\varphi(a)}^{\varphi(b)} f(x)dx$. □

Die Idee dabei ist, dass man durch eine geschickte "Substitution" $x = \varphi(t)$ die ursprüngliche Funktion $f(x)$ in eine (hoffentlich) einfacher zu behandelnde Funktion $\tilde{f}(t) := f(\varphi(t))$ in der Variablen t überführt. Betrachten wir als Beispiel noch einmal $f(x) := \sqrt{1-x^2}$ für $t \in [0,1]$. Wir setzen $\varphi(t) := \sin(t)$ für $x \in [0, \frac{\pi}{2}]$. Dann ist $\varphi(t) \in [0,1]$ für alle $t \in [0, \frac{\pi}{2}]$, sowie $\varphi(0) = 0$ und $\varphi(\frac{\pi}{2}) = 1$. Nun ist $\cos(t) \geqslant 0$ für $t \in [0, \frac{\pi}{2}]$, also $\sqrt{1-\sin(t)^2} = \sqrt{\cos(t)^2} = \cos(t)$. Es folgt $\int_0^1 f(x)dx = \int_0^{\pi/2} \varphi'(t)f(\varphi(t))dt = \int_0^{\pi/2} \cos(t)^2 dt$. Mit dem Additionstheorem für cos gilt $\cos(2t) = \cos(t)^2 - \sin(t)^2 = 2\cos(t)^2 - 1$, also $\cos(t)^2 = \frac{1}{2}(\cos(2t) + 1)$. Eine Stammfunktion ist gegeben durch $G(t) := \frac{1}{4}(\sin(2t) + 2t)$. Damit erhalten wir

$$\int_0^1 \sqrt{1-x^2}dx = \int_0^{\pi/2} \cos(t)^2 dt = G(\pi/2) - G(0) = \tfrac{1}{4}(\sin(\pi) + \pi) - 0 = \tfrac{1}{4}(0 + \pi) = \pi/4.$$

35. *Uneigentliche Integrale, Bogenlänge und höhere Ableitungen*

Dieser Abschnitt enthält eine (kleine) Auswahl an Ergänzungen zu Integralen und Ableitungen. Analog zur bestimmten Divergenz von Folgen (siehe Definition 27.16) und Funktionen (siehe Definition 31.14) haben wir die folgende Definition für Integrale:

Definition 35.1 (***Uneigentliche Integrale***, 1. Fall)**.** Gegeben sei $f\colon \mathbb{R}_{\geqslant a} \to \mathbb{R}$ (mit $a \in \mathbb{R}$), so dass f integrierbar über jedem Intervall $[a,b]$ mit $a < b \in \mathbb{R}$ ist; sei dann $F(b) := \int_a^b f(x)dx$. Existiert der Grenzwert von $F(b)$ für $b \to +\infty$, so schreiben wir

$$\int_a^\infty f(x)dx := \lim_{b\to+\infty} \int_a^b f(x)dx.$$

(Analog definiert man $\int_{-\infty}^a f(x)dx$ für eine Funktion $f\colon \mathbb{R}_{\leqslant a} \to \mathbb{R}$.)

Beispiel 35.2. (1) Es gilt $\int_a^b \exp(x)dx = \exp(b) - \exp(a)$ für alle $a \leqslant b$. Sei nun $b = 0$. Nach Beispiel 31.14(b) gilt $\lim\limits_{a\to-\infty} \exp(a) = 0$. Also folgt $\int_{-\infty}^0 \exp(x)dx = \exp(0) - 0 = 1$.

(2) Sei $f(x) := x^{-3/2} = 1/\sqrt{x}^3$ für $x \geqslant 1$. Eine Stammfunktion ist $F(x) := -2x^{-1/2}$. Also ist $\int_1^b f(x)dx = F(b) - F(1) = 2 - 2b^{-1/2}$. Wegen $\lim\limits_{b\to+\infty}(b^{-1/2}) = 0$ folgt $\int_1^\infty f(x)dx = 2$.

Beispiel 35.3. Sei $f(t) := \exp(-t^2)$ für $t \in \mathbb{R}$. Dann ist f stetig, also integrierbar. Man kann zeigen, dass $\int_0^\infty \exp(-t^2)dt = \frac{\sqrt{\pi}}{2}$ gilt; siehe zum Beispiel §20, (20.11), im 1. Band des Buches von Forster. Die ***Gaußsche Fehlerfunktion*** ist definiert durch

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)dt \qquad \text{für } x \in \mathbb{R}.$$

("erf" für "error function".) Diese Funktion spielt eine wichtige Rolle in der Wahrscheinlichkeitstheorie und Statistik. Es gibt keine einfache Formel für $\operatorname{erf}(x)$ (siehe Bemerkung 33.11); zur numerischen Berechnung von $\operatorname{erf}(x)$ gibt es die Reihen-Entwicklung in Tabelle 5 (S. 69).

Definition 35.4 (***Uneigentliche Integrale***, 2. Fall)**.** Gegeben sei $f\colon [a,b) \to \mathbb{R}$, mit $a < b$, so dass f integrierbar über jedem Intervall $[a, b-\varepsilon]$ mit $a < b - \varepsilon$ und $\varepsilon > 0$ ist; sei dann $F(\varepsilon) := \int_a^{b-\varepsilon} f(x)dx$. Existiert der Grenzwert $\lim\limits_{\varepsilon\to 0} F(\varepsilon)$, so schreiben wir

$$\int_a^b f(x)dx := \lim_{\varepsilon\to 0} \int_a^{b-\varepsilon} f(x)dx.$$

(Analog definiert man $\int_a^b f(x)dx$ für eine Funktion $f\colon (a,b] \to \mathbb{R}$.)

Beispiel 35.5. (1) Es gilt $\int_\varepsilon^1 \frac{1}{x}dx = \log(1) - \log(\varepsilon) = -\log(\varepsilon)$ für $0 < \varepsilon < 1$. Wegen $\lim\limits_{\varepsilon\to 0} \log(\varepsilon) = -\infty$ existiert der Grenzwert $\int_0^1 \frac{1}{x}dx$ nicht.

(2) Sei $f(x) := 1/\sqrt{1-x^2}$ für $0 \leqslant x < 1$. Eine Stammfunktion ist gegeben durch $F(x) := \arcsin(x)$ (siehe Tabelle 4). Also ist $\int_0^{1-\varepsilon} f(x)dx = F(1-\varepsilon) - F(0) = \arcsin(1-\varepsilon)$. Wegen $\lim\limits_{\varepsilon\to 0} \arcsin(1-\varepsilon) = \arcsin(1) = \pi/2$ folgt $\int_0^1 f(x)dx = \pi/2$.

Wir zeigen nun, dass man Integrale nicht nur dazu benutzen kann, um Flächeninhalte zu berechnen, sondern auch die Länge von Kurvenstücken. Wir betrachten hier nur den Spezialfall des Graphs einer Funktion $f\colon [a,b] \to \mathbb{R}$. (Für allgemeinere Situationen siehe zum Beispiel §4 im 2. Band des Buches von Forster). Durch die Funktion

$$\tilde{f}\colon [a,b] \to \mathbb{R}\times\mathbb{R}, \qquad x \mapsto \big(x, f(x)\big)$$

wird eine Kurve beschrieben, die genau den Graph $\mathcal{G}_f[a,b] := \{(x, f(x)) \mid x \in [a,b]\}$ von f durchläuft. Um die Länge des Kurvenstücks $\mathcal{G}_f[a,b]$ zu bestimmen, gehen wir ähnlich wie bei Integralen vor. Sei P eine Zerlegung von $[a,b]$, mit $|P| = n+1$ und $\underline{P}\colon a = t_0 < x_1 < x_2 < \ldots < x_n = b$. Zwischen zwei Punkten x_{k-1} und x_k ersetzen wir die Funktion $f(x)$ durch das Geradenstück durch die Punkte $(x_{k-1}, f(x_{k-1})) \in \mathcal{G}_f[a,b]$ und $(x_k, f(x_k)) \in \mathcal{G}_f[a,b]$. Nach dem Satz von Pythagoras ist die Länge dieses Geradenstücks gegeben durch

$$l_k := \sqrt{(x_k - x_{k-1})^2 + (f(x_k) - f(x_{k-1}))^2} \geqslant 0.$$

Wir können die Summe $A_f(P) := \sum_{k=1}^n l_k$ als eine Approximation an die Länge von $\mathcal{G}_f[a,b]$ betrachten. Sei nun $\mathcal{A}_f := \{A_f(P) \mid P \text{ Zerlegung von } [a,b]\} \subseteq \mathbb{R}$; diese Menge ist durch 0 nach unten beschränkt. Die Kurve $\mathcal{G}_f[a,b]$ heißt ***rektifizierbar***, wenn $\mathcal{A}_f$ auch nach oben beschränkt ist; in diesem Fall heißt $|\mathcal{G}_f[a,b]| := \sup(\mathcal{A}_f)$ die ***Bogenlänge*** von $\mathcal{G}_f[a,b]$.

Satz 35.6. *Sei* $f\colon [a,b] \to \mathbb{R}$ *und die Ableitung* $f'\colon [a,b] \to \mathbb{R}$ *stetig. Dann ist die Kurve* $\mathcal{G}_f[a,b] = \{(x,f(x)) \mid x \in [a,b]\}$ *rektifizierbar und ihre Bogenlänge gegeben durch das Integral*

$$|\mathcal{G}_f[a,b]| = \int_a^b \sqrt{1 + f'(x)^2}\,dx.$$

Beweis. Da $f'\colon [a,b] \to \mathbb{R}$ stetig ist, ist auch $g(x) := \sqrt{1+f'(x)^2}$ für $x \in [a,b]$ stetig und damit integrierbar; also existiert $L := \int_a^b \sqrt{1+f'(x)^2}\,dx = \int_a^b g(x)dx$. Sei nun $\varepsilon > 0$ beliebig. Nach dem Riemann–Kriterium gibt es eine Zerlegung P von $[a,b]$ mit $U_g(P) - L_g(P) < \varepsilon$; sei $\underline{P}\colon a = x_0 < x_1 < \ldots < x_n = b$ wie oben. Nach dem Mittelwertsatz der Differentialrechnung gibt es $y_k \in (x_{k-1}, x_k)$ mit $f(x_k) - f(x_{k-1}) = f'(y_k)(x_k - x_{k-1})$, für $k = 1, \ldots, n$. Damit folgt

$$A_f(P) = \sum_{k=1}^n \sqrt{(x_k - x_{k-1})^2 + f'(y_k)^2(x_k - x_{k-1})^2} = \sum_{k=1}^n \sqrt{1+f'(y_k)^2}\,(x_k - x_{k-1}).$$

Setze $m_k := \inf\{g(x) \mid x \in [x_{k-1}, x_k]\}$ und $M_k := \sup\{g(x) \mid x \in [x_{k-1}, x_k]\}$ für $k = 1, \ldots, n$. Dann ist $m_k \leqslant g(y_k) \leqslant M_k$ und damit $L_g(P) \leqslant A_f(P) \leqslant U_g(P)$. Nun ist auch $L_g(P) \leqslant L \leqslant U_g(P)$. Also folgt $A_f(P) - L \leqslant A_f(P) - L_g(P) \leqslant U_g(P) - L_g(P) < \varepsilon$. Da $\varepsilon > 0$ beliebig war, folgt $A_f(P) \leqslant L$. Insbesondere ist $\mathcal{A}_f$ nach oben beschränkt mit $|\mathcal{G}_f[a,b]| = \sup(\mathcal{A}_f) \leqslant L$. Andererseits ist $U_g(P) - A_f(P) \leqslant U_g(P) - L_g(P) < \varepsilon$, also $|\mathcal{G}_f[a,b]| \geqslant A_f(P) > U_g(P) - \varepsilon \geqslant L - \varepsilon$. Wiederum, da $\varepsilon > 0$ beliebig war, folgt auch $|\mathcal{G}_f[a,b]| \geqslant L$ und damit $|\mathcal{G}_f[a,b]| = L$. $\square$

Beispiel 35.7. (a) Sei $f(x) := x$. Die Bogenlänge des Geradenstücks von 0 bis 1 ist gegeben durch $|\mathcal{G}_f[0,1]| = \int_0^1 \sqrt{1+f'(x)^2}\,dx = \int_0^1 \sqrt{2}\,dx = \sqrt{2}$ (wie zu erwarten).

(b) Sei $f(x) := x^2$ für $x \in [0,c]$ mit $c \in \mathbb{R}_{>0}$. Die Bogenlänge des Parabelstücks von 0 bis c ist gegeben durch $|\mathcal{G}_f[0,c]| = \int_0^c \sqrt{1+4x^2}\,dx$. Mit Sage finden wir die Stammfunktion $F(x) = \frac{x}{2}\sqrt{1+4x^2} + \frac{1}{4}\mathrm{Arsinh}(2x)$. Also ist $|\mathcal{G}_f[0,c]| = F(c)$, ein ziemlich komplizierter Ausdruck!

Beispiel 35.8. Sei $K_1 := \{(a,b) \in \mathbb{R} \times \mathbb{R} \mid a^2 + b^2 = 1\}$ der Einheitskreis. Die Funktion $f(x) := \sqrt{1-x^2}$ für $x \in [0,1]$ beschreibt genau ein Viertel von K_1 (in der rechten oberen Ebene). In Beispiel 34.9 haben wir bereits gesehen, dass $\pi = 4\int_0^1 f(x)dx$ gilt. Sei nun $(a,b) \in K_1$ mit $0 \leqslant a, b \leqslant 1$. Nach Beispiel 31.13 gibt es einen eindeutigen "Winkel" $\theta \in [0, \pi/2]$ mit $a = \cos(\theta)$ und $b = \sin(\theta)$. Die Bogenlänge des Kreisstücks zwischen den Punkten $(a,b) \in K_1$ und $(1,0) \in K_1$ ist dann gegeben durch

$$|\mathcal{G}_f[a,1]| = \int_a^1 \sqrt{1+f'(x)^2}\,dx = \int_a^1 \sqrt{1+\Big(-\frac{x}{\sqrt{1-x^2}}\Big)^2}\,dx = \int_a^1 \frac{1}{\sqrt{1-x^2}}\,dx.$$

Um das Integral auf der rechten Seite auszuwerten, müssen wir wieder den Grenzwert für $x \to 1$ betrachten, wie in Beispiel 35.5(2). Mit Tabelle 4 erhalten wir dann

$$|\mathcal{G}_f[a,1]| = \arcsin(1) - \arcsin(a) = \arccos(a) - \arccos(1) = \arccos(a) = \theta.$$

Damit haben wir schließlich eine formal korrekte Definition des Winkels θ im "***Bogenmaß***", als die Länge des Kreisstücks zwischen $(a,b) \in K_1$ und $(1,0) \in K_1$. Insbesondere für $a = 0$ ist $\theta = \pi/2$, also ein Viertel des Kreisumfangs — und damit 2π der gesamte Kreisumfang.

Beispiel 35.9. Sei $c \in \mathbb{R}$, $c > 0$, fest. Dann ist $K_c := \{(x,y) \in \mathbb{R}\times\mathbb{R} \mid (cx)^2 + y^2 = c^2\}$ eine Ellipse mit den Halbachsen 1 (auf der x-Achse) und c (auf der y-Achse); es gilt $(\pm 1, 0) \in K_c$ und $(0, \pm c) \in K_c$. Die Funktion $f(x) := c\sqrt{1-x^2}$ für $x \in [0,1]$ beschreibt wieder genau ein Viertel von K_c. Der Flächeninhalt der gesamten Ellipse ist gegeben durch

$$4\int_0^1 f(x)dx = 4\int_0^1 c\sqrt{1-x^2}dx = 4c\int_0^1 \sqrt{1-x^2}dx = c\pi,$$

wobei wir die Formel in Beispiel 34.9 benutzen. Der Umfang der Ellipse ist gegeben durch

$$4|\mathcal{G}_f[0,1]| = 4\int_0^1 \sqrt{1+f'(x)^2}\,dx = 4\int_0^1 \sqrt{1+\Big(-\frac{cx}{\sqrt{1-x^2}}\Big)^2}\,dx = 4\int_0^1 \sqrt{\frac{1+(c^2-1)x^2}{1-x^2}}\,dx.$$

Für $c \neq 1$ gibt es im Allgemeinen keine geschlossene Formel für dieses Integral! (Für mehr dazu siehe auch https://de.wikipedia.org/wiki/Elliptisches_Integral.)

Definition 35.10 (***Höhere Ableitungen***)**.** Sei $f\colon D \to \mathbb{R}$ differenzierbar auf D; wir erhalten also eine neue Funktion $f'\colon D \to \mathbb{R}$. Ist f' differenzierbar in $a \in D$, so schreiben wir $f''(a)$ oder $f^{(2)}(a)$ für $(f')'(a)$; in diesem Fall heißt $f''(a)$ die 2-te Ableitung von f in a.
Allgemeiner definieren wir $f^{(n)}\colon D \to \mathbb{R}$ für $n \in \mathbb{N}_0$ rekursiv wie folgt: Für $n = 0$ sei $f^{(0)} := f$; ist $n \geqslant 1$ und $f^{(n-1)}$ bereits definiert und differenzierbar auf D, so setze $f^{(n)}(x) := (f^{(n-1)})'(x)$ für $x \in D$; die Funktionen $f^{(n)}$ heißen ***n-te Ableitungen*** von f. Schließlich heißt f ***unendlich oft differenzierbar***, wenn $f^{(n)}\colon D \to \mathbb{R}$ für alle $n \in \mathbb{N}$ existiert.

Zum Beispiel ist die Exponential-Funktion unendlich oft differenzierbar, mit $\exp^{(n)}(x) = \exp(x)$ für alle $x \in \mathbb{R}$ und $n \in \mathbb{N}$.

Beispiel 35.11. Sei $\alpha \in \mathbb{R}$ fest; betrachte die Potenzfunktion $f(x) = x^\alpha$ für $x > 0$. Es gilt $f'(x) = \alpha x^{\alpha-1}$, $f''(x) = \alpha(\alpha-1)x^{\alpha-2}$ und allgemein

$$f^{(n)}(x) = \alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)\,x^{\alpha-n} \qquad \text{für alle } n \geqslant 1.$$

Gegeben sei nun eine Polynomfunktion $p(x) = a_0 + a_1x + \cdots + a_dx^d$, wobei $a_i \in \mathbb{R}$ und $d \in \mathbb{N}_0$. Mit den obigen Formeln folgt $p^{(n)}(x) = n!a_n +$ Linear-Kombination von höheren Potenzen von x. Dies ergibt $p^{(n)}(0) = n!a_n$ für $0 \leqslant n \leqslant d$. Insbesondere ist $p(x)$ selbst eindeutig durch die höheren Ableitungen von $p(x)$ an der Stelle $a = 0$ bestimmt; es gilt:

$$p(x) = \sum_{n=0}^{d} \frac{p^{(n)}(0)}{n!}\, x^n = p(0) + p'(0)x + \frac{p''(0)}{2!}x^2 + \cdots + \frac{p^{(d)}(0)}{d!}x^d \qquad \text{für alle } x \in \mathbb{R}.$$

Diese Idee läßt sich auf allgemeinere Funktionen wie folgt übertragen. Sei $f\colon D \to \mathbb{R}$ eine n-mal differenzierbare Funktion; sei $a \in D$. Für $x \in D$ definieren wir das ***Taylor–Polynom*** (vom Grad n) sowie den zugehörigen Restterm durch

$$T_{n,a}f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(x-a)^k \qquad \text{bzw.} \qquad R_{n,a}f(x) = f(x) - T_{n,a}f(x).$$

(Es gilt $T_{n,a}f(a) = f(a)$, $R_{n,a}f(a) = 0$; außerdem $T_{0,a}f(x) = f(a)$, $R_{0,a}f(x) = f(x) - f(a)$.)

Satz 35.12 (***Taylor–Formel***)**.** *Sei $D \subseteq \mathbb{R}$ ein Intervall und $f\colon D \to \mathbb{R}$ eine $(n+1)$-mal differenzierbare Funktion, so dass $f^{(n+1)}\colon D \to \mathbb{R}$ noch stetig auf D ist. Sei $a \in D$. Dann gelten folgende Formeln für den Restterm $R_{n,a}f(x)$:*

$$R_{n,a}f(x) = \frac{1}{n!}\int_a^x f^{(n+1)}(t)\,(x-t)^n\,dt, \qquad \text{“Integral-Form”},$$

$$R_{n,a}f(x) = \frac{f^{(n+1)}(c_x)}{(n+1)!}\,(x-a)^{n+1}, \qquad \text{“Lagrange–Form”},$$

wobei $x \in D$ und $c_x \in D$ mit $|c_x - a| \leqslant |x - a|$.

Beweis. Sei $a \neq x \in D$ fest und setze $F(t) := -f(x) + \sum_{k=0}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k$ für $t \in D$. Dann gilt $F(x) = -f(x) + f(x) = 0$ und $F(a) = -f(x) + T_{n,a}f(x)$. Wir leiten F nach t ab und erhalten

$$F'(t) = f'(t) + \sum_{k=1}^{n} \frac{f^{(k+1)}(t)}{k!}(x-t)^k - \sum_{k=1}^{n} \frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} = \ldots = \frac{f^{(n+1)}(t)}{n!}(x-t)^n.$$

Weil $f^{(n+1)}$ stetig ist, ist auch $F'(t)$ stetig und damit integrierbar. Also gilt nach dem 1. Teil des Hauptsatzes der Differential- und Integralrechnung:

$$R_{n,a}f(x) = f(x) - T_{n,a}f(x) = F(x) - F(a) = \int_a^x F'(t)dt = \frac{1}{n!}\int_a^x f^{(n+1)}(t)\,(x-t)^n\,dt.$$

Lagrange–Form: Setze $H(t) := -F(t)(x-a)^{n+1} - \big(F(x) - F(a)\big)(x-t)^{n+1}$ für $t \in D$. Dann ist H differenzierbar für alle $t \in D$. Man rechnet sofort nach, dass $H(x) = H(a)$ gilt. Nach dem Mittelwertsatz der Differentialrechnung gibt es ein $c_x \in D$ zwischen a und x mit $0 = H(x) - H(a) = H'(c_x)(x-a)$, also $H'(c_x) = 0$. Nun folgt

$$\begin{aligned} 0 = H'(c_x) &= -F'(c_x)(x-a)^{n+1} - \big(F(x) - F(a)\big)\big(-(n+1)(x-c_x)^n\big) \\ &= -\tfrac{f^{(n+1)}(c_x)}{n!}(x-c_x)^n(x-a)^{n+1} + R_{n,a}f(x)(n+1)(x-c_x)^n, \end{aligned}$$

und damit die gewünschte Formel. $\square$

Ist f unendlich oft differenzierbar auf D und $x \in D$, so heißt die unendliche Reihe

$$T_{\infty,a}f(x) := \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}\,(x-a)^n$$

die ***Taylor–Reihe*** zu f um die Entwicklungsstelle $a \in D$. Diese konvergiert genau dann gegen den Funktionswert $f(x)$, wenn $\lim\limits_{n\to\infty} R_{n,a}f(x) = 0$ gilt.

Tabelle 5. Taylor–Reihen für einige oft benutzte Funktionen

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \qquad \text{für alle } x \in \mathbb{R},$$

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\,x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \cdots \qquad \text{für alle } x \in \mathbb{R},$$

$$\cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\,x^{2n} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \cdots \qquad \text{für alle } x \in \mathbb{R},$$

$$\tan(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}4^n(4^n-1)B_{2n}}{(2n)!}\,x^{2n-1} = x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \cdots \qquad \text{für } |x| < \tfrac{\pi}{2},$$

wobei B_{2n} die ***Bernoulli–Zahlen*** sind, rekursiv definiert durch $B_0 := 1$ und $B_0 + \binom{n+1}{1}B_1 + \binom{n+1}{2}B_2 + \cdots + \binom{n+1}{n-1}B_{n-1} + \binom{n+1}{n}B_n = 0$ für $n > 0$,

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n}\,x^n \qquad \text{für } |x| < 1 \text{ und } \alpha \in \mathbb{R} \text{ beliebig},$$

wobei $\binom{\alpha}{0} := 1$ und $\binom{\alpha}{n} := \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{1\cdot 2\cdot 3\cdot\ldots\cdot n}$ für $n \geqslant 1$,

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{n!(2n+1)} = \frac{2}{\sqrt{\pi}}\Big(x - \frac{x^3}{3} + \frac{x^5}{10} - \frac{x^7}{42} + \frac{x^9}{216} \pm \ldots\Big) \qquad \text{für alle } x \in \mathbb{R},$$

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\,x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \pm \cdots \qquad \text{für } x \in (-1,1],$$

$$\arcsin(x) = \sum_{n=0}^{\infty} \frac{(2n)!}{4^n(n!)^2}\,\frac{x^{2n+1}}{2n+1} = x + \frac{1}{2}\,\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\,\frac{x^5}{5} + \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\,\frac{x^7}{7} + \cdots \qquad \text{für } |x| \leqslant 1,$$

$$\arccos(x) = \frac{\pi}{2} - \arcsin(x) \qquad \text{für } |x| \leqslant 1,$$

$$\arctan(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}\,x^{2n+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \pm \cdots \qquad \text{für } |x| \leqslant 1.$$

Beispiel 35.13. (a) Die Funktion $f(x) := \sin(x)$ ist unendlich oft differenzierbar, mit $\sin^{(n)}(x) = (-1)^{(n-1)/2}\cos(x)$ (falls n ungerade) und $\sin^{(n)}(x) = (-1)^{n/2}\sin(x)$ (falls n gerade). Insbesondere ist $|f^{(n)}(x)| \leqslant 1$ für alle $n \in \mathbb{N}_0$. Mit diesen Formeln folgt dann

$$T_{2m+2,0}f(x) = \sum_{k=0}^{m} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \quad \text{und} \quad \big|R_{2m+2,0}f(x)\big| = \frac{|f^{(2m+3)}(c_x)||x|^{2m+3}}{(2m+3)!} \leqslant \frac{|x|^{2m+3}}{(2m+3)!}$$

für alle $x \in \mathbb{R}$ und $m \in \mathbb{N}_0$, also eine etwas bessere Abschätzung für den Rest- oder Fehlerterm $|R_{2m+2,0}f(x)|$ als in Satz 30.9. Analog für $g(x) := \cos(x)$:

$$T_{2m+1,0}g(x) = \sum_{k=0}^{m} (-1)^k \frac{x^{2k}}{(2k)!} \qquad \text{und} \qquad \left|R_{2m+1,0}g(x)\right| \leqslant \frac{|x|^{2m+2}}{(2m+2)!}.$$

(b) Sei $h(x) := \log(1+x)$ für $x > -1$. Mit vollständiger Induktion nach n zeigt man leicht, dass $h^{(n)}(x) = (-1)^{n-1}(n-1)!/(1+x)^n$ für alle $n \in \mathbb{N}$ gilt. Folglich ist $h^{(0)}(0) = \log(1) = 0$ und $h^{(n)}(0) = (-1)^{n-1}(n-1)!$ für $n \geqslant 1$. Wir erhalten damit die Taylor–Reihe

$$T_{\infty,0}h(x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(n-1)!}{n!} x^n = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \pm \ldots.$$

Für $x = 1$ ist diese konvergent (siehe Beispiel 28.15(b)), für $x = -1$ divergent (siehe Beispiel 28.5). Also hat die obige Potenzreihe den Konvergenzraduius $R_0 = 1$ (siehe Satz 29.1). Dies zeigt, dass eine Taylor–Reihe nicht den gleichen Definitionsbereich wie die Funktion selbst haben muss. — Wir müssen jetzt noch den Restterm $R_{n,0}h(x)$ für $n \geqslant 1$ abschätzen. Ist $0 \leqslant x \leqslant 1$, so benutzen wir die Lagrange–Form $R_{n,0}h(x) = \frac{h^{(n+1)}(c_x)}{(n+1)!}x^{n+1} = (-1)^n \frac{x^{n+1}}{n(1+c_x)^{n+1}}$, wobei $0 \leqslant c_x \leqslant x$. Dann ist $1 + c_x \geqslant 1$ und damit $|R_{n,0}h(x)| \leqslant \frac{x^{n+1}}{n} \leqslant \frac{1}{n} \to 0$ für $n \to \infty$. Also gilt $\log(1+x) = T_{\infty,0}h(x)$ für $0 \leqslant x \leqslant 1$. Ist $-1 < x \leqslant 0$, so kann man wie folgt argumentieren: Nach Bemerkung 33.12 ist $T_{\infty,0}h(x)$ für $x \in (-1,1)$ differenzierbar, mit $(T_{\infty,0}h)'(x) = \sum_{n=1}^{\infty}(-1)^{n-1}x^{n-1}$. Nach Beispiel 28.2 ist dies gleich $1/(1+x) = \log'(1+x)$. Wegen $T_{\infty,0}h(0) = 0 = \log(1)$ folgt also $\log(1+x) = T_{\infty,0}h(x)$ für $x \in (-1,1]$ (siehe Folgerung 32.13). Wir sehen jetzt auch, dass $\log(1+x) = -f(-x)$ für $-1 < x \leqslant 1$ gilt, wobei $f(x)$ die Potenzreihe in Beispiel 28.15(b) ist; insbesondere haben wir die Abschätzung des Fehlerterms in Bemerkung 29.3, die besser ist als die, die man aus Satz 35.12 erhält. Speziell für $x = \frac{1}{2}$ erhalten wir die unendliche Reihe:

$$f(\tfrac{1}{2}) = -\log(1-\tfrac{1}{2}) = -\log(\tfrac{1}{2}) = \log(2) = 1 - \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} \pm \ldots \approx 0{,}6931471805599453;$$

wie in Bemerkung 29.3.

Beispiel 35.14. Sei $f(x) := \exp(-1/x^2)$ für $x \in \mathbb{R} \setminus \{0\}$. Es gilt $\lim\limits_{x\to 0}(-1/x^2) = -\infty$; mit Beispiel 31.15(a) folgt $\lim\limits_{x\to 0} f(x) = 0$. Setzen wir also auch $f(0) := 0$, so erhalten wir eine Funktion $f\colon \mathbb{R} \to \mathbb{R}$, die auf ganz $\mathbb{R}$ stetig ist. Man kann dann sogar zeigen, dass f unendlich oft differenzierbar ist, mit $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$; siehe zum Beispiel §22, (22.1), im 1. Band des Buches von Forster. Damit können wir die Taylor–Reihe zu f um die Entwicklungsstelle $a = 0$ bilden und erhalten $T_{\infty,0}f(x) = 0$ für $x \in \mathbb{R}$. Diese Potenzreihe konvergiert also (sogar mit Konvergenzradius $R_0 = \infty$), aber nicht gegen den Funktionswert $f(x)$. — Um eine gegebene Funktion durch eine Taylor–Reihe darzustellen genügt es also nicht, die Taylor–Reihe auf Konvergenz zu untersuchen, sondern man muss auch noch zeigen, dass für $x \in D$ der Restterm $R_{n,a}f(x)$ auch gegen 0 konvergiert für $n \to \infty$.

Ab hier Woche 11

Kapitel VIII: Partielle Ableitungen und Differentialgleichungen

Wir betrachten nun reelle Funktionen in mehreren Variablen, und stellen uns ähnliche Fragen wie zuvor, also zum Beispiel: Kann man sinnvoll Stetigkeit und Ableitungen definieren, so dass analoge Aussagen gelten wie zuvor? Erhalten wir dadurch auch gute Aussagen über das Verhalten der Funktion? Wir werden sehen, dass sich vieles von dem, was wir für Funktionen einer Variablen gemacht haben, auf diesen Fall verallgemeinern lässt. Dadurch eröffnen sich auch weitreichende Anwendungsmöglichkeiten für analytische Methoden, die schießlich in Abschnitten zu Differentialgleichungen zumindest angedeutet werden.

Bezeichnungen: Sei $n \geqslant 1$. Wie bisher besteht $\mathbb{R}^n$ aus Spaltenvektoren. Aber für eine Funktion $f\colon D \to \mathbb{R}$ mit $D \subseteq \mathbb{R}^n$ schreiben wir zur Vereinfachung die Argumente einfach als Tupel, also etwa $f(x) = f(x_1, \dots, x_n)$. Analog für Funktionen $f\colon D \to \mathbb{R}^m$, wobei $m \geqslant 1$.

36. ***Reelle Funktionen von mehreren Variablen***

Bevor wir zu Stetigkeit und Ableitungen von Funktionen in mehreren Variablen kommen, benötigen wir zuerst einige Grundbegriffe zu Folgen und Grenzwerten in $\mathbb{R}^n$. Zur Erinnerung: Nach der Betrachtung von Grenzwerten von Folgen von reellen Zahlen haben wir in Kapitel VI, §30, bereits Grenzwerte für Folgen von komplexen Zahlen definiert; dazu wurde der reelle durch den komplexen Absolutbetrag $|z| = \sqrt{z\bar{z}}$ ersetzt. Völlig analog kann man auch für Folgen in $\mathbb{R}^n$ vorgehen, indem man die übliche Euklidische Norm $\|x\|$ für $x \in \mathbb{R}^n$ benutzt. Es ist zweckmäßig, die weiteren Definitionen und Aussagen gleich im allgemeineren Rahmen von "metrischen Räumen" zu formulieren.

Sei M eine nicht-leere Menge und $d\colon M \times M \to \mathbb{R}$ eine Abbildung. Dann heißt d eine ***Metrik*** auf M, und (M, d) ein ***metrischer Raum***, wenn folgende Eigenschaften erfüllt sind:

Symmetrie: $d(x, y) = d(y, x)$ für alle $x, y \in M$;
Definitheit: $d(x, y) = 0$ genau dann, wenn $x = y$;
Dreiecksungleichung: $d(x, z) \leqslant d(x, y) + d(y, z)$ für alle $x, y, z \in M$.

Beachte: Für $x, y \in M$ folgt automatisch $d(x, y) \geqslant 0$.
(Denn mit obigen 3 Eigenschaften folgt $0 = d(x, x) \leqslant d(x, y) + d(y, x) = 2d(x, y)$.)

Unsere Standard-Beispiele für metrische Räume (M, d) sind:

$M = \mathbb{R}$ mit $d(x, y) := |x - y|$ (reeller Absolutbetrag);
$M = \mathbb{C}$ mit $d(z, z') := |z - z'|$ (komplexer Absolutbetrag);
$M = \mathbb{R}^n$ mit $d(v, w) := \|v - w\|$ (Euklidische Norm).

Im dritten Fall folgt die Dreiecksungleichung aus der Cauchy–Schwarz–Ungleichung in Kapitel IV, Satz 19.2; die Definitheit wurde bereits zu Beginn von Kapitel IV bemerkt.

Sei $(x_k)_{k\in\mathbb{N}_0}$ eine Folge mit $x_k \in M$ für alle $k \in \mathbb{N}_0$. Dann heißt $(x_k)_{k\in\mathbb{N}_0}$ konvergent, wenn es ein $\alpha \in M$ gibt, so dass die analoge Bedingung wie in Kapitel VI, Definition 27.1 gilt, also:

(∗) *Zu jedem* $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, *gibt es ein* $k_0 \in \mathbb{N}_0$ *mit* $d(x_k, \alpha) < \varepsilon$ *für alle* $k \geqslant k_0$.

In diesem Fall schreiben wir $\lim\limits_{k\to\infty}(x_k) = \alpha$; wie in Kapitel IV, Lemma 27.4, sieht man, dass der Grenzwert α dann eindeutig bestimmt ist.

Definition 36.1 (Vgl. Kap. VI, Def. 32.1)**.** Seien (M, d) und (M', d') metrische Räume. Sei $\varnothing \neq D \subseteq M$ und $f\colon D \to M'$ eine Abbildung. Sei $a \in M$ (also nicht unbedingt $a \in D$).
Sei $c \in M'$. Wir schreiben $\lim\limits_{x\to a} f(x) = c$ (oder auch $f(x) \to c$ für $x \to a$), wenn für jede konvergente Folge $(x_k)_{k\in\mathbb{N}_0}$ mit Grenzwert a und $x_k \in D$ für alle $k \in \mathbb{N}_0$ auch die Folge der Funktionswerte $(f(x_k))_{k\in\mathbb{N}_0}$ konvergiert, und zwar immer mit $\lim\limits_{k\to\infty}(f(x_k)) = c$. In diesem Fall heißt c der ***Grenzwert*** von $f(x)$ für $x \to a$.
Ist $a \in D$, so heißt f stetig in a, wenn $\lim\limits_{x\to a} f(x) = f(a)$ gilt. Ist f stetig in jedem $a \in D$, so sagen wir kurz, dass f ***stetig auf*** D ist.

[Beachte: Damit diese Definitionen wirklich sinnvoll sind (vor allem falls $a \notin D$), müssen wir wieder voraussetzen, dass a ein ***Häufungspunkt*** von D ist, d.h., es gibt mindestens eine konvergente Folge $(x_n)_{n\in\mathbb{N}_0}$ mit Grenzwert a und $a \neq x_n \in D$ für alle $n \in \mathbb{N}_0$. Auch hier werden wie dies nicht besonders thematisieren; in allen konkreten Beispielen wird a stets ein Häufungspunkt von D sein.]

Damit haben wir die Grundlagen zusammen, um Stetigkeit und Ableitungen von Funktionen mit mehreren Variablen zu betrachten. Im Folgenden sei stets $\mathbb{R}^n$ der metrische Raum mit der üblichen Euklidischen Norm. Wir halten fest:

Lemma 36.2. *Sei* $(v_k)_{k\in\mathbb{N}_0}$ *eine Folge in* $\mathbb{R}^n$. *Für jedes* k *seien* $v_{k1}, \ldots, v_{kn} \in \mathbb{R}$ *die Komponenten von* v_k. *Genau dann ist* $(v_k)_{k\in\mathbb{N}_0}$ *konvergent in* $\mathbb{R}^n$, *wenn alle* n *reellen Folgen der Komponenten* $(v_{k1})_{k\in\mathbb{N}_0}, \ldots, (v_{kn})_{k\in\mathbb{N}_0}$ *in* $\mathbb{R}$ *konvergieren. In diesem Fall sind die Komponenten von* $\lim\limits_{k\to\infty}(v_k)$ *gegeben durch* $\lim\limits_{k\to\infty}(v_{k1}), \ldots, \lim\limits_{k\to\infty}(v_{kn})$.

Beweis. Dies ist analog zu Kapitel VI, Lemma 30.1, wo der Fall $n = 2$ behandelt wurde (mit der Identifikation von $\mathbb{R}^2$ mit $\mathbb{C}$); der allgemeine Fall $n \geqslant 2$ geht genauso. □

Beispiel 36.3. Gegeben sei die Folge $(v_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^2$ mit $v_k = \begin{bmatrix} 1 - \frac{1}{k} \\ \frac{-2k}{k+1} \end{bmatrix}$ für alle $k \in \mathbb{N}$. Die Folgen der Komponenten sind also gegeben durch $v_{k1} = 1 - \frac{1}{k}$ und $v_{k2} = \frac{-2k}{k+1}$ für $k \in \mathbb{N}$.
Die erste konvergiert gegen 1 und die zweite gegen -2. Also folgt $\lim\limits_{k\to\infty}(v_k) = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \in \mathbb{R}^2$.

Ist dagegen $(v_k)_{k\in\mathbb{N}}$ gegeben durch $v_k = \begin{bmatrix} 1-\frac{1}{k} \\ (-1)^k \end{bmatrix} \in \mathbb{R}^2$ für alle $k \in \mathbb{N}$, so ist die Folge nicht konvergent. Denn die Folge der ersten Komponenten konvergiert zwar weiterhin mit Grenzwert 1, aber die Folge der zweiten Komponenten ist nicht konvergent.

Definition 36.4. Sei $n \geqslant 1$, $D \subseteq \mathbb{R}^n$ und $f\colon D \to \mathbb{R}$. Sei $a = (a_1, \ldots, a_n) \in D$.
Für $1 \leqslant j \leqslant n$ sei $D_{(j)} := \{x \in \mathbb{R} \mid (a_1, \ldots, a_{j-1}, x, a_{j+1}, \ldots, a_n) \in D\}$. Betrachte dann die Funktion $f_{(j)}\colon D_{(j)} \to \mathbb{R}$ (einer reellen Variablen!) mit

$$f_{(j)}(x) := f(a_1, \ldots, a_{j-1}, x, a_{j+1}, \ldots, a_n) \qquad \text{für } x \in D_{(j)} \subseteq \mathbb{R}.$$

(Beachte $D_{(j)} \neq \varnothing$, denn $a_j \in D_{(j)}$.) Ist $f_{(j)}$ differenzierbar in $a_j \in D_{(j)}$, so bezeichnen wir $f'_{(j)}(a_j)$ als die j-te ***partielle Ableitung*** von f im Punkt a.

Schreibweisen: $\qquad D_j f(a) := f'_{(j)}(a_j) \qquad$ oder $\qquad \dfrac{\partial f}{\partial x_j}(a) := f'_{(j)}(a_j).$

Existieren alle n partiellen Ableitungen von f in a, so heißt

$$\nabla f(a) := \big[D_1 f(a), \ldots, D_n f(a)\big] \in \mathbb{R}^{1\times n} \qquad \textbf{\textit{Gradient}} \text{ von } f.$$

Beispiel 36.5. (a) Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x,y) := x^2 \sin(xy)$ für alle $x, y \in \mathbb{R}$. Sei $(a,b) \in \mathbb{R}^2$ fest; wir bestimmen die partiellen Ableitungen $D_1 f(a,b)$ und $D_2 f(a,b)$. Halten wir $y = b$ fest, so ergibt sich die Funktion $f_{(1)}(x) = x^2 \sin(xb)$. Wir können alle Regeln aus dem vorherigen Kapitel anwenden, um $f'_{(1)}(x)$ zu berechnen. Mit Produkt- und Kettenregel ergibt sich $D_1 f(x,b) = f'_{(1)}(x) = 2x \sin(xb) + x^2 b \cos(xb)$ für alle $x \in R$. Halten wir analog $x = a$ fest, so ergibt sich $D_2 f(a,y) = f'_{(2)}(y) = a^3 \cos(ay)$ für alle $y \in \mathbb{R}$. Also

$$\nabla f(a,b) = \big[\ 2a \sin(ab) + a^2 b \cos(ab),\ a^3 \cos(ab)\ \big] \in \mathbb{R}^{1\times 2}.$$

(b) Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x,y) := \sqrt{x^2+y^2}$ für alle $x, y \in \mathbb{R}$. Sei $(a,b) \in \mathbb{R}^2$. Halten wir $y = b$ fest, so ergibt sich die Funktion $f_{(1)}(x) = \sqrt{x^2+b^2}$ für $x \in \mathbb{R}$. Für $b = 0$ ist $f_{(1)}(x) = \sqrt{x^2} = |x|$, also $f'_{(1)}(x) = 1$ für $x > 0$ und $f'_{(1)}(x) = -1$ für $x < 0$; aber $f_{(1)}$ ist nicht differenzierbar in $x = 0$. Sei nun $b \neq 0$. Dann ist $f_{(1)}$ differenzierbar mit $D_1 f(x,b) = \frac{x}{\sqrt{x^2+b^2}}$ für alle $x \in R$. Ein analoges Bild ergibt sich für $D_2 f$. Wir sehen also, dass unter Umständen mehr Fallunterscheidungen nötig sind als bei Funktionen einer Variablen.

(c) Sei $f\colon \mathbb{R}^n \to \mathbb{R}$ gegegen durch $f(x_1, \ldots, x_n) = \sqrt{x_1^2 + \ldots + x_n^2}$ (Euklidische Norm).
Behauptung: Für $0_n \neq a \in \mathbb{R}^n$ ist f partiell differenzierbar mit $D_i f(a) = a_i/\|a\|$ für $i = 1, \ldots, n$ (wobei a_i die i-te Komponente von a ist).
Dazu: Hier ist $f_{(i)}(x) = \sqrt{a_1^2 + \ldots + a_{i-1}^2 + x^2 + a_{i+1}^2 + \ldots + a_n^2}$. Für die Funktion $g(x) := \sqrt{x}$ (mit $x \in \mathbb{R}_{>0}$) gilt $g'(x) = 1/(2\sqrt{x})$; mit der Kettenregel erhalten wir also $f'_{(i)}(x) = (2x)/(2f_{(i)}(x))$, d.h., $f'_{(i)}(a_i) = a_i/f(a)$.

Definition 36.6. Sei $m \geqslant 1$ und $f\colon D \to \mathbb{R}^m$ eine Abbildung. Für jedes $x \in D$ seien $f_1(x), \ldots, f_m(x) \in \mathbb{R}$ die Komponenten von $f(x)$. Die Funktionen $f_i\colon D \to \mathbb{R}$ (für $1 \leqslant i \leqslant m$)

bezeichnen wir als die ***Komponenten*** von f. Sei $a \in D$. Existieren die partiellen Ableitungen $D_j f_i(a)$ für alle i, j, so heißt die Matrix

$$\mathcal{J}_f(a) := \left[D_j f_i(a)\right]_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n} \in \mathbb{R}^{m \times n} \qquad \textbf{\textit{Jacobi–Matrix}} \text{ von } f \text{ im Punkt } a.$$

(Die Zeilen dieser Matrix bestehen also genau aus den Gradienten $\nabla f_i(a)$ für $i = 1, \ldots, m$.)

Beispiel 36.7. (a) Sei $m = 3$, $n = 2$ und $f\colon \mathbb{R}^2 \to \mathbb{R}^3$ gegeben durch

$$f(x, y) := \begin{bmatrix} \sin(2x^3 - 5y^7) \\ x^2 + x + 1 \\ e^{x-y} \end{bmatrix} \qquad \text{für alle } x, y \in \mathbb{R}.$$

Hier sind die Komponenten-Funktionen gegeben durch $f_1(x, y) = \sin(2x^3 - 5y^7)$, $f_2(x, y) = x^2 + x + 1$ und $f_3(x, y) = e^{x-y}$. Alle partiellen Ableitungen können mit den üblichen Regeln berechnet werden; wir erhalten für $(a, b) \in \mathbb{R}^2$:

$$\mathcal{J}_f(a, b) = \begin{bmatrix} 6a^2 \cos(2a^3 - 5b^7) & -35b^6 \cos(2a^3 - 5b^7) \\ 2a + 1 & 0 \\ e^{a-b} & -e^{a-b} \end{bmatrix} \in \mathbb{R}^{3 \times 2}.$$

(b) Sei $A \in \mathbb{R}^{m \times n}$ und $f\colon \mathbb{R}^n \to \mathbb{R}^m$ definiert durch $f(x) := A \cdot x$ für $x \in \mathbb{R}^n$. (Beachte: Für das Produkt $A \cdot x$ ist x ein Spaltenvektor.) Die i-te Komponente $f_i\colon \mathbb{R}^n \to \mathbb{R}$ von f ist gegeben durch $f_i(x_1, \ldots, x_n) = \sum_{j=1}^n a_{ij} x_j = a_{i1} x_1 + \ldots + a_{in} x_n$ für alle $x_j \in \mathbb{R}$. Für $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ folgt $D_j f_i(a) = a_{ij}$ für alle i, j. Also ist $\mathcal{J}_f(a) = A$ für alle $a \in \mathbb{R}^n$.

In Kapitel VI, Satz 32.5 hatten wir gezeigt, dass eine differenzierbare Funktion (einer reellen Variablen) auch stetig ist. Das folgende Beispiel zeigt, dass die Situation für Funktionen mehrerer Variablen komplizierter ist.

Beispiel 36.8. Definiere $f\colon \mathbb{R}^2 \to \mathbb{R}$ durch $f(x, y) := \begin{cases} 1 & \text{falls } x = y \neq 0, \\ 0 & \text{sonst.} \end{cases}$

Wir bestimmen die partiellen Ableitungen in $(0, 0)$. Betrachte zuerst $f_{(1)}(x) := f(x, 0)$; es gilt $f_{(1)}(x) = 0$ für alle $x \in \mathbb{R}$, also $f'_{(1)}(0) = 0$. Genauso ist $f_{(2)}(y) = f(0, y) = 0$ für alle $y \in \mathbb{R}$, also auch $f'_{(2)}(0) = 0$. Also existieren die partiellen Ableitungen in $(0, 0)$, mit $\mathcal{J}_f(0, 0) = [0\ 0]$.

Aber f ist nicht stetig in $(0, 0)$. Denn für die Null-Folge $(1/n, 1/n)_{n \in \mathbb{N}}$ ist $f(1/n, 1/n) = 1$ für alle $n \in \mathbb{N}$, also $\lim\limits_{n \to \infty} f(1/n, 1/n) = 1$; aber für die Null-Folge $(1/n, 0)_{n \in \mathbb{N}}$ ist $f(1/n, 0) = 0$ für alle $n \in \mathbb{N}$, also $\lim\limits_{n \to \infty} f(1/n, 0) = 0$.

Die Existenz aller partiellen Ableitungen ist also zu schwach, um Stetigkeit zu garantieren. Es gibt allerdings eine stärkere Definition für die Differenzierbarkeit einer Funktion in mehreren Variablen, die wir kurz diskutieren wollen. Um diese stärkere Definition zu motivieren, wiederholen wir die genaue Definition der Ableitung für eine Funktion einer reellen Variablen. Sei also $D \subseteq \mathbb{R}$ und $f\colon D \to \mathbb{R}$; sei $a \in D$ und $D_a := D \setminus \{a\}$. Definiere

$$F_a\colon D_a \to \mathbb{R} \qquad \text{mit} \qquad F_a(x) := \frac{f(x)-f(a)}{x-a} \qquad \text{für alle } x \in D_a.$$

Dann ist f differenzierbar in a, wenn $\lim\limits_{x\to a} F_a(x)$ existiert (wobei wir stillschweigend annehmen, dass a ein Häufungspunkt von D ist); siehe Kapitel VII, Definition 32.3. Der folgende Hilfssatz enthält eine leichte Umformulierung der obigen Bedingung:

Lemma 36.9. *Sei $D \subseteq \mathbb{R}$, $f\colon D \to \mathbb{R}$ und $a \in D$ wie oben. Für $c \in \mathbb{R}$ definiere $r_c\colon D_a \to \mathbb{R}$ durch $r_c(x) := f(x) - f(a) - c(x-a) \in \mathbb{R}$ für alle $x \in D_a$. Dann gilt*

$$\begin{array}{c} f \textit{ differenzierbar in } a \\ (\textit{mit } f'(a) = c) \end{array} \quad \Leftrightarrow \quad \lim_{x\to a} \frac{r_c(x)}{x-a} = 0 \quad \Leftrightarrow \quad \lim_{x\to a} \frac{|r_c(x)|}{|x-a|} = 0.$$

Beweis. Für $x \in D_a$ ist $\frac{r_c(x)}{x-a} = F_a(x) - c$. Daraus folgt sofort die erste Äquivalenz. Für die zweite Äquivalenz betrachte eine konvergente Folge $(x_n)_{n\in\mathbb{N}_0}$ mit $x_n \in D_a$ für alle $n \in \mathbb{N}_0$ und Grenzwert a; setze $r_n := \frac{r_c(x_n)}{x_n - a}$ für alle $n \in \mathbb{N}_0$. Dann gilt $\lim\limits_{x\to a} \frac{r_c(x)}{x-a} = 0$ genau dann, wenn $(r_n)_{n\in\mathbb{N}_0}$ eine Null-Folge ist, und dies gilt genau dann, wenn $(|r_n|)_{n\in\mathbb{N}_0}$ eine Null-Folge ist, was wiederum äquivalent zu $\lim\limits_{x\to a} \frac{|r_c(x)|}{|x-a|} = 0$ ist. □

Es ist die dritte Seite der obigen Äquivalenzen, die es erlaubt, Differenzierbarkeit für Funktionen in mehreren Variablen zu definieren; wir ersetzen dazu die Absolutbeträge durch die entsprechenden Euklidischen Normen.

Definition 36.10. Seien $n, m \geqslant 1$. Sei $D \subseteq \mathbb{R}^n$ und $f\colon D \to \mathbb{R}^m$. Sei $a \in D$ und $D_a := D \setminus \{a\}$. Dann heißt f ***(total) differenzierbar*** in a, wenn es eine Matrix $\Delta \in \mathbb{R}^{m\times n}$ gibt mit:

$$\lim_{x\to a} \frac{\|r_\Delta(x)\|}{\|x-a\|} = 0 \quad \text{wobei} \quad r_\Delta(x) := f(x) - f(a) - \Delta\cdot(x-a) \in \mathbb{R}^m \qquad \text{für alle } x \in D_a.$$

Hier ist $\|r_\Delta(x)\|$ die übliche Euklidische Norm von $r_\Delta(x) \in \mathbb{R}^m$ und $\|x-a\|$ die übliche Euklidische Norm von $x - a \in \mathbb{R}^n$. (Beachte auch: Für das Produkt $\Delta\cdot(x-y)$ sind $x, y \in \mathbb{R}^n$ Spaltenvektoren.) Wir schreiben dann $Df(a) := \Delta$.

Satz 36.11. *Sei $D \subseteq \mathbb{R}^n$ und $f\colon D \to \mathbb{R}^m$, mit Komponenten $f_1, \ldots, f_m\colon D \to \mathbb{R}$. Sei f (total) differenzierbar in $a \in D$ und $\Delta \in \mathbb{R}^{m\times n}$ wie in Definition 36.10. Dann gilt:*

(a) *Die Funktion f ist stetig in a;*

(b) *Alle Komponenten f_i sind partiell differenzierbar in a;*

(c) *Die Matrix $\Delta \in \mathbb{R}^{m\times n}$ ist eindeutig bestimmt: Es gilt $\Delta = \mathcal{J}_f(a)$ (Jacobi–Matrix).*

Beweis. Siehe zum Beispiel §6, Satz 1, im 2. Band des Buches von Forster. □

Beispiel 36.12. Sei $n = 1$, $m \geqslant 1$. Sei $D \subseteq \mathbb{R}$ und $f\colon D \to \mathbb{R}^m$; seien $f_1, \ldots, f_m\colon D \to \mathbb{R}$ die Komponenten von f. Mit Lemma 36.2 folgt dann sofort:

(a) f ist stetig in $a \in D$, wenn jedes f_i stetig in a ist für $i = 1, \ldots, m$.

(b) f ist (total) differenzierbar in $a \in D$, wenn die übliche Ableitung $f_i'(a)$ existiert für $i = 1, \dots, m$. In diesem Fall ist $\mathcal{J}_f(a) = \begin{bmatrix} f_1'(a) \\ \vdots \\ f_m'(a) \end{bmatrix} \in \mathbb{R}^{m \times 1}$.

Satz 36.13. *Sei* $D \subseteq \mathbb{R}^n$ *und* $f \colon D \to \mathbb{R}$ *partiell differenziebar auf* D. *Sind alle partiellen Ableitungen* $D_i f$ *stetig in* $a \in D$, *so ist* f *(total) differenzierbar und stetig in* $a \in D$.

Beweis. Siehe zum Beispiel §6, Satz 2, im 2. Band des Buches von Forster. □

Sind wie im obigen Satz alle partiellen Ableitungen $D_i f \colon D \to \mathbb{R}$ stetig für alle $a \in D$, so bezeichnen wir f als ***stetig partiell differenziebar***. Wir halten also fest für $f \colon D \to \mathbb{R}$:

$$\boxed{\text{stetig partiell differenzierbar} \quad \Rightarrow \quad \text{(total) differenzierbar} \quad \Rightarrow \quad \text{partiell differenzierbar}}$$

aber die Umkehrungen gelten nicht, siehe Beispiel 36.8 und das folgende Beispiel:

Beispiel 36.14. Definiere $f \colon \mathbb{R}^2 \to \mathbb{R}$ durch

$$f(x,y) := \begin{cases} (x^2+y^2)\sin\left(\frac{1}{x^2+y^2}\right) & \text{falls } (x,y) \neq (0,0), \\ 0 & \text{falls } (x,y) = (0,0). \end{cases}$$

Auf $\mathbb{R}^2 \setminus \{0,0\}$ kann man alle partiellen Ableitungen mit den üblichen Formeln berechnen und sieht, dass diese stetig sind. Also ist f (total) differenzierbar für $(0,0) \neq (a,b) \in \mathbb{R}^2$. Betrachte nun $(a,b) = (0,0)$. Mit $\Delta := [0\ 0] \in \mathbb{R}^{1\times 2}$ folgt außerdem $|r_\Delta(x,y)| = |f(x,y) - f(0,0)| \leqslant x^2 + y^2 = \|(x,y)\|^2$ für $(x,y) \neq (0,0)$, und damit $0 \leqslant \frac{|r_\Delta(x,y)|}{\|(x,y)\|} \leqslant \|(x,y)\| \to 0$ für $(x,y) \to (0,0)$. Also ist f (total) differenzierbar in $(0,0)$ mit $\mathcal{J}_f(0,0) = [0\ 0] \in \mathbb{R}^{1\times 2}$; aber man überzeugt sich leicht, dass die partiellen Ableitungen nicht stetig in $(0,0)$ sind.

Bemerkung 36.15. Später werden wir auch Funktionen in einer reellen Variablen aber mit Werten in $\mathbb{C}$ betrachten. Sei also $D \subseteq \mathbb{R}$ und $f \colon D \to \mathbb{C}$. Wir schreiben $f(x) = f_1(x) + f_2(x)i$ wobei $f_1(x), f_2(x) \in \mathbb{R}$ für alle $x \in D$. Damit erhalten wir $\dot{f} \colon D \to \mathbb{R}^2$, $x \mapsto \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix}$. Wir bezeichnen f als differenzierbar in $a \in D$, wenn $\dot{f}$ im oben definierten Sinn differenzierbar in a ist, also f_1, f_2 im üblichen Sinne differenzierbar sind. In diesem Fall schreiben wir wieder

$$f'(a) := f_1'(a) + f_2'(a)i \in \mathbb{C}.$$

Ist f differenzierbar für alle $a \in D$, so erhalten wir die Ableitungsfunktion $f' \colon D \to \mathbb{C}$.
Es handelt sich also eigentlich um nichts Neues; aber dadurch, dass $\mathbb{C}$ auch selbst ein Körper ist, ergeben sich einige vereinfachte Formeln, wie im folgenden Beispiel.

Beispiel 36.16. In Kapitel VI, §30, haben wir $e^z = \exp(z) \in \mathbb{C}$ für alle $z \in \mathbb{C}$ definiert. Sei nun $\lambda \in \mathbb{C}$ fest und definiere die Funktion $f \colon \mathbb{R} \to \mathbb{C}$ durch $f(x) := \exp(\lambda x)$ für alle $x \in \mathbb{R}$. Wir behaupten, dass f auf ganz $\mathbb{R}$ differenzierbar ist mit

$$f'(x) = \lambda \exp(\lambda x) = \lambda f(x) \in \mathbb{C} \qquad \text{für alle } x \in \mathbb{R}.$$

Dazu: Schreibe $\lambda = \lambda_1 + \lambda_2 i$ mit $\lambda_1, \lambda_2 \in \mathbb{R}$. Mit der Funktionalgleichung für exp folgt $f(x) = \exp(\lambda_1 x + \lambda_2 x i) = \exp(\lambda_1 x) \cdot \exp(\lambda_2 x i)$. Kombiniert mit der Eulergleichung $\exp(xi) = \cos(x) + \sin(x)i$ ergibt sich $f(x) = f_1(x) + f_2(x)i$, wobei

$$f_1(x) = \exp(\lambda_1 x)\cos(\lambda_2 x) \qquad \text{und} \qquad f_2(x) = \exp(\lambda_1 x)\sin(\lambda_2 x).$$

Mit Produkt- und Kettenregel sind f_1, f_2 auf ganz $\mathbb{R}$ differenzierbar, wobei

$$f_1'(x) = \lambda_1 \exp(\lambda_1 x)\cos(\lambda_2 x) - \lambda_2 \exp(\lambda_1 x)\sin(\lambda_2 x),$$
$$f_2'(x) = \lambda_1 \exp(\lambda_1 x)\sin(\lambda_2 x) + \lambda_2 \exp(\lambda_1 x)\cos(\lambda_2 x).$$

Andererseits $\lambda f(x) = (\lambda_1 + \lambda_2 i)(f_1(x) + f_2(x)i) = \big(\lambda_1 f_1(x) - \lambda_2 f_2(x)\big) + \big(\lambda_1 f_2(x) + \lambda_2 f_1(x)\big)i$, und man rechnet sofort nach, dass dies gleich $f'(x) = f_1'(x) + f_2'(x)i$ ist. □

Satz 36.17 (***Kettenregel***; vgl. Satz 33.1)**.** *Gegeben seien* $f\colon D \to \mathbb{R}^m$ *und* $g\colon E \to \mathbb{R}^p$, *wobei* $D \subseteq \mathbb{R}^n$ *und* $f(D) \subseteq E \subseteq \mathbb{R}^m$*; wir können also* $g \circ f\colon D \to \mathbb{R}^p$ *bilden. Sei* $a \in D$. *Ist* f *(total) differenzierbar in* a *und* g *(total) differenzierbar in* $f(a) \in E$, *so ist* $g \circ f$ *(total) differenzierbar in* a *und es gilt* $\mathcal{J}_{g\circ f}(a) = \mathcal{J}_g(f(a)) \cdot \mathcal{J}_f(a) \in \mathbb{R}^{p\times m}$ *(Produkt der* $p \times m$*-Matrix* $\mathcal{J}_g(f(a))$ *und der* $m \times n$*-Matrix* $\mathcal{J}_f(a)$*).*

Beweis. Siehe §6, Satz 3, im 2. Band des Buches von Forster. □

Beispiel 36.18. Sei oben $n = p = 1$, also $f\colon D \to \mathbb{R}^m$ mit $D \subseteq \mathbb{R}$, und $g\colon E \to \mathbb{R}$ mit $f(D) \subseteq E \subseteq \mathbb{R}^m$. Seien $f_1, \dots, f_m\colon D \to \mathbb{R}$ die Komponenten von f. Ist f (total) differenzierbar in a, so ist $\mathcal{J}_f(a) \in \mathbb{R}^m$ ein Spaltenvektor mit Komponenten $f_1'(a), \dots, f_m'(a)$ (wobei f_i' die übliche Ableitung einer Funktion in einer reellen Variablen ist). Ist g (total) differenzierbar in $f(a) \in E$, so ist $\mathcal{J}_g(f(a)) = \nabla g(f(a)) \in \mathbb{R}^{1\times m}$ (Zeilenvektor). Dann ist $g \circ f\colon D \to \mathbb{R}$ und

$$(g \circ f)'(a) = \mathcal{J}_{g\circ f}(a) = \mathcal{J}_g(f(a)) \cdot \mathcal{J}_f(a) = \sum_{i=1}^{n} D_i g(f(a)) \cdot f_i'(a).$$

Sei zum Beispiel $f(t) := \begin{bmatrix} 1 - t^2 \\ t \end{bmatrix} \in \mathbb{R}^2$ für $t \in [0,1]$, und $g(x) = \|x\|$ für $x \in \mathbb{R}^2$. Dann ist $(g \circ f)(t) = \sqrt{(1-t^2)^2 + t^2} = \sqrt{t^4 - t^2 + 1}$ für $t \in [0,1]$. Man kann $(g \circ f)'(t)$ natürlich direkt ausrechnen. Oder man benutzt $f'(t) = \begin{bmatrix} -2t \\ 1 \end{bmatrix}$ und $D_i g(x) = \|x\|^{-1} x_i$ für $i = 1, 2$; siehe Beispiel 36.5(c). Obige Formel ergibt $(g \circ f)'(t) = \|f(t)\|^{-1}((1-t^2)(-2t) + t \cdot 1) = \|f(t)\|^{-1}(2t^3 - t)$.

Ab hier Woche 12

37. *Extrema, inkl. mit Nebenbedingungen*

Das folgende Beispiel ist ein sehr einfacher Modell-Fall für allgemeinere “Extremalprobleme”, die in den unterschiedlichsten Anwendungen vorkommen.

Beispiel 37.1. Wir wollen eine Box aus Pappmaterial (ohne Deckel) basteln mit Breite $a > 0$, Tiefe $b > 0$ und Höhe $c > 0$. Das Volumen dieser Box ist also $V = abc$. Wir

wollen die Box so basteln, dass das Volumen möglichst groß wird; uns steht allerdings nur eine bestimmte Fläche, sagen wir $N > 0$, von Pappmaterial zur Verfügung. Wie müssen wir a, b, c wählen, damit $V = abc$ maximal wird? Wir modellieren dies wie folgt:
Die Oberfläche der Box ist gegeben durch $ab + 2bc + 2ac$. (Denn es gibt den Boden der Fläche ab, zwei Seitenwände der Fläche bc und zwei Seitenwände der Fläche ac.) Die Formel für die Oberfläche definiert eine Funktion $g\colon \mathbb{R}^3_{>0} \to \mathbb{R}$ mit $g(x, y, z) := xy + 2yz + 2xz$ für alle $x, y, z > 0$; sei $D := \{(x, y, z) \in \mathbb{R}^3_{>0} \mid g(x, y, z) = N\} \subseteq \mathbb{R}^3$ und $V\colon D \to \mathbb{R}$ definiert durch $V(x, y, z) := xyz$. Dann suchen wir also ein Extremum dieser Funktion V.

Das obige Beispiel zeigt vor Allem, wie Funktionen in mehreren Variablen auf natürliche Weise vorkommen, wobei die Anzahl n der Variablen durch das gegebene Problem bestimmt wird, im obigen Fall also $n = 3$. Aber in anderen Situationen können genauso gut mehr Variablen eine Rolle spielen; neben den Ortskoordinaten x, y, z etwa die Zeit, Temperatur, Geschwindigkeit, elektrische Ladung, und so weiter. Wir werden also gleich allgemeine Funktionen $f\colon D \to \mathbb{R}$ betrachten, wobei D eine Teilmenge von $\mathbb{R}^n$ mit $n \geqslant 1$ ist.

Definition 37.2 (Vgl. Kap. VI, Def. 32.9)**.** Sei $D \subseteq \mathbb{R}^n$, $f\colon D \to \mathbb{R}$ und $a \in D$. Wir sagen, dass f in a ein ***lokales Maximum*** hat, wenn es ein $\varepsilon > 0$ gibt mit

$$U_\varepsilon(a) := \{x \in \mathbb{R}^n \mid \|x - a\| < \varepsilon\} \subseteq D \quad \text{und} \quad f(x) \leqslant f(a) \quad \text{für alle } x \in U_\varepsilon(a).$$

Analog sagen wir, dass f in a ein ***lokales Minimum*** hat, wenn es ein $\varepsilon > 0$ gibt mit $U_\varepsilon(a) \subseteq D$ und $f(x) \geqslant f(a)$ für alle $x \in U_\varepsilon(a)$.
Hat f in a ein lokales Maximum oder ein lokales Minimum, so sagen wir, dass f in a ein ***lokales Extremum*** besitzt.
[Beachte: Für $n = 1$ ist $U_\varepsilon(a)$ genau das offene Intervall $(a - \varepsilon, a + \varepsilon) \subseteq \mathbb{R}$ in Definition 32.9.]

Satz 37.3 (Vgl. Kap. VI, Satz 32.10)**.** *Sei* $D \subseteq \mathbb{R}^n$ *und* $a \in D$*. Die Funktion* $f\colon D \to \mathbb{R}$ *besitze in* a *ein lokales Extremum. Existieren alle partiellen Ableitungen von* f *in* a*, so gilt* $D_i f(a) = 0$ *für* $i = 1, \dots, n$*.*

Beweis. Nehmen wir an, es liegt ein lokales Maximum in a vor. (Der Beweis für ein lokales Minimum ist analog.) Für $1 \leqslant i \leqslant n$ betrachte die Funktion $f_{(i)}\colon D_{(i)} \to \mathbb{R}$. Dann hat diese in a_i ein lokales Maximum; also folgt $D_i f(a) = f'_{(i)}(a_i) = 0$ nach Kapitel VI, Satz 32.10. □

Wie zuvor heißt $a \in D$ ein ***kritischer Punkt*** von f, wenn $D_i f(a) = 0$ gilt für $i = 1, \dots, n$.

Um auch ein hinreichendes Kriterium für lokale Extrema zu finden, benötigen wir eine Version des Satzes von Taylor für mehrere Variablen. Wir formulieren nur einen Spezialfall dazu.

Definition 37.4. Sei $D \subseteq \mathbb{R}^n$ und $f\colon D \to \mathbb{R}$ partiell differenzierbar.
(a) Ist jede partielle Ableitung $D_i f\colon D \to \mathbb{R}$ wiederum partiell differenzierbar, so bezeichnen

wir f als ***zweimal partiell differenzierbar***; man kann also die partiellen Ableitungen 2. Ordnung $D_{ij}f(a) := D_i(D_jf)(a)$ für alle $a \in D$ und $1 \leqslant i, j \leqslant n$ bilden.
(b) Ist f zweimal partiell differenziebar und sind die partiellen Ableitungen 1. Ordnung und 2. Ordnung stetig, so bezeichnen wir f als ***zweimal stetig partiell differenzierbar***.

Satz 37.5 (***Taylor–Formel***; vgl. Satz 35.12)**.** *Sei* $D \subseteq \mathbb{R}^n$ *und* $f\colon D \to \mathbb{R}$ *zweimal stetig partiell differenzierbar. Sei* $a = (a_1, \ldots, a_n) \in D$*; es gebe ein* $\varepsilon > 0$ *mit* $U_\varepsilon(a) \subseteq D$*. Zu jedem* $x = (x_1, \ldots, x_n) \in U_\varepsilon(a)$ *gibt es dann ein* $c_x \in U_\varepsilon(a)$ *mit* $\|c_x - a\| \leqslant \|x - a\|$ *und*

$$f(x) = f(a) + \sum_{i=1}^{n} D_if(a) \cdot (x_i - a_i) + \frac{1}{2}\sum_{i,j=1}^{n} D_{ij}f(c_x) \cdot (x_i - a_i)(x_j - a_j).$$

Beweis (Skizze). Sei $x \in U_\varepsilon(a)$ fest. Betrachte die Funktion $g\colon [0,1] \to \mathbb{R}$ mit $g(t) := f(a+t(x-a))$. Beachte: Für $t \in [0,1]$ gilt $\|x-(a+t(x-a))\| = \|(x-a)(1-t)\| \leqslant \|x-a\|$, also $a+t(x-a) \in U_\varepsilon(a) \subseteq D$. Dann ist g zweimal stetig differenzierbar (als Funktion einer reellen Variablen). Wir bilden das Taylor–Polynom zu g vom Grad 1 und den zugehörigen Restterm (wie in Kapitel VI, §35):

$$T_{1,0}g(t) := g(0) + g'(0)t \qquad \text{und} \qquad R_{1,0}g(t) := g(t) - T_{1,0}g(t).$$

Nun ist $g(0) = f(a)$; außerdem berechnet man $g'(t) = \sum_{j=1}^{n} D_jf(a + t(x-a)) \cdot (x_j - a_j)$ und $g''(t) = \sum_{i,j=1}^{n} D_{ij}f(a + t(x-a)) \cdot (x_i - a_i)(x_j - a_j)$ für $t \in [0,1]$. (Dazu benötigt man die Kettenregel in Satz 36.17, angewandt wie in Beispiel 36.18.) Die Aussage folgt dann aus der Taylor–Formel in Kapitel VI, Satz 35.12, mit der Lagrange–Form des Restterms; für den Punkt c_x gilt sogar $c_x = a + t(x-a)$ mit $t \in [0,1]$.
Für weitere Details siehe §7, Sätze 1 und 2, im 2. Band des Buches von Forster. □

Für die oben auftretenden partiellen Ableitungen 2. Ordnung gilt:

Satz 37.6 (H. A. Schwarz)**.** *Sei* $D \subseteq \mathbb{R}^n$ *und* $f\colon D \to \mathbb{R}$ *zweimal stetig partiell differenzierbar. Dann gilt* $D_{ij}f(a) = D_{ji}f(a)$ *für alle* $a \in D$ *und* $1 \leqslant i, j \leqslant n$*. D.h., die wie folgt definierte* ***Hesse–Matrix*** *von* f *ist symmetrisch:*

$$H_f(a) := \begin{bmatrix} D_{11}f(a) & D_{12}f(a) & \ldots & D_{1n}f(a) \\ D_{21}f(a) & D_{22}f(a) & \ldots & D_{2n}f(a) \\ \vdots & \vdots & & \vdots \\ D_{n1}f(a) & D_{n2}f(a) & \ldots & D_{nn}f(a) \end{bmatrix} \in M_n(\mathbb{R}).$$

Beweis. Siehe zum Beispiel §5, Satz 1, im 2. Band des Buches von Forster. □

Definition 37.7. Sei $A = [a_{ij}]_{1 \leqslant i,j \leqslant n} \in M_n(\mathbb{R})$ eine symmetrische Matrix. Die zu A gehörige ***quadratische Form*** $Q_A\colon \mathbb{R}^n \to \mathbb{R}$ ist definiert durch

$$Q_A(x) = \sum_{i,j=1}^{n} a_{ij}x_ix_j = x^{\,\mathrm{tr}} \cdot A \cdot x \qquad \text{für alle } x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n.$$

Die Matrix A heißt ***positiv-definit***, wenn $Q_A(x) > 0$ für alle $0_n \neq x \in \mathbb{R}^n$ gilt.
Die Matrix A heißt ***negativ-definit***, wenn $Q_A(x) < 0$ für alle $0_n \neq x \in \mathbb{R}^n$ gilt.
Die Matrix A heißt ***indefinit***, wenn es $x, y \in \mathbb{R}^n$ gibt mit $Q_A(x) > 0$ und $Q_A(y) < 0$.

Bemerkung 37.8. Sei $A \in M_n(\mathbb{R})$ symmetrisch. Wie können wir entscheiden, ob A positiv-definit, negativ-definit oder indefinit ist? Dazu: Nach dem Spektralsatz gibt es eine orthogonale Matrix $T \in M_n(\mathbb{R})$, so dass $D := T^{tr} \cdot A \cdot T$ eine Diagonalmatrix ist. Seien $d_1, \ldots, d_n \in \mathbb{R}$ die Diagonaleinträge von D; dies sind genau die Eigenwerte von A. Ist nun $x \in \mathbb{R}^n$ (Spaltenvektor), so definiere $y := T^{tr} \cdot x = T^{-1} \cdot x$. Dann folgt $x = T \cdot y$ und damit

$$Q_A(x) = x^{tr} \cdot A \cdot x = (T \cdot y)^{tr} \cdot A \cdot (T \cdot y) = y^{tr} \cdot (T^{tr} \cdot A \cdot T) \cdot y = y^{tr} \cdot D \cdot y = \textstyle\sum_{i=1}^n d_i y_i^2.$$

$$\text{Wir sehen nun:} \qquad \begin{cases} \text{Gilt } d_i > 0 \text{ für alle } i \text{, so ist } A \text{ positiv-definit.} \\ \text{Gilt } d_i < 0 \text{ für alle } i \text{, so ist } A \text{ negativ-definit.} \\ \text{Gibt es } i, j \text{ mit } d_i > 0 \text{ und } d_j < 0 \text{, so ist } A \text{ indefinit.} \end{cases}$$

Kennt man also die Eigenwerte von A, so ist es einfach zu entscheiden, ob A positiv-definit, negativ-definit oder indefinit ist. Wenn die Eigenwerte nicht leicht zu bestimmen sind, so hilft oft das folgende ***Determinanten-Kriterium*** weiter. Setzen wir

$$A_k := \begin{bmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kk} \end{bmatrix} \in M_k(\mathbb{R}) \qquad \text{für } k = 1, 2, \ldots, n,$$

so heißen die Determinanten $\det(A_k)$ für $k = 1, \ldots, n$ die ***Hauptminoren*** von A. Dann gilt:

$$\boxed{A \text{ positiv-definit} \quad \Leftrightarrow \quad \text{alle Hauptminoren von } A \text{ sind positiv}}$$
$$\boxed{A \text{ negativ-definit} \quad \Leftrightarrow \quad -A \text{ positiv-definit}}$$

(Für einen Beweis siehe zum Beispiel Satz 8.3.20 im Buch von Huppert–Willems.)

Folgerung 37.9 (***Hinreichende Bedingung für lokale Extrema***)**.** *Sei $D \subseteq \mathbb{R}^n$ und die Funktion $f \colon D \to \mathbb{R}$ zweimal stetig partiell differenzierbar. Sei $a \in D$ ein kritischer Punkt von f, also $D_i f(a) = 0$ für $i = 1, \ldots, n$. Es gebe ein $\varepsilon > 0$ mit $U_\varepsilon(a) \subseteq D$. Dann gilt:*

(a) *Ist die Hesse–Matrix $H_f(a)$ positiv-definit, so besitzt f in a ein lokales Minimum.*
(b) *Ist die Hesse–Matrix $H_f(a)$ negativ-definit, so besitzt f in a ein lokales Maximum.*
(c) *Ist die Hesse–Matrix $H_f(a)$ indefinit, so liegt in a kein lokales Extremum vor.*

Beweis (Skizze). Nehmen wir an, wir sind im Fall (a), d.h., $H_f(a)$ ist positiv-definit. Dann überlegt man sich, dass es ein $0 < \varepsilon' \leqslant \varepsilon$ gibt, so dass $H_f(x)$ positiv-definit ist für alle $x \in U_{\varepsilon'}(a)$. Nach Voraussetzung ist $D_i f(a) = 0$ für alle i. Für $x \in U_{\varepsilon'}(a)$ hat also die Taylor–Formel in Satz 37.5 die folgende Form:

$$f(x) = f(a) + \frac{1}{2} \sum_{i,j=1}^{n} D_{ij} f(c_x) \cdot (x_i - a_i)(x_j - a_j) = f(a) + \frac{1}{2} Q_{H_f(c_x)}(x - a),$$

wobei $c_x \in D$ und $\|c_x - a\| \leqslant \|x - a\|$ gilt. Wegen $x \in U_{\varepsilon'}(a)$ ist $\|c_x - a\| \leqslant \|x - a\| < \varepsilon'$, also auch $c_x \in U_{\varepsilon'}(a)$. Ist also $x \neq a$, so folgt $Q_{H_f(c_x)}(x - a) > 0$ und $f(x) > f(a)$, d.h., es liegt ein lokales Minimum vor. Die Fälle (b) und (c) behandelt man analog.
(Für weitere Details siehe auch §7, Satz 4, im 2. Band des Buches von Forster.) $\square$

Beispiel 37.10. Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ definiert durch $f(x, y) := 4x^2 - 3xy$. Wir bestimmen die kritischen Punkte: Es gilt $D_1 f(a, b) = 8a - 3b$ und $D_2 f(a, b) = -3a$. Setzen wir diese Gleichungen gleich 0, so erhalten wir $a = b = 0$; also ist $(0, 0)$ der einzige kritische Punkt. Es gilt $H_f(0, 0) = \begin{bmatrix} 8 & -3 \\ -3 & 0 \end{bmatrix} \in M_2(\mathbb{R})$. Das charakteristische Polynom dieser Matrix ist $X^2 - 8X - 9 = (X + 1)(X - 9)$; also gibt es einen positiven und einen negativen Eigenwert. Die Matrix ist also indefinit; nach Folgerung 37.9(c) liegt kein lokales Extremum vor.

Beispiel 37.11. Betrachte die Funktion $V\colon D \to \mathbb{R}$ mit $V(x, y, z) = xyz$ in Beispiel 37.1. Man kann hier Folgerung 37.9 nicht direkt anwenden, denn man sieht leicht, dass es zu $a \in D$ kein $\varepsilon > 0$ gibt mit $U_\varepsilon(a) \subseteq D$. Aber wir können wie folgt argumentieren. Es ist $D = \{(x, y, z) \in \mathbb{R}^3_{>0} \mid xy + 2yz + 2xz = N\}$. Für $(x, y, z) \in D$ folgt also $2z(x + y) + xy = N$ und damit $z = \frac{N - xy}{2(x+y)}$. Setzen wir dies in $V(x, y, z) = xyz$ ein, so erhalten wir die Funktion

$$f\colon \mathbb{R}^2_{>0} \to \mathbb{R} \qquad \text{mit} \qquad f(x, y) = \frac{xy(N - xy)}{2(x + y)}.$$

Hier sind nun alle Voraussetzungen von Folgerung 37.9 erfüllt. Es gilt

$$\nabla f(x, y) = \Big(\frac{y^2(N - 2xy - x^2)}{2(x + y)^2}, \ \frac{x^2(N - 2xy - y^2)}{2(x + y)^2} \Big)$$

für alle $x > 0$, $y > 0$. Außerdem berechnet man

$$D_{11} f(x, y) = -\frac{y^2(N + y^2)}{(x + y)^3}, \qquad D_{22} f(x, y) = -\frac{x^2(N + x^2)}{(x + y)^3},$$
$$D_{12} f(x, y) = D_{21} f(x, y) = \frac{xy(N - x^2 - 3xy - y^2)}{(x + y)^3}.$$

Bestimmen wir die kritischen Punkte: Sei $\nabla f(a, b) = [0\ 0]$; also $N - 2ab - a^2 = 0$ und $N - 2ab - b^2 = 0$. Es folgt $a^2 = b^2$; wegen $a, b > 0$ also $a = b$. Schließlich ergibt sich $N - 3a^2 = 0$, also $a = b = \sqrt{N/3}$.
Also: Der einzige kritische Punkt ist $(a, b) = (\sqrt{N/3}, \sqrt{N/3})$. Setzen wir dies in die obigen Formeln für $D_{ij} f$ ein, so erhalten wir

$$H_f(a, b) = \begin{bmatrix} -\sqrt{N}/(2\sqrt{3}) & -\sqrt{N}/(4\sqrt{3}) \\ -\sqrt{N}/(4\sqrt{3}) & -\sqrt{N}/(2\sqrt{3}) \end{bmatrix} \quad \text{mit} \quad \det(H_f(a, b)) = N/16 > 0.$$

Mit dem Determinanten-Kriterium in Bemerkung 37.8 folgt, dass $-H_f(a, b)$ positiv-definit ist, also ist $H_f(a, b)$ negativ-definit und es liegt ein lokales Maximum vor.

Für die gesuchte Box maximalen Volumens muss man also folgende Maße nehmen:

$$\boxed{\text{Breite } a = \sqrt{N/3}, \quad \text{Tiefe } b = a = \sqrt{N/3}, \quad \text{Höhe } c = \tfrac{N - ab}{2(a+b)} = \ldots = \sqrt{N/12} = a/2}$$

Im obigen Beispiel war es relativ einfach, das ursprüngliche Problem so umzuformulieren, dass man danach Folgerung 37.9 direkt anwenden kann. Es gibt allerdings auch eine Methode, um derartige Umformulierungen zu vermeiden.

Sei $D \subseteq \mathbb{R}^n$ und eine stetig differenzierbare Funktion $f\colon D \to \mathbb{R}$ gegeben. Seien außerdem $g_1, \dots, g_r\colon D \to \mathbb{R}$ stetig differenzierbar. Gesucht sind dann die lokalen Extrema $a \in D$ von f "unter den Nebenbedingungen $g_1(a) = \dots = g_r(a) = 0$".
Etwas formaler: Sei $D' := \{x \in D \mid g_1(x) = \dots = g_r(x) = 0\}$. Wir betrachten dann die Einschränkung $f|_{D'}\colon D' \to \mathbb{R}$ und wollen herausfinden, ob $f|_{D'}$ ein ***relatives lokales Extremum*** besitzt. D.h.: Gibt es ein $a \in D'$ und $\varepsilon > 0$ mit $U_\varepsilon(a) \subseteq D$ und $f(x) \leqslant f(a)$ (bzw. $f(x) \geqslant f(a)$) für alle $x \in D' \cap U_\varepsilon(a)$? Sei dazu $g\colon D \to \mathbb{R}^r$ definiert durch $g(x) := (g_1(x), \dots, g_r(x))$ für alle $x \in D$. Sei $\mathcal{J}_g(x) \in \mathbb{R}^{r\times n}$ die Jacobi–Matrix von g in $x \in D$.

Satz 37.12 (***Methode der Lagrange–Multiplikatoren***)**.** *Seien* $f, g_1, \dots, g_r\colon D \to \mathbb{R}$ *und* g, D' *wie oben definiert. In* $a \in D'$ *liege ein relatives lokales Extremum von* $f|_{D'}$ *vor.*
Gilt $\mathrm{Rang}(\mathcal{J}_g(a)) = r$*, so gibt es Konstanten* $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ *mit* $\nabla f(a) = \sum_{j=1}^r \lambda_j \nabla g_j(a)$.

Beweis. Siehe §9, Satz 4, im 2. Band des Buches von Forster. ☐

Beispiel 37.13. Betrachte noch einmal die Funktion $V\colon \mathbb{R}^3_{>0} \to \mathbb{R}$ mit $V(x, y, z) = xyz$ in Beispiel 37.1. Wir haben hier eine Nebenbegingung, nämlich $g_1(x, y, z) = xy+2yz+2xz-N = 0$. Also bilde $D' := \{(x, y, z) \in \mathbb{R}^3_{>0} \mid g_1(x, y, z) = 0\}$. Hier ist

$$\nabla f(x, y, z) = (yz, xz, xy) \qquad \text{und} \qquad \nabla g_1(x, y, z) = (y + 2z, x + 2z, 2y + 2x).$$

Mit $\lambda_1 \in \mathbb{R}$ führt dies auf die Gleichungen:

$$yz = \lambda_1(y + 2z), \qquad xz = \lambda_1(x + 2z), \qquad xy = \lambda_1(2y + 2x).$$

Wir können diese wie folgt lösen: Bilde die Differenz der ersten und zweiten Gleichung: $(y - x)z = \lambda_1(y - x)$. Annahme: $x \neq y$; dann folgt $\lambda_1 = z > 0$, also $yz = z(y + 2z)$ und $y = y + 2z$, woraus sich $z = 0$ ergibt, Widerspruch. Also ist $x = y$. Aus $xy = \lambda_1(2y + 2x)$ und $x = y$ folgt dann $\lambda_1 = x/4$. Einsetzen in $xz = \lambda_1(x + 2z)$ und $yz = \lambda_1(y + 2z)$ ergibt $4z = x + 2z$ und $4z = y + 2z$, also schließlich $x = y = 2z$. Nun ist $0 = g_1(2z, 2z, z) = 4z^2 + 4z^2 + 4z^2 - N = 12z^2 - N$ und damit $z = \sqrt{N/12}$, wie zuvor.

Beispiel 37.14. Wir betrachten noch einmal die Funktion $f(x, y) = 4x^2 - 3xy$ wie in Beispiel 37.10, aber diesmal mit Definitionsbereich $D := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leqslant 1\}$ (Kreischeibe mit Radius 1). Um die lokalen Extrema von f auf D zu bestimmen, zerlegen wir D als $D = D^\circ \cup D'$, wobei $D^\circ := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ und $D' := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. Wie in Beispiel 37.10 ist $(0, 0)$ der einzige kritische Punkt von f auf D°, aber es liegt dort kein lokales Extremum vor. Um mögliche relative lokale Extrema von $f\colon D' \to \mathbb{R}$ zu finden, benutzen wir Lagrage–Multiplikatoren, mit $g_1 = x^2 + y^2 - 1$. Hier ist

$$\nabla f(x,y) = (8x - 3y, -3x) \qquad \text{und} \qquad \nabla g_1(x,y) = (2x, 2y).$$

Mit $\lambda_1 \in \mathbb{R}$ führt dies auf die beiden Gleichungen: $8x - 3y = 2\lambda_1 x$ und $-3x = 2\lambda_1 y$. Dies ergibt $x = -2\lambda_1 y/3$; wäre $y = 0$, so auch $x = 0$, Widerspruch. Also ist $y \neq 0$. Nun ist $-16\lambda_1 y/3 - 3y = -4\lambda_1^2 y/3$ also $(4\lambda_1^2 - 16\lambda_1 - 9)y = 0$. Wegen $y \neq 0$ folgt $4\lambda_1^2 - 16\lambda_1 - 9 = 0$, mit den beiden Lösungen $\lambda_1 = -1/2$ oder $\lambda_1 = 9/2$. Dann folgt $x = y/3$ oder $x = -3y$. Einsetzen in $g_1(x,y) = 0$ ergibt mögliche relative lokale Extrema an den vier Punkten

$$(-1/\sqrt{10}, -3/\sqrt{10}), \quad (1/\sqrt{10}, 3/\sqrt{10}), \quad (3/\sqrt{10}, -1/\sqrt{10}), \quad (-3/\sqrt{10}, 1/\sqrt{10}).$$

Schließlich muss man sich noch genauer anschauen, ob hier tatsächlich relative lokale Extrema vorliegen; man kann zeigen, dass an den ersten beiden Punkten relative lokale Minima vorliegen, und an den letzten beiden Punkten relative lokale Maxima.

Beispiel 37.15. Sei $A \in M_n(\mathbb{R})$ eine symmetrische Matrix. In Bemerkung 24.1 haben wir gesehen, dass A einen reellen Eigenwert besitzt (indem wir den Fundamentalsatz der Algebra angewandt haben). Mit Hilfe der Lagrange–Multiplikatoren kann man auch einen anderen Beweis geben, direkt in $\mathbb{R}$; siehe §9, Beispiel (9.2), im 2. Band des Buches von Forster, oder

`https://math.stackexchange.com/questions/814467/`.

38. *Differentialgleichungen: Elementare Methoden*

Unter einer ***Differentialgleichung*** (kurz DGL) versteht man eine Gleichung, in der eine gesuchte Funktion $y = f(x)$ zusammen mit ihren Ableitungen (auch höheren Ableitungen) vorkommt. Als "Ordnung" der DGL wird dabei die höchste auftretende Ableitung bezeichnet. Ist $y = f(x)$ eine reelle Funktion einer Variablen x, so spricht man von einer ***gewöhnlichen Differentialgleichung***. Einfache Beispiele sind:

$$y' = x^3 \cos(2x) \quad \text{(1. Ordnung)} \qquad \text{oder} \qquad y''' + 7y^2 = (x^2 - 5)y' \quad \text{(3. Ordnung)}.$$

Man sucht dann alle Funktionen $y = f(x)$, welche die DGL erfüllen. Im ersten Beispiel handelt es sich einfach um das Problem, eine Stammfunktion von $g(x) := x^3 \cos(2x)$ zu finden (wie in §33); im zweiten Beispiel suchen wir alle dreimal differenzierbaren Funktionen $f(x)$, so dass $f'''(x) + 7f(x)^2 = (x^2 - 5)f'(x)$ gilt, für alle $x \in \mathbb{R}$ im Definitionsbereich von f. Die Entwicklung von Lösungsmethoden ist eines der zentralen Gebiete der gesamten Analysis. In diesem und dem nächsten Abschnitt geben wir eine erste Einführung dazu, wobei wir uns hauptsächlich auf die Behandlung von vielen konkreten Beispielen konzentrieren.

Differentialgleichungen kommen in den unterschiedlichsten Anwendungen vor, in Physik, Biologie, Informatik, Ökonomie, ... Zum Beispiel geht es oft darum, die Entwicklung einer gegebenen Größe $y = y(t)$ im Laufe der Zeit t zu modellieren. (Denken Sie etwa an Bakterienkulturen, Kometenbahnen etc.) Typischerweise fließen dabei bestimmte Regeln ein, die für die Änderungsrate der Größe von einem Zeitpunkt zum nächsten Zeitpunkt gelten (das kann natürlich je nach Anwendungsfall unterschiedlich sein); rücken die beiden Zeitpunkte immer näher zusammen, so wird die Änderungsrate im Grenzfall durch die Ableitung $y'(t)$ beschrieben: Auf diese Weise erhält man also eine DGL für $y(t)$.

Beispiel 38.1 (***Exponentielles Wachstum***)**.** Gegeben sei eine Größe $y = y(t)$ als Funktion der Zeit t. Es gelte die Regel: *In jedem Zeitintervall mit einer festen Länge* $\delta > 0$ *ändern sich die Werte von* $y(t)$ *stets um den gleichen Faktor*; in Formeln: Es gibt ein $\rho \in \mathbb{R}_{>0}$ mit $y(t+\delta) = \rho y(t)$ für alle $t \geqslant 0$. (Ein solches Verhalten ist zum Besipiel aus Corona-Zeiten Allgemeinwissen geworden durch Aussagen wie: "In einer Woche verdoppelt sich die Anzahl der Infizierten"; also hier $\delta =$ "1 Woche" und $\rho = 2$.) Es folgt dann $y(t+2\delta) = y((t+\delta)+\delta) = \rho y(t+\delta) = \rho^2 y(t)$ und analog $y(t+n\delta) = \rho^n y(t)$ für alle $n \in \mathbb{N}$; daher die Bezeichnung "exponentielles Wachstum". Geht man zu Bruchteilen δ/n des Zeitintervalls über, so ergibt sich entsprechend $y(t+\delta/n) = \rho^{1/n} y(t)$ für alle $t \geqslant 0$. Damit erhalten wir einen Ansatz, um die Ableitung von $y(t)$ ins Spiel zu bringen. Denn es folgt

$$\frac{y(t+\delta/n) - y(t)}{\delta/n} = \frac{\rho^{1/n} y(t) - y(t)}{\delta/n} = \frac{y(t)}{\delta} n\big(\rho^{1/n} - 1\big).$$

Für $n \to \infty$ geht die linke Seite gegen $y'(t)$; auf der rechten Seite geht $n(\rho^{1/n} - 1)$ gegen $\lim\limits_{x\to 0}(\rho^x - 1)/x = \log(\rho)$. (Dies folgt leicht mit Beispiel 32.2(b) und den Formeln in Definition 33.5.) Wir erhalten also die DGL $\boxed{y'(t) = c y(t)}$ wobei $c := \log(\rho)/\delta \in \mathbb{R}$ fest ist.

Ab hier Woche 13

Wir kennen bereits eine Lösung der obigen DGL 1. Ordnung $y' = cy$ (wobei $c \in \mathbb{R}$ fest ist). Sei nämlich $f(x) := \exp(cx)$ für $x \in \mathbb{R}$; dann gilt $f'(x) = c\exp(cx) = cf(x)$. Typische Fragestellungen: Gibt es noch weitere Lösungen? Wenn ja, durch welche Bedingungen kann man Eindeutigkeit der Lösung erreichen?

Lemma 38.2. *Sei* $c \in \mathbb{R}$ *und* $f\colon \mathbb{R} \to \mathbb{R}$ *differenzierbar mit* $f'(x) = cf(x)$ *für alle* $x \in \mathbb{R}$*. Dann gibt es eine Konstante* $a \in \mathbb{R}$ *mit* $f(x) = a\exp(cx)$ *für alle* $x \in \mathbb{R}$*. Geben wir uns den Wert* $f(0) \in \mathbb{R}$ *vor, so ist* $a = f(0)$*, also die Lösung eindeutig bestimmt.*

Beweis. Wir definieren $F\colon \mathbb{R} \to \mathbb{R}$ durch $F(x) := \exp(-cx)f(x)$ für alle $x \in \mathbb{R}$. Mit der Produktregel folgt, dass F differenzierbar ist; für alle $x \in \mathbb{R}$ gilt

$$F'(x) = -c\exp(-cx)f(x) + \exp(-cx)f'(x) = -c\exp(-cx)f(x) + c\exp(-cx)f(x) = 0.$$

Nach Folgerung 32.12 gibt es also ein $a \in \mathbb{R}$ mit $F(x) = a$ für alle $x \in \mathbb{R}$. Also gilt $f(x) = F(x)\exp(cx) = a\exp(cx)$ für alle $x \in \mathbb{R}$. Wegen $\exp(0) = 1$ muss hier $f(0) = a$ gelten. □

Die obige Aussage zeigt, dass die Funktion exp als Lösung der sehr einfachen DGL $y' = y$ charakterisiert werden kann. (Hätten wir exp noch nicht definiert, so wäre man nun auf die Idee gekommen, diese Funktion zu konstruieren, um die DGL $y' = y$ lösen zu können.)

Eine analoge Charakterisierung gibt es auch für sin und cos. Sei dazu wieder $c \in \mathbb{R}$ und $f(x) := \sin(cx)$ für $x \in \mathbb{R}$. Dann gilt $f'(x) = c\cos(cx)$ und $f''(x) = -c^2\sin(cx) = -c^2 f(x)$, d.h., die Funktion $f(x) = \sin(cx)$ ist eine Lösung der folgenden DGL 2. Ordnung: $y'' = -c^2 y$. Analog sieht man, dass auch die Funktion $g(x) = \cos(cx)$ eine Lösung dieser DGL ist.

Lemma 38.3. *Sei $0 \neq c \in \mathbb{R}$ und $f\colon \mathbb{R} \to \mathbb{R}$ zweimal differenzierbar mit $f''(x) = -c^2 f(x)$ für alle $x \in \mathbb{R}$. Dann gibt es $a, b \in \mathbb{R}$ mit $f(x) = a\sin(cx) + b\cos(cx)$ für alle $x \in \mathbb{R}$. Geben wir uns die Werte $f(0)$, $f'(0)$ vor, so ist $a = f'(0)/c$ und $b = f(0)$, also die Lösung eindeutig.*

Beweis. Die Funktionen $\cos(cx)$ und $\sin(cx)$ sind Lösung der DGL, also auch die Funktion $a\sin(cx) + b\cos(cx)$ für beliebige $a, b \in \mathbb{R}$, und schließlich die Funktion $h(x) := f(x) - (a\sin(cx) + b\cos(cx))$. Setzen wir $b := f(0)$ und $a := f'(0)/c$, so gilt $h(0) = f(0) - b = 0$; wegen $h'(x) = f'(x) - (ac\cos(cx) - bc\sin(cx))$ folgt auch $h'(0) = f'(0) - ac = 0$. Jetzt kommt der Trick des Beweises: Definiere die Funktion $H\colon \mathbb{R} \to \mathbb{R}$ durch $H(x) := c^2 h(x)^2 + h'(x)^2$ für alle $x \in \mathbb{R}$. Nach der Kettenregel gilt

$$H'(x) = 2c^2 h'(x)h(x) + 2h''(x)h'(x) = 2c^2 h'(x)h(x) - 2c^2 h(x)h'(x) = 0,$$

also ist H konstant. Wegen $H(0) = c^2 h(0)^2 + h'(0)^2 = 0$ folgt $H(x) = 0$ und damit auch $h(x) = h'(x) = 0$ für alle $x \in \mathbb{R}$. Insbesondere gilt $f(x) = a\sin(cx) + b\cos(cx)$. □

Bemerkung 38.4. In den obigen Beispielen sehen wir, dass es im Allgemeinen mehrere Lösungen einer DGL geben kann, aber die Lösung $y = f(x)$ eindeutig wird, sobald man die Werte $f(x_0), f'(x_0), \ldots, f^{(d-1)}(x_0)$ an einem bestimmten Punkt x_0 vorgibt (in den obigen Beispielen stets $x_0 = 0$), wobei $d \geqslant 1$ die Ordnung der DGL ist. Das Problem, die Lösungen einer DGL unter Vorgabe solcher "Anfangsbedingungen" zu bestimmen, bezeichnet man dann auch als ***Anfangswertproblem***. — Idealerweise sollte ein solches Anfangswertproblem eine eindeutige Lösung besitzen, aber dies ist nicht immer der Fall, wie folgendes Beispiel zeigt.

Beispiel 38.5. Gegeben sei die DGL $y' = \sqrt[3]{y^2}$. Die Funktion $f_0(x) := 0$ (für $x \in \mathbb{R}$) ist sicherlich eine Lösung. Sei nun $a \in \mathbb{R}$ beliebig und $f_a(x) := (x-a)^3/27$ für $x \in \mathbb{R}$. Dann gilt $f_a'(x) = (x-a)^2/9 = \sqrt[3]{f_a(x)^2}$ für alle $x \in \mathbb{R}$; also ist auch $y = f_a(x)$ eine Lösung. Für $a \neq 0$ sind dann aber $f_0(x)$ und $f_a(x)$ zwei verschiedene Lösungen des Anfangswertproblems mit $f_0(a) = f_a(a) = 0$.

Wir werden nun eine Reihe von Beispielen von DGLen behandeln, in denen man mit elementaren Methoden eindeutige Lösungen für die jeweiligen Anfangswertprobleme erhält.

Erstes weiteres Beispiel: Betrachte eine DGL der Form $y' = g(x)y + h(x)$, wobei die Funktionen $g, h\colon D \to \mathbb{R}$ vorgegeben sind. Eine solche DGL heißt ***lineare Differentialgleichung*** 1. Ordnung; und zwar homogen falls $h(x) = 0$ für alle $x \in D$ gilt, und inhomogen sonst. Der folgende Satz ist eine Verallgemeinerung von Lemma 38.2.

Satz 38.6 (Lineare DGL 1. Ordnung)**.** *Gegeben seien Funktionen* $g, h\colon D \to \mathbb{R}$*. Es gebe eine Stammfunktion* $G\colon D \to \mathbb{R}$ *von* g*, d.h.,* G *ist differenzierbar mit* $G'(x) = g(x)$*. Setze damit* $u(x) := \exp(-G(x))h(x)$ *für* $x \in D$*. Dann ist die allgemeine Lösung der linearen* DGL $y' = g(x)y + h(x)$ *gegeben durch* $y = f(x) = \exp(G(x))U(x)$*, wobei* $U\colon D \to \mathbb{R}$ *eine Stammfunktion von* u *ist. Insbesondere: Im homogenen Fall ist* $u(x) = h(x) = 0$ *und die allgemeine Lösung gegeben durch* $y = f(x) = C\exp(G(x))$ *mit einer Konstanten* $C \in \mathbb{R}$*.*

Beweis. Sei zuerst $U\colon D \to \mathbb{R}$ eine Stammfunktion von u; dann ist $U'(x) = u(x) = \exp(-G(x))h(x)$. Für $f(x) := \exp(G(x))U(x)$ folgt mit der Produkt- und Kettenregel:

$$\begin{aligned} f'(x) &= G'(x)\exp(G(x))U(x) + \exp(G(x))U'(x) = \exp(G(x))\big(g(x)U(x) + U'(x)\big) \\ &= g(x)\exp(G(x))U(x) + \exp(G(x))\exp(-G(x))h(x) = g(x)f(x) + h(x). \end{aligned}$$

Also ist $y = f(x)$ eine Lösung der DGL $y' = g(x)y + h(x)$. Sei umgekehrt $y = f(x)$ eine beliebige Lösung. Setze dann $U(x) := \exp(-G(x))f(x)$ für $x \in D$. Analog wie oben folgt

$$\begin{aligned} U'(x) &= -G'(x)\exp(-G(x))f(x) + \exp(-G(x))f'(x) \\ &= \exp(-G(x))\big(-g(x)f(x) + f'(x)\big) = \exp(-G(x))h(x) = u(x) \end{aligned}$$

für alle $x \in D$; also ist U eine Stammfunktion von u und es gilt $f(x) = \exp(G(x))U(x)$. $\square$

Gegeben sei $y' = g(x)y + h(x)$ mit einer Lösung $y = f(x) = \exp(G(x))U(x)$ wie oben. Sei $x_0 \in D$ und $c := f(x_0)$ vorgegeben. Dann ist $c = f(x_0) = \exp(G(x_0))U(x_0)$, also $U(x_0) = c\exp(-G(x_0))$. Nun ist U eine Stammfunktion von u, also selbst bis auf die Addition einer Konstanten $C' \in \mathbb{R}$ eindeutig bestimmt. Da $U(x_0) = c\exp(-G(x_0))$ vorgegeben ist, ist daher auch U eindeutig bestimmt. Also hat hier jedes Anfangswertproblem eine eindeutige Lösung.

Beispiel 38.7. Gegeben sei die lineare DGL $y' = x^{-1}y + x$, mit $g(x) = x^{-1}$ und $h(x) = x$ für $x \in \mathbb{R}_{>0}$. Hier ist $G(x) = \log(x)$ eine Stammfunktion von g. Nun gilt $u(x) = \exp(-\log(x)h(x) = x^{-1}x = 1$ für $x \in \mathbb{R}_{>0}$. Eine beliebige Stammfunktion von u ist also gegeben durch $U(x) = x + C$ mit einer Konstanten C. Die allgemeine Lösung der DGL ist folglich gegeben durch $f(x) = \exp(G(x))U(x) = x(x + C)$ für $x \in \mathbb{R}_{>0}$. Geben wir uns den Anfangswert $t := f(1) \in \mathbb{R}$ vor, so ist $t = f(1) = 1 + C$, also $C = t - 1$. Die eindeutige Lösung des Anfangswertproblems ist also $y = f(x) = x(x + t - 1) = x^2 + (t - 1)x$.

Beispiel 38.8. Betrachte die lineare DGL $y' = 2xy + x^3$, mit $g(x) = 2x$ und $h(x) = x^3$ für $x \in \mathbb{R}$. Hier ist $G(x) = x^2$ eine Stammfunktion von g. Wir benötigen nun eine Stammfunktion von $u(x) = \exp(-x^2)h(x) = \exp(-x^2)x^3$ für $x \in \mathbb{R}$. Zum Beispiel mit Sage erhalten wir:

```
sage: integral(exp(-x^2)*x^3,x)
-1/2*(x^2 + 1)*e^(-x^2)
```

(Oder benutze die Substitution $t = x^2$ und partielle Integration; Beweis selbst oder Übung.) Also gibt es eine Konstante $C \in \mathbb{R}$ mit $U(x) = -\frac{1}{2}(x^2 + 1)\exp(-x^2) + C$. Die allgemeine

Lösung der DGL ist also gegeben durch:

$$y = f(x) = \exp(G(x))U(x) = \exp(x^2)\bigl(-\tfrac{1}{2}(x^2+1)\exp(-x^2) + C\bigr) = C\exp(x^2) - \tfrac{1}{2}(x^2+1).$$

Geben wir zum Beispiel den Wert $3 = f(0) \in \mathbb{R}$ vor, so ist $3 = f(0) = C\exp(0) - \frac{1}{2}(0+1) = C - \frac{1}{2}$, also $C = 3 + \frac{1}{2} = \frac{7}{2}$ eindeutig bestimmt.

Es folgen weitere Beispiele von gewöhnlichen DGLen, für die man Lösungen finden kann.

Beispiel 38.9 (Trennung der Variablen)**.** Wir bezeichnen eine DGL 1. Ordnung als "***separierbar***", wenn sie von der folgenden Form ist: $y' = g(x)h(y)$, wobei $g\colon D \to \mathbb{R}$ und $h\colon D' \to \mathbb{R}$ vorgegebene Funktionen sind mit $D, D' \subseteq \mathbb{R}$. Gesucht sind hier also alle differenzierbaren Funktionen $f\colon D \to \mathbb{R}$ mit $f(D) \subseteq D'$ und $f'(x) = g(x)h\bigl(f(x)\bigr)$ für $x \in D$.

<u>1. Fall:</u> Es gebe ein $y_0 \in D'$ mit $h(y_0) = 0$. Dann ist die konstante Funktion $f(x) := y_0$ (für $x \in D$) offensichtlich eine Lösung (denn $f'(x) = 0$ für alle $x \in D$, und $h(f(x)) = h(y_0) = 0$).

<u>2. Fall:</u> Sei $f\colon D \to \mathbb{R}$ eine differenzierbare Funktion mit $f(D) \subseteq D'$ und $h(f(x)) \neq 0$ für alle $x \in D$. Ist f eine Lösung der DGL, so erhalten wir nach dem folgenden Verfahren Bedingungen an f, die es uns unter Umständen erlauben, f explizit zu bestimmen. Sei dazu $E := f(D) = \{f(x) \mid x \in D\}$. Wir nehmen nun an:

- Die Funktion g sei auf D definiert und besitze eine Stammfunktion $G\colon D \to \mathbb{R}$;
- die Funktion $\tilde{h}\colon E \to \mathbb{R}$ mit $\tilde{h}(x) := 1/h(x)$ besitze eine Stammfunktion $\tilde{H}\colon E \to \mathbb{R}$.

D.h., G und $\tilde{H}$ sind differenzierbar mit $G'(x) = g(x)$ für $x \in D$, und $\tilde{H}'(x) = 1/h(x)$ für $x \in E$. Mit der Kettenregel folgt dann $(\tilde{H} \circ f)'(x) = f'(x)\tilde{H}'(f(x)) = f'(x)/h(f(x)) = g(x)$ für alle $x \in D$, d.h., $\tilde{H} \circ f$ ist auch eine Stammfunktion von g. Also gibt es eine Konstante $C \in \mathbb{R}$ mit $(\tilde{H} \circ f)(x) = G(x) + C$, d.h., $\tilde{H}\bigl(f(x)\bigr) = G(x) + C$ für $x \in D$. Wenn $\tilde{H}$ nicht zu kompliziert ist, so kann man versuchen, diese Gleichung nach f "aufzulösen".

Beispiel 38.10. Gegeben sei die DGL $y' = 2xy^2$. Diese ist separierbar mit $g(x) = 2x$ und $h(x) = x^2$; beide Funktionen sind zunächst auf ganz $\mathbb{R}$ definiert. (Die genaue Form von D, D' wie oben stellt sich erst nach und nach heraus.) Wir unterscheiden 2 Fälle wie oben:

<u>1. Fall:</u> Für $y_0 := 0$ ist $h(y_0) = 0$, also ist $f(x) := 0$ eine Lösung.

<u>2. Fall:</u> Sei $D' := \mathbb{R} \setminus \{0\}$; dann ist $h(x) \neq 0$ für alle $x \in D'$. Nun ist $G(x) := x^2$ eine Stammfunktion von g, und $\tilde{H}(x) := -1/x$ eine Stammfunktion von $\tilde{h} = 1/h(x) = 1/x^2$. Es muss also gelten $-1/f(x) = \tilde{H}(f(x)) = G(x) + C = x^2 + C$ mit einer Konstanten $C \in \mathbb{R}$. Damit haben wir die folgenden Kandidaten für f gefunden:

$$f(x) = \frac{-1}{x^2 + C} \qquad \text{für } x \in \mathbb{R} \text{ mit } x^2 \neq -C;$$

Der Definitionsbereich für f ist $D := \{x \in \mathbb{R} \mid x^2 \neq -C\}$. Eine einfache Probe ergibt, dass $y = f(x)$ auch tatsächlich eine Lösung der DGL $y' = 2xy^2$ ist. Geben wir einen Wert $f(x_0)$ vor (mit $x_0 \in D$), so wird dadurch die Konstante C eindeutig bestimmt, d.h., das

Anfangswertproblem mit Vorgabe von $f(x_0)$ hat eine eindeutige Lösung.
Suchen wir also zum Beispiel die Lösung $y = f(x)$ mit $f(0) = 1$, so muss $1 = f(0) = -1/C$ gelten, d.h., $C = -1$. Also ist $f(x) = 1/(1-x^2)$ die eindeutige Lösung der DGL mit $f(0) = 1$.

Beispiel 38.11 (***Logistische DGL***)**.** Gegeben sei eine DGL der Form $y' = c{\cdot}y{\cdot}(1 - \frac{y}{M})$ mit vorgegebenen $c, M \in \mathbb{R}_{>0}$. Diese modelliert sogenanntes "logistisches Wachstum", welches eine Erweiterung des Modells des exponentiellen Wachstums ist. Man interpretiert wieder $y = y(t)$ als eine Funktion der Zeit t, wobei wie zuvor c eine konstante Wachstumsrate ist, und zusätzlich M eine obere Schranke für die möglichen Werte für $y(t)$. (Ist zum Beispiel $y(t)$ die Anzahl der Corona-Infektionen, so ist $y(t)$ durch die Gesamtbevölkerungszahl nach oben beschränkt, und dies muss in die Modellrechnungen eingebaut werden; siehe auch https://de.wikipedia.org/wiki/Logistische_Funktion.) Beachte: Für $M \to +\infty$ geht obige DGL über in die vorherige DGL $y' = c{\cdot}y$, die reines exponentielles Wachstum beschreibt.
Zur Lösung der DGL: Diese ist separierbar mit $g(t) = \frac{c}{M}$ und $h(t) = t(M-t)$ für $t \in \mathbb{R}$. Sei nun $D' := \mathbb{R}_{>0}$. Es ist $G(t) := \frac{c}{M}t$ eine Stammfunktion von g. Es ist $\tilde{h}(t) = 1/h(t) = \frac{1}{t(M-t)} = \frac{1}{M}\big(\frac{1}{t} + \frac{1}{M-t}\big)$; also ist eine Stammfunktion gegeben durch $\tilde{H}(t) = \frac{1}{M}\big(\log(t) - \log(M-t)\big) = \frac{1}{M}\log\big(\frac{t}{M-t}\big)$. Folglich gibt es eine Konstante $C \in \mathbb{R}$ mit

$$\frac{1}{M}\log\Big(\frac{y(t)}{M-y(t)}\Big) = \tilde{H}(y(t)) = G(t) + C = \frac{c}{M}t + C.$$

Damit ergibt sich $\log\big(\frac{y(t)}{M-y(t)}\big) = ct + CM$ und $\frac{y(t)}{M-y(t)} = \exp(CM)\exp(ct)$. Wir bilden den Kehrwert und erhalten $\frac{M}{y(t)} - 1 = \exp(-CM)\exp(-ct)$; also schließlich

$$y(t) = \frac{M}{1 + \exp(-CM)\exp(-ct)} \qquad \text{für } t \geqslant 0.$$

(Eine Probe zeigt, dass dies auch tatsächlich eine Lösung der DGL ist.) Sei nun der Startwert $M_0 := y(0) > 0$ vorgegeben (also inbesondere $M_0 \leqslant M$). Dann gilt $M_0 = y(0) = \frac{M}{1+\exp(-CM)}$. Umformen ergibt $\exp(-CM) = \frac{M}{M_0} - 1$. Also ist die Lösung gegeben durch

$$\boxed{y(t) = \frac{M}{1 + \big(\frac{M}{M_0} - 1\big)\exp(-ct)} \qquad \text{für alle } t \geqslant 0}$$

Beachte: Wegen $M_0 \leqslant M$ und $\exp(-ct) > 0$ ist der Nenner im obigen Ausdruck $\geqslant 1$, also $y(t) \leqslant M$; außerdem ist $y(t) \to M$ für $t \to \infty$. Langfristig nähert sich also $y(t)$ der oberen Schranke M an, aber das Wachstum wird immer langsamer je größer $y(t)$ bereits ist.

Beispiel 38.12. Gegeben seien $p, q \in \mathbb{R}$. Wir betrachten dann die ***homogene lineare DGL 2. Ordnung*** der Form: $y'' + py' + qy = 0$. Um die Lösungen zu bestimmen, bilden wir das Polynom $F := X^2 + pX + q \in \mathbb{R}[X]$ und unterscheiden die 3 möglichen Fälle:

1. Fall: $F = (X - \lambda_1) * (X - \lambda_2)$ mit $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 \neq \lambda_2$.
2. Fall: $F = (X - z) * (X - \overline{z})$ mit $z \in \mathbb{C} \setminus \mathbb{R}$.

3. Fall: $F = (X - \lambda)^2$ mit $\lambda \in \mathbb{R}$.

Beachte nun zunächst: Sind $y_i = f_i(x)$ Lösungen der DGL für $i = 1, 2$, so ist auch $y = f(x) = af_1(x) + bf_2(x)$ für beliebige $a, b \in \mathbb{R}$ eine Lösung, wie man sofort nachrechnet. Wir beschreiben nun in jedem der obigen Fälle jeweils zwei spezielle Lösungsfunktionen f_1, f_2 und werden anschließend zeigen, dass sich *alle* Lösungen der DGL aus diesen kombinieren lassen.

Zum 1. Fall: Für $i = 1, 2$ setze $f_i(x) := \exp(\lambda_i x)$ für alle $x \in \mathbb{R}$. Dann ist $f_i'(x) = \lambda_i f_i(x)$ und $f_i''(x) = \lambda_i^2 f_i(x)$; also $f_i''(x) + pf_i(x) + qf_i(x) = (\lambda_i^2 + p\lambda_i + q)f_i(x) = F(\lambda_i)f_i(x) = 0$ für alle $x \in \mathbb{R}$; also sind $f_1(x)$ und $f_2(x)$ Lösungen. Wir erhalten also Lösungen der Form:

$$\boxed{f(x) = C_1 \exp(\lambda_1 x) + C_2 \exp(\lambda_2 x) \text{ für alle } x \in \mathbb{R}, \text{ mit Konstanten } C_1, C_2 \in \mathbb{R}}$$

Zum 2. Fall: Mit den Regeln in Beispiel 36.16 sieht man wieder, dass $f(x) := \exp(zx)$ und $g(x) := \exp(\bar{z}x)$ für $x \in \mathbb{R}$ Lösungen sind — aber f, g haben Werte in $\mathbb{C}$! Schreibe $z = \alpha + \beta i$ mit $\alpha, \beta \in \mathbb{R}$. Dann ist $f(x) = f_1(x) + f_2(x)i$ und $g(x) = f_1(x) - f_2(x)i$ mit $f_1(x) := \exp(\alpha x)\cos(\beta x)$ und $f_2(x) := \exp(\alpha x)\sin(\beta x)$ (siehe noch einmal Beispiel 36.16). Wegen $f_1(x) = \frac{1}{2}(f(x) + g(x))$ und $f_2(x) = \frac{1}{2i}(f(x) - g(x))$ erhalten wir also Lösungen der Form:

$$\boxed{f(x) = C_1 \exp(\alpha x)\cos(\beta x) + C_2 \exp(\alpha x)\sin(\beta x) \text{ für alle } x \in \mathbb{R}, \text{ mit } C_1, C_2 \in \mathbb{R}}$$

(Dies kann man nun auch direkt nachrechnen, ohne zuerst über die komplexen Zahlen zu gehen; aber ohne den Umweg über $\mathbb{C}$ wäre man vermutlich nicht so schnell auf die obigen reellen Lösungen gekommen.)

Zum 3. Fall: Hier ist $f_1(x) := \exp(\lambda x)$ wie im 1. Fall eine Lösung. Setze nun $f_2(x) := x\exp(\lambda x)$ für $x \in \mathbb{R}$. Dann ist $f_2'(x) = \exp(\lambda x) + \lambda x \exp(\lambda x) = (1 + \lambda x)\exp(\lambda x)$ und $f_2''(x) = \lambda \exp(\lambda x) + (1 + \lambda x)\lambda \exp(\lambda x) = (\lambda^2 x + 2\lambda)\exp(\lambda x)$. Es folgt $f_2''(x) + pf_2'(x) + qf_2(x) = \big(\lambda^2 x + p\lambda x + qx + 2\lambda + p\big)\exp(\lambda x)$. Hier ist $F(\lambda) = \lambda^2 + p\lambda + q = 0$; außerdem $X^2 + pX + q = (X - \lambda)^2 = X^2 - 2\lambda X + \lambda^2$, also $p = -2\lambda$. Also ist $f_2(x)$ eine Lösung der DGL. Wir erhalten hier also Lösungen der Form:

$$\boxed{f(x) = (C_1 + C_2 x)\exp(\lambda x) \text{ für alle } x \in \mathbb{R}, \text{ mit } C_1, C_2 \in \mathbb{R}}$$

Zum Beispiel ist die DGL $y'' = -c^2 y$ in Lemma 38.3 im 2. Fall; denn hier ist $p = 0$ und $q = c^2$, also $F = X^2 + c^2 = (X + ci) * (X - ci)$ (keine reelle Nullstelle).

Satz 38.13. *Seien* $p, q \in \mathbb{R}$ *fest. Dann sind alle Lösungen der DGL* $y'' + py' + q = 0$ *durch* $y = C_1 f_1(x) + C_2 f_2(x)$ *gegeben, mit* $C_1, C_2 \in \mathbb{R}$ *und* f_1, f_2 *wie in den obigen* 3 *Fällen.*

Beweis (Skizze). Sind $f_1(x)$ und $f_2(x)$ jeweils wie oben definiert, so haben wir bereits gesehen, dass $f(x) := C_1 f_1(x) + C_2 f_2(x)$ für $C_1, C_2 \in \mathbb{R}$ eine Lösung der DGL ist. Sei nun umgekehrt $y = f(x)$ eine beliebige Lösung der DGL.

Im ersten Fall bilde dann $h(x) := \exp(-\lambda_1 x)f(x)$. Man rechnet nun nach, dass $h'(x) = -\lambda_1 h(x) + \exp(-\lambda_1 x)f'(x)$ und $h''(x) = -(2\lambda_1 + p)h'(x)$ gilt. Setze dann $u(x) := h'(x)$; wegen $\lambda_1 + \lambda_2 = -p$ ist $2\lambda_1 + p = \lambda_1 - \lambda_2$. Also gilt $u'(x) = -(2\lambda_1 + p)u(x) = (\lambda_2 - \lambda_1)u(x)$.

Mit Lemma 38.2 folgt $h'(x) = u(x) = a\exp((\lambda_2 - \lambda_1)x)$ für eine Konstante $a \in \mathbb{R}$ und dann $h(x) = (\lambda_2 - \lambda_1)^{-1} a \exp((\lambda_2 - \lambda_1)x) + b$ für eine weitere Konstante $b \in \mathbb{R}$. Also erhalten wir $f(x) = \exp(\lambda_1 x)h(x) = a(\lambda_2 - \lambda_1)^{-1}\exp(\lambda_2 x) + b\exp(\lambda_1 x)$, wie gewünscht.

Im zweiten Fall bilde $h(x) := \exp(-\alpha x)f(x)$ und rechne nach, dass $h''(x) = -(2\alpha + p)h'(x) - F(\alpha)h$ gilt. Nun ist $p = -2\alpha$, also $h''(x) = -F(\alpha)h$. Aus $F(z) = 0$ folgt $\beta^2 = \alpha^2 + p\alpha + q = F(\alpha)$, d.h., $h''(x) = -\beta^2 h(x)$. Nun wende Lemma 38.3 an und schließe wie oben.

Im dritten Fall bilde wieder $h(x) = \exp(-\lambda x)f(x)$. Wiederum folgt $h''(x) = -(2\lambda + p)h'$. Nun ist $p = -2\lambda$, also $h''(x) = 0$, d.h., es gilt $h(x) = ax + b$ mit Konstanten $a, b \in \mathbb{R}$. □

Bemerkung 38.14. Sei $d \geqslant 2$ und betrachte die folgende DGL d-ter Ordnung:

$$y^{(d)} + a_{d-1}y^{(d-1)} + \ldots + a_1 y^{(1)} + a_0 y = 0 \qquad \text{mit vorgegebenen } a_0, a_1, \ldots, a_{d-1} \in \mathbb{R}.$$

Wie oben bilde das Polynom $F := X^d + a_{d-1}X^{d-1} + \ldots + a_1 X + a_0 \in \mathbb{R}[X]$. Wie in Satz 33.13 (Partialbruchzerlegung) können wir F schreiben in der Form:

$$F = \prod_{i=1}^{r}(X - \lambda_i)^{r_i} \cdot \prod_{i=1}^{s}(X^2 + p_i X + q_i)^{s_i} \qquad (\text{mit } r, s \in \mathbb{N}_0 \text{ und } r_i, s_i \in \mathbb{N}),$$

wobei $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ die verschiedenen reellen Nullstellen sind und $p_i, q_i \in \mathbb{R}$ jeweils so, dass $f_i := X^2 + p_i X + q_i \in \mathbb{R}[X]$ keine reelle Nullstelle hat (also $4q_i - p_i^2 > 0$). Schreibe $f_i = (X - z_i) * (X - \bar{z}_i)$ mit $z_i \in \mathbb{C} \setminus \mathbb{R}$; sei $\alpha_i \in \mathbb{R}$ der Realteil von z_i und $\beta_i \in \mathbb{R}$ der Imaginärteil. Es gilt dann folgende Verallgemeinerung von Satz 38.13: Alle Lösungen $y = f(x)$ der DGL sind gegeben als Linearkombinationen der folgenden, insgesamt d Funktionen:

$$x^j \exp(\lambda_i x) \quad \begin{pmatrix} i = 1, \ldots, r \\ j = 0, 1, \ldots, r_i - 1 \end{pmatrix} \quad \text{und} \quad \begin{matrix} x^j \exp(\alpha_i x)\cos(\beta_i x) \\ x^j \exp(\alpha_i x)\sin(\beta_i x) \end{matrix} \quad \begin{pmatrix} i = 1, \ldots, s \\ j = 0, 1, \ldots, s_i - 1 \end{pmatrix}.$$

(Siehe §15, Satz 2, im 2. Band des Buches von Forster.) Dazu 3 Beispiele:

1) Sei $d = 4$ und $y^{(4)} - y = 0$; dann ist $F := X^4 - 1 = (X - 1) * (X + 1) * (X^2 + 1)$ wobei $X^2 + 1$ die beiden Nullstellen $\pm i$ hat. Also ist die allgemeine Lösung der DGL gegeben durch $y = f(x) = C_1\exp(x) + C_2\exp(-x) + C_3\cos(x) + C_4\sin(x)$, mit Konstanten $C_1, C_2, C_3, C_4 \in \mathbb{R}$.

2) Sei $d = 3$ und $y^{(3)} - 6y^{(2)} + 12y^{(1)} - 8y = 0$; dann ist $F = X^3 - 6X^2 + 12X - 8 = (X - 2)^3$. Also ist die allgemeine Lösung der DGL gegeben durch $y = f(x) = (C_1 + C_2 x + C_3 x^2)\exp(2x)$, mit Konstanten $C_1, C_2, C_3 \in \mathbb{R}$.

3) Sei $d = 4$ und $y^{(4)} + 2y^{(3)} + 3y^{(2)} + 2y^{(1)} + y = 0$; dann ist $F = X^4 + 2X^3 + 3X^2 + 2X + 1 = (X^2 + X + 1)^2$. Hier ist $X^2 + X + 1 = (X - z_1) * (X - \bar{z}_1)$ mit $z_1 = (-1 + \sqrt{3}i)/2$, d.h., $\alpha_1 = -1/2$ und $\beta_1 = \sqrt{3}/2$. Also ist die allgemeine Lösung der DGL gegeben durch

$$\begin{aligned} y = f(x) = {} & C_1\exp(-x/2)\cos(\sqrt{3}x/2) + C_2 x\exp(-x/2)\cos(\sqrt{3}x/2) \\ & + C_3\exp(-x/2)\sin(\sqrt{3}x/2) + C_4 x\exp(-x/2)\sin(\sqrt{3}x/2), \end{aligned}$$

mit Konstanten $C_1, C_2, C_3, C_4 \in \mathbb{R}$.

Es gibt eine Reihe von weiteren Differentialgleichungen speziellen Typs, zum Beispeil die "Bernoullische" oder die "Riccatische" DGL; für solche Fälle finden Sie in der `wikipedia` oft erste Hinweise auf weitere Referenzen.

Ab hier Woche 14

39. *Systeme von Differentialgleichungen*

Bisher haben wir stets einzelne Differentialgleichungen behandelt. Allgemeiner betrachtet man Systeme von Differentialgleichungen, in denen n Funktionen $y_1 = f_1(x), \ldots, y_n = f_n(x)$ (in einer reellen Variablen x) und ihre Ableitungen vorkommen[2].

Solche Systeme treten in den unterschiedlichsten Anwendungen auf. Hier ein Beispiel:

Beispiel 39.1. Gegeben sei das folgende System von 2 Differentialgleichungen in Funktionen $y_1 = y_1(t)$ und $y_2 = y_2(t)$ (in einer reellen Variablen t):

$$y_1'(t) = ay_1(t) - by_1(t)y_2(t), \qquad y_2'(t) = cy_1(t)y_2(t) - dy_2(t),$$

mit Konstanten $a, b, c, d \in \mathbb{R}_{>0}$. Solche Systeme, auch ***Lotka–Volterra Gleichungen*** (oder "Räuber–Beute–Gleichungen") genannt, treten auf in der Modellierung von biologischen Systemen, in denen zwei Populationen von Lebewesen interagieren.

Es sei etwa $y_1(t)$ die Anzahl von "Beute-Lebewesen" und $y_2(t)$ die Anzahl von "Räubern" zum Zeitpunkt t (zum Beispiel Beuten=Mäuse, Räuber=Katzen). Dann entsprechen $y_1'(t)$ und $y_2'(t)$ der Änderung von $y_1(t)$ und $y_2(t)$ in Abhängigkeit von der Zeit t. Die weiteren Terme haben folgende Interpretation: a beschreibt die natürliche "Geburtenrate" der Beuten und d die natürliche "Sterberate" der Räuber. Durch die gemischten Terme $y_1(t)y_2(t)$ werden Effekte durch Begegnungen von Räubern und Beuten modelliert: Für Beuten werden diese mit einer gewissen Wahrscheinlichkeit negativ ausgehen, daher der Faktor $-b$; die Auswirkungen auf die Räuber mögen zwar nicht so klar sein, werden aber jedenfalls als positiv angenommen, mit dem Faktor c. Für weitere Erläuterungen siehe `https://de.wikipedia.org/wiki/Lotka-Volterra-Gleichungen`.

Solche Systeme zu lösen ist nicht einfach; es gibt zum Teil ganze Bücher zu einzelnen Systemen! Dabei spielen auch numerische Verfahren eine wichtige Rolle. Hier beschäftigen wir uns nun mit einigen Systemen, wo man tatsächlich komplett die Lösungen bestimmen kann.

Bemerkung 39.2. Das System der Differentialgleichungen 1. Ordnung habe folgende Form: $y_i' = g_i(x, y_1, \ldots, y_n)$ für $i = 1, \ldots, n$, wobei die Funktionen g_i (in $n+1$ Variablen) gegeben

[2] Man kann natürlich auch noch Funktionen von mehreren Variablen erlauben und dann Systeme von Differentialgleichungen betrachten, in denen diese Funktionen zusammen mit ihren partiellen Ableitungen, und auch höheren partiellen Ableitungen, vorkommen. Dann spricht man von Systemen von "partiellen Differentialgleichungen" (englisch: PDE für "partial differential equation"). Solche Systeme kommen ebenfalls in vielen Anwendungen vor; siehe zum Beispiel `https://de.wikipedia.org/wiki/Maxwell-Gleichungen`. In dieser Einführung bleiben wir hier bei dem Fall von Funktionen einer reellen Variablen.

sind; gesucht sind alle Funktionen $y_i = f_i(x)$, die dieses System erfüllen, d.h., es soll gelten:

$$f_i'(x) = g_i\big(x, f_1(x), \ldots, f_n(x)\big) \qquad \text{für } i = 1, \ldots, n$$

und alle $x \in \mathbb{R}$ im gemeinsamen Definitionsbereich der f_i. Der grundlegende ***Existenz- und Eindeutigkeitssatz von Picard–Lindelöf*** (um 1890) besagt:

*Sind ein Punkt x_0 und Werte $c_1, \ldots, c_n \in \mathbb{R}$ vorgegeben, so gibt es stets eine eindeutige Lösung $f_1, \ldots, f_n$ des Systems mit den **Anfangsbedingungen** $f_i(x_0) = c_i$ für $i = 1, \ldots, n$.*

Dabei müssen x_0 und die vorgegebenen Funktionen $g_1, \ldots, g_n$ gewisse Bedingungen erfüllen, die wir hier nicht in aller Allgemeinheit formulieren. Zum Beispiel sind diese Bedingungen erfüllt, wenn die Funktionen $g_i(x_0, x_1, \ldots, x_n)$ in den Variablen $x_1, \ldots, x_n$ stetig partiell differenzierbar sind; siehe dazu §12 (dort Satz 2) im 2. Band des Buches von Forster.

Beispiel 39.3. Gegeben seien Funktionen $\alpha_{ij}\colon D \to \mathbb{R}$ und $\beta_i\colon D \to \mathbb{R}$ für $i, j = 1, \ldots, n$. Dann erhalten wir ein lineares System von Differentialgleichungen 1. Ordnung für n differenzierbare Funktionen $y_1 = y_1(x), \ldots, y_n = y_n(x)$ (in einer reellen Variablen x) durch:

$$\begin{array}{rcccccccc}
y_1'(x) & = & \alpha_{11}(x)y_1(x) & + & \ldots & + & \alpha_{1n}(x)y_n(x) & + & \beta_1(x) \\
y_2'(x) & = & \alpha_{21}(x)y_1(x) & + & \ldots & + & \alpha_{2n}(x)y_n(x) & + & \beta_2(x) \\
\vdots & & \vdots & & & & \vdots & & \vdots \\
y_n'(x) & = & \alpha_{n1}(x)y_1(x) & + & \ldots & + & \alpha_{nn}(x)y_n(x) & + & \beta_n(x)
\end{array}$$

Ist $y(x) \in \mathbb{R}^n$ der Spaltenvektor mit Komponenten $y_1(x), \ldots, y_n(x)$ und $\beta(x)$ der Spaltenvektor mit Komponenten $\beta_1(x), \ldots, \beta_n(x)$, so können wir dies auch kurz schreiben als

$$\boxed{y'(x) = A(x) \cdot y(x) + \beta(x) \qquad \text{wobei} \qquad A(x) := \big[\alpha_{ij}(x)\big]_{1 \leqslant i,j \leqslant n}}$$

Hier ist $y'(x)$ der Spaltenvektor mit Komponenten $y_1'(x), \ldots, y_n'(x)$.

Beispiel 39.4. Gegeben sei das System $y_1 = y_1 + xy_2$ und $y_2' = (x+1)^{-1}y_2$ für $D := \mathbb{R}\backslash\{-1\}$. Es handelt sich also um ein lineares System wie oben, mit Matrix $A(x) = \begin{bmatrix} 1 & x \\ 0 & (x+1)^{-1} \end{bmatrix}$.

Versuchen wir, dieses zu lösen. Die zweite Gleichung des Systems ist eine separierbare DGL für y_2. Mit der Methode in Beispiel 38.9 findet man die allgemeine Lösung $y_2(x) = a(x+1)$, wobei $a \in \mathbb{R}$ eine Konstante ist. Setzen wir dies in die erste Gleichung ein, so erhalten wir die lineare DGL $y_1' = y_1 + ax(x+1)$. Mit Satz 38.6 ergibt sich die allgemeine Lösung als $y_1(x) = -a(x^2 + 3x + 3) + b\exp(x)$ mit einer Konstanten $b \in \mathbb{R}$. Also ist die allgemeine Lösung des Systems gegeben durch

$$y(x) = \begin{bmatrix} -a(x^2+3x+3) + b\exp(x) \\ a(x+1) \end{bmatrix} = a \begin{bmatrix} -(x^2+3x+3) \\ x+1 \end{bmatrix} + b \begin{bmatrix} \exp(x) \\ 0 \end{bmatrix} \quad \text{für } x \neq -1.$$

Bemerkung 39.5. Differentialgleichungen höherer Ordnung können in bestimmten Fällen auf ein System 1. Ordnung wie in obiger Bemerkung zurückgeführt werden. Für $d \geqslant 2$ sei eine DGL d-ter Ordnung der Form $y^{(d)} = g(x, y, y', \ldots, y^{(d-1)})$ gegeben. Betrachte dann das folgende System von d Differentialgleichungen für Funktionen $z_1 = z_1(x), \ldots, z_d = z_d(x)$:

$$z_1' = z_2, \qquad z_2' = z_3, \qquad \ldots, \qquad z_{d-1}' = z_d, \qquad z_d' = g(x, z_1, z_2, \ldots, z_d).$$

Bilden die Funktionen $z_1(x), \ldots, z_d(x)$ eine Lösung für dieses System, so ist $f(x) := z_1(x)$ eine Lösung der ursprünglichen DGL $y^{(d)} = g(x, y, y', \ldots, y^{(d-1)})$. Ist umgekehrt $f(x)$ eine Lösung der ursprünglichen DGL, so bilden die Funktionen $z_1(x) := f(x)$, $z_2(x) := f'(x)$, ..., $z_d'(x) := f^{(d-1)}(x)$ eine Lösung des obigen Systems.

Mit dem Satz von Picard–Lindelöf folgt also: Sind ein Punkt x_0 und Werte $c_0, \ldots, c_{d-1} \in \mathbb{R}$ vorgegeben, so gibt es genau eine Lösung $y = f(x)$ mit $f^{(d)}(x) = g\big(x, f(x), f'(x), \ldots, f^{(d-1)}(x)\big)$ für alle $x \in \mathbb{R}$ im Definitionsbereich von g, und $f^{(i)}(x_0) = c_i$ für $i = 0, 1, \ldots, d-1$.

Beispiel 39.6. Betrachte die DGL 2. Ordnung $y'' + py' + qy = 0$ in Beispiel 38.12. Nach obigem Schema können wir diese umwandeln in ein System von zwei Differentialgleichungen 1. Ordnung für zwei Funktionen $y_1 = y_1(x)$ und $y_2 = y_2(x)$, nämlich: $y_1' = y_2$ und $y_2' = -qy_1 - py_2$. Betrachte allgemeiner für $d \geqslant 2$ die DGL d-ter Ordnung:

$$y^{(d)} + a_{d-1}y^{(d-1)} + \ldots + a_1 y^{(1)} + a_0 y = 0 \quad \text{mit vorgegebenen } a_0, a_1, \ldots, a_{d-1} \in \mathbb{R}.$$

Diese können wir analog umwandeln in ein System für d Funktionen $y_i = y_i(x)$ wobei

$$y_1' = y_2, \;\; y_2' = y_3, \;\; \ldots, \;\; y_{d-1}' = y_d \quad \text{und} \quad y_d' = -a_0 y_1 - a_1 y_2 - \ldots - a_{d-1} y_d.$$

Mit obigen Bezeichnungen können wir dies auch kompakt schreiben als

$$y'(x) = A \cdot y(x) \qquad \text{mit} \qquad A = \begin{bmatrix} 0 & 1 & 0 & \ldots & & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ 0 & 0 & \ldots & 0 & 1 & 0 \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & & \ldots & -a_{d-1} \end{bmatrix} \in M_d(\mathbb{R}).$$

Wir untersuchen im Folgenden lineare Systeme von Differentialgleichungen 1. Ordnung der Form $y'(x) = A \cdot y(x)$ für $x \in D$, wobei $A \in M_n(\mathbb{R})$ eine beliebige, aber konstante Matrix ist. Dazu ist es nützlich, nicht nur Funktionen f mit Definitionsbereich $D \subseteq \mathbb{R}$ und Werten in $\mathbb{R}^n$ zu betrachten, sondern auch Werte in $\mathbb{C}^n$ zuzulassen. Es ist also

$$f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix} \in \mathbb{C}^n \qquad \text{mit Funktionen} \quad f_k \colon D \to \mathbb{C} \quad \text{für } k = 1, \ldots, n.$$

Indem wir jedes $f_k(x)$ in Real- und Imaginärteil zerlegen, können wir f_k auch als Funktion $f_k \colon D \to \mathbb{R}^2$ auffassen und dann insgesamt f als Funktion $f \colon D \to \mathbb{R}^{2n}$. Die Funktion f ist differenzierbar, wenn alle $2n$ Komponentenfunktionen im üblichen Sinne differenzierbar sind. In diesem Fall sei $f'(x) \in \mathbb{C}^n$ der Spaltenvektor mit Komponenten $f_1'(x), \ldots, f_n'(x)$. Wie in Bemerkung 36.15 ergibt sich also eigentlich nichts Neues; aber wiederum dadurch, dass $\mathbb{C}$ selbst ein Körper ist, vereinfachen sich manche Formeln (so wie in Beispiel 36.16).

Ein lineares System 1. Ordnung wie oben sieht formal so aus wie eine einzelne lineare DGL wie in Satz 38.6. Diese formale Analogie funktioniert tatsächlich: Wir werden nun zeigen, dass man auch $\exp(A)$ für Matrizen $A \in M_n(\mathbb{C})$ bilden kann und damit Lösungen des Systems $y'(x) = A \cdot y(x)$ erhält. Da dies in einführenden Lehrbüchern oft etwas knapp behandelt wird, gehen wir die Definition von $\exp(A) \in M_n(\mathbb{C})$ im Detail durch.

Zur Erinnerung: In §23 haben wir die Norm eines Spaltenvektors $v \in \mathbb{C}^n$ definiert durch $\|v\| = \sqrt{|v_1|^2 + \ldots + |v_n|^2}$, wobei $v_1, \ldots, v_n \in \mathbb{C}$ die Komponenten von v sind und $|v_i|$ jeweils der komplexe Absolutbetrag ist. Wir setzen nun analog:

$$\|A\| := \sqrt{\sum\nolimits_{1\leqslant k,l\leqslant n} |a_{kl}|^2} \qquad \text{für eine Matrix } A = \left[a_{kl}\right]_{1\leqslant k,l\leqslant n} \in M_n(\mathbb{C}).$$

Lemma 39.7. *Es gilt* $\|v + w\| \leqslant \|v\| + \|w\|$ *für alle* $v, w \in \mathbb{C}^n$ *und* $\|A \cdot v\| \leqslant \|A\|\|v\|$ *für alle* $A \in M_n(\mathbb{C})$ *und* $v \in \mathbb{C}^n$.

Beweis. Seien $v_1, \ldots, v_n \in \mathbb{C}$ die Komponenten von v und $w_1, \ldots, w_n \in \mathbb{C}$ die Komponenten von w. Mit der Dreiecksungleichung für den komplexen Absolutbetrag gilt $|v_k + w_k| \leqslant |v_k| + |w_k|$ für alle k. Damit folgt

$$\|v + w\|^2 = \sum_{k=1}^{n} |v_k + w_k|^2 \leqslant \sum_{k=1}^{n} \big(|v_k| + |w_k|\big)^2 = \sum_{k=1}^{n} \big(|v_k|^2 + 2|v_k||w_k| + |w_k|^2\big).$$

Weil die Terme $|v_k|$, $|w_k|$ alle reell sind, können wir die Cauchy–Schwarz–Ungleichung in Kapitel IV, §19, anwenden und erhalten $\sum_{k=1}^{n} |v_k||w_k| \leqslant \big(\sum_{k=1}^{n} |v_k|^2\big)^{1/2} \cdot \big(\sum_{k=1}^{n} |w_k|^2\big)^{1/2} = \|v\| \cdot \|w\|$. Also ist die rechte Seite der obigen Ungleichung $\leqslant \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2$ und es folgt $\|v + w\| \leqslant \|v\| + \|w\|$.

Nun zu $\|A \cdot v\|$: Die k-te Komponente von $A \cdot v$ ist gegeben durch $\sum_{l=1}^{n} a_{kl} v_l$. Indem wir zunächst wieder die Dreiecksungleichung für den komplexen Absolutbetrag und die Cauchy–Schwarz–Ungleichung benutzen, folgt:

$$\begin{aligned}\|A \cdot v\|^2 &= \sum_{k=1}^{n} \Big|\sum_{l=1}^{n} a_{kl} v_l\Big|^2 \leqslant \sum_{k=1}^{n} \Big(\sum_{l=1}^{n} |a_{kl}||v_l|\Big)^2 \leqslant \sum_{k=1}^{n} \Big(\sum_{l=1}^{n} |a_{kl}|^2\Big)\Big(\sum_{l'=1}^{n} |v_{l'}|^2\Big) \\ &= \sum_{k,l,l'=1}^{n} |a_{kl}|^2 |v_{l'}|^2 = \Big(\sum_{k,l=1}^{n} |a_{kl}|^2\Big)\Big(\sum_{l'=1}^{n} |v_{l'}|^2\Big) = \|A\|^2 \cdot \|v\|^2.\end{aligned}$$

Also gilt auch $\|A \cdot v\| \leqslant \|A\| \cdot \|v\|$. □

Folgerung 39.8. *Setzen wir* $d(A, B) := \|A - B\|$ *für* $A, B \in M_n(\mathbb{C})$, *so wird* $M_n(\mathbb{C})$ *ein metrischer Raum. Es gilt außerdem* $\|A \cdot B\| \leqslant \|A\| \cdot \|B\|$ *für alle* $A, B \in M_n(\mathbb{C})$.

Beweis. Zunächst sind Symmetrie und Definitheit von $d(A, B)$ klar. Die Dreiecksungleichung folgt ebenso leicht aus der Dreicksungleichung in Lemma 39.7 für die (Euklidische) Norm auf $\mathbb{C}^{n^2}$, indem wir $M_n(\mathbb{C})$ mit $\mathbb{C}^{n^2}$ identifizieren (also die n^2 Einträge einer Matrix einfach als Spaltenvektor mit n^2 Komponenten schreiben). Damit ist $d(A, B)$ eine Metrik auf $M_n(\mathbb{C})$.

Nun zu $\|A\cdot B\|$: Die k-te Spalte von $A\cdot B$ ist gegeben durch $A\cdot b_k$. Also folgt mit Lemma 39.7:
$$\|A\cdot B\|^2=\sum_{k=1}^n\|A\cdot b_k\|^2\leqslant\sum_{k=1}^n\|A\|^2\|b_k\|^2=\|A\|^2\sum_{k=1}^n\|b_k\|^2=\|A\|^2\|B\|^2,$$
und damit auch $\|A\cdot B\|\leqslant\|A\|\cdot\|B\|$. □

Sei nun $(A_r)_{r\in\mathbb{N}_0}$ eine Folge von Matrizen in $M_n(\mathbb{C})$. Für $r\in\mathbb{N}_0$ definieren wir die zugehörige Partialsumme durch $\mathcal{A}_r:=A_0+A_1+\ldots+A_r\in M_n(\mathbb{C})$. Dann heißt (wie zuvor) die unendliche Reihe $\sum_{r=0}^\infty A_r$ konvergent, wenn die Folge der Partialsummen $(\mathcal{A}_r)_{r\in\mathbb{N}_0}$ in $M_n(\mathbb{C})$ konvergiert.

Folgerung 39.9. *Sei $(A_r)_{r\in\mathbb{N}_0}$ eine Folge von Matrizen in $M_n(\mathbb{C})$. Es gebe eine Konstante $C\in\mathbb{R}$ mit $\mathcal{A}_r^*:=\|A_0\|+\|A_1\|+\ldots+\|A_r\|\leqslant C$ für alle $r\in\mathbb{N}_0$. Dann konvergiert die unendliche Reihe $\sum_{r=0}^\infty A_r$ in $M_n(\mathbb{C})$.*

Beweis. Schreibe $A_r=\big[a_{kl}^{(r)}\big]_{1\leqslant k,l\leqslant n}\in M_n(\mathbb{C})$. Seien zunächst $k,l\in\{1,\ldots,n\}$ fest. Bilde die Partialsumme $Z_{kl}^{(r)}:=a_{kl}^{(0)}+\ldots+a_{kl}^{(r)}\in\mathbb{C}$ sowie $Z_{kl}^{(r)*}:=|a_{kl}^{(0)}|+\ldots+|a_{kl}^{(r)}|\in\mathbb{R}$ für $r\in\mathbb{N}_0$. Nun ist $|a_{kl}^{(r)}|\leqslant\|A_r\|$ für alle $r\in\mathbb{N}_0$, also auch $Z_{kl}^{(r)*}\leqslant\mathcal{A}_r^*\leqslant C$ für alle $r\in\mathbb{N}_0$. Folglich existiert $\hat{Z}_{kl}:=\lim\limits_{r\to\infty}(Z_{kl}^{(r)})\in\mathbb{C}$; siehe Satz 30.3. Definiere nun die Matrix $\hat{Z}:=\big[\hat{Z}_{kl}\big]_{1\leqslant k,l\leqslant n}\in M_n(\mathbb{C})$. Wir zeigen, dass $(\mathcal{A}_r)_{r\in\mathbb{N}_0}$ konvergiert, mit Grenzwert $\hat{Z}$. Sei $\varepsilon>0$. Es gibt ein $r_0\in\mathbb{N}_0$ mit $|Z_{kl}^{(r)}-\hat{Z}_{kl}|<\varepsilon/n$ für alle $r\geqslant r_0$ und alle k,l. Es folgt $\|\mathcal{A}_r-\hat{Z}\|^2=\sum_{k,l=1}^n|Z_{kl}^{(r)}-\hat{Z}_{kl}|^2<\varepsilon^2$ und damit $\|\mathcal{A}_r-\hat{Z}\|<\varepsilon$ für alle $r\geqslant r_0$. Also konvergiert $(\mathcal{A}_r)_{r\in\mathbb{N}_0}$ in $M_n(\mathbb{C})$ gegen $\hat{Z}$. □

Satz 39.10 (***Exponential-Funktion für Matrizen***)**.** *Für jedes $A\in M_n(\mathbb{C})$ konvergiert die unendliche Reihe $\sum_{r=0}^\infty\frac{A^r}{r!}$ und wir setzen $\exp(A):=\sum_{r=0}^\infty\frac{A^r}{r!}\in M_n(\mathbb{C})$. Es gilt:*

(a) $$\Big\|\exp(A)-\sum_{r=0}^s\frac{A^r}{r!}\Big\|\leqslant\frac{2\|A\|^{s+1}}{(s+1)!}\qquad\textit{für jedes } s\in\mathbb{N}_0 \textit{ mit } s\geqslant 2\|A\|-2,$$

(b) $\exp(A+B)=\exp(A)\cdot\exp(B)$ *für alle $A,B\in M_n(\mathbb{C})$ mit $A\cdot B=B\cdot A$.*

(c) $\exp(A)$ *ist invertierbar mit* $\exp(A)^{-1}=\exp(-A)$.

Beweis. Sei $A\in M_n(\mathbb{C})$ fest und $\alpha:=\|A\|\in\mathbb{R}_{\geqslant 0}$. Aus Folgerung 39.9 erhalten wir zunächst $\|A^r\|\leqslant\|A\|^r=\alpha^r$ für alle $r\in\mathbb{N}_0$. Also folgt $\mathcal{A}_r^*=\sum_{s=0}^r\|\frac{A^s}{s!}\|\leqslant\sum_{s=0}^r\frac{\alpha^s}{s!}$ für alle $r\in\mathbb{N}_0$. Die rechts Seite ist die r-te Partialsumme zu der Reihe, durch die $\exp(\alpha)\in\mathbb{R}$ definiert ist. Wegen $\alpha\geqslant 0$ folgt also $\mathcal{A}_r^*\leqslant C:=\exp(\alpha)$ für alle $r\in\mathbb{N}_0$; also ist die unendliche Reihe $\sum_{r=0}^\infty\frac{A^r}{r!}$ konvergent. Die Abschätzung in (a) zeigt man nun genauso wie im Beweis von Satz 30.5. Zu (b): Man überzeugt sich, dass die Aussage über das Cauchy–Produkt in Satz 28.13 auch für Reihen in $M_n(\mathbb{C})$ gilt (mit wörtlich fast dem gleichen Beweis). Die Gleichung $\exp(A+B)=\exp(A)\cdot\exp(B)$ folgt dann genauso wie in Satz 30.5; dabei wird die Voraussetzung $A\cdot B=B\cdot A$ benutzt, um die Binomische Formel anwenden zu können (so wie im Beweis von Satz 29.7). Zu (c): Mit der Definition sieht man sofort, dass $\exp(0_{n\times n})=I_n$ gilt. Mit (b) folgt also $I_n=\exp(0_{n\times n})=\exp(A+(-A))=\exp(A)\cdot\exp(-A)$. □

Sei $A \in M_n(\mathbb{C})$ fest. Für $v \in \mathbb{C}^n$ definieren wir nun eine Funktion $\mathfrak{f}_v \colon \mathbb{R} \to \mathbb{C}^n$ durch

$$\boxed{\mathfrak{f}_v(x) = \exp(xA) \cdot v \in \mathbb{C}^n \qquad \text{für alle } x \in \mathbb{R}}$$

Damit können wir den folgenden ***Hauptsatz über lineare Systeme von Differentialgleichungen 1. Ordnung mit konstanten Koeffizienten*** formulieren:

Satz 39.11. *Sei $A \in M_n(\mathbb{C})$ fest. Für $v_0 \in \mathbb{C}^n$ ist die Funktion $\mathfrak{f}_{v_0} \colon \mathbb{R} \to \mathbb{C}^n$ die eindeutige Lösung des Systems $y'(x) = A \cdot y(x)$ mit Anfangswert v_0 für $x = 0$. Bilden $v_1, \ldots, v_n \in \mathbb{C}^n$ eine Basis von $\mathbb{C}^n$, so ist jede Lösung des Systems eine eindeutige Linearkombination von $\mathfrak{f}_{v_1}, \ldots, \mathfrak{f}_{v_n}$. Wir bezeichnen dann $\{\mathfrak{f}_{v_1}, \ldots, \mathfrak{f}_{v_n}\}$ als ein **Fundamentalsystem** von Lösungen.*

Beweis. Nachdem die obigen Sätze zur Norm auf $M_n(\mathbb{C})$ und zu Eigenschaften der Exponential-Funktion für Matrizen zur Verfügung stehen, geht der Beweis im Wesentlichen genauso wie in Lemma 38.2. Für die Details siehe zum Beispiel Kapitel 6, Hauptsatz 6.4.3, im Buch von Huppert–Willems, oder Kapitel 4, Satz (8.14), im Buch von Artin. □

Für konkrete Rechnungen in Beispielen muss man jetzt noch sehen, wie man $\exp(xA)$ explizit auswerten kann. Dazu benutzt man den ***Struktursatz für zerfallende Matrizen***, den wir am Ende des 1. Semesters behandelt haben.

Bemerkung 39.12. Sei $A \in M_n(\mathbb{C})$. Seien $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ die paarweise verschiedenen Eigenwerte von A. Nach Kapitel IV, Satz 21.2, gibt es eine invertierbare Matrix $T \in M_n(\mathbb{C})$ so dass $\hat{A} := T^{-1} \cdot A \cdot T$ eine Blockdiagonalgestalt hat mit Diagonalblöcken $A_1, \ldots, A_r$, wobei $A_k \in M_{n_k}(\mathbb{C})$ und $\chi_{A_k} = (\lambda_k - X)^{n_k}$ für $k = 1, \ldots, r$ gilt; hier ist jeweils $n_k \geqslant 1$ die Dimension des Hauptraums H_{λ_k} zu λ_k. Entsprechend gilt $x\hat{A} = T^{-1} \cdot (xA) \cdot T$ für $x \in \mathbb{R}$. Mit der Definition von $\exp(xA)$ und $\exp(x\hat{A})$ sieht man sofort:

(a) $\qquad \exp(xA) = T \cdot \exp(x\hat{A}) \cdot T^{-1}$ für alle $x \in \mathbb{R}$.

Nun beachte: Setzen wir $N_k := A_k - \lambda_k I_{n_k} \in M_{n_k}(\mathbb{C})$, so ist jedes N_k eine nilpotente Matrix (siehe Kapitel IV, Bemerkung 21.6). Indem wir jeden Diagonalblock in $\hat{A}$ als $A_k = \lambda_k I_{n_k} + N_k$ schreiben, erhalten wir insgesamt auch eine Zerlegung $\hat{A} = D + N$, wobei $D \in M_n(\mathbb{C})$ eine Diagonalmatrix und $N \in M_n(\mathbb{C})$ eine nilpotente Matrix ist; außerdem gilt hier $D \cdot N = N \cdot D$ (weil die entsprechende Vertauschbarkeit offensichtlich in jedem Diagonalblock A_k gilt)[3]. Entsprechend gilt $x\hat{A} = xD + xN$ für $x \in \mathbb{R}$, wobei weiterhin xD diagonal ist, xN nilpotent und $(xD) \cdot (xN) = (xN) \cdot (xD)$ gilt. Also folgt mit Satz 39.10(b):

(b) $\qquad \exp(x\hat{A}) = \exp(xD) \cdot \exp(xN)$ für alle $x \in \mathbb{R}$.

[3] Umgekehrt ist $A = D' + N'$ wobei $D' = T \cdot D \cdot T^{-1}$ diagonalisierbar ist und $N' = T \cdot N \cdot T^{-1}$ nilpotent, sowie weiterhin $D' \cdot N' = N' \cdot D'$ gilt. Dies bezeichnet man als ***Jordan–Chevalley–Zerlegung*** von A; die Matrizen D' und N' sind hier eindeutig bestimmt. Siehe zum Beispiel Kapitel 8, §5, im Buch von Koecher.

Lemma 39.13. (a) *Sei* $D \in M_n(\mathbb{C})$ *eine Diagonalmatrix mit* $z_1, \ldots, z_n \in \mathbb{C}$ *auf der Diagonalen. Für* $x \in \mathbb{R}$ *ist dann auch* $\exp(xD) \in M_n(\mathbb{C})$ *eine Diagonalmatrix, mit Diagonaleinträgen* $\exp(z_k x) \in \mathbb{C}$ *für* $k = 1, \ldots, n$.
(b) *Ist* $N \in M_n(\mathbb{C})$ *nilpotent und* $m \geqslant 1$ *so, dass* $N^m = 0_{n\times n}$ *gilt, so ist* $\exp(xN)$ *für* $x \in \mathbb{R}$ *gleich der endlichen Summe* $\sum_{r=0}^{m-1} \frac{N^r}{r!} x^r$.

Beweis. Dies folgt leicht aus Satz 39.10. □

Beispiel 39.14. (a) Gegeben sei $A = \begin{bmatrix} 3 & 1 & 1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix} \in M_n(\mathbb{C})$. Hier ist $\chi_A = (X-2)^4$; es gibt also nur einen Eigenwert $\lambda_1 = 2$. Für $N = A - 2I_4$ berechnet man $N^3 = 0_{4\times 4}$. Damit ergibt sich einfach $\exp(xA) = \exp(2xI_4 + xN) = \exp(2x)\exp(xN)$ mit

$$\exp(xN) = I_4 + xN + N^2x^2/2 = \begin{bmatrix} x{+}1 & x & x & -x \\ x^2/2 & x^2/2{-}x{+}1 & x^2/2{-}x & -x^2/2{+}x \\ 0 & 0 & 1 & 0 \\ x^2/2{+}x & x^2/2 & x^2/2 & -x^2/2{+}1 \end{bmatrix}.$$

(b) Sei $A = \begin{bmatrix} 2 & 6 & 6 & 4 \\ 1 & 6 & 4 & -2 \\ -1 & -7 & -5 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \in M_4(\mathbb{C})$. Hier ist $\chi_A = (X-2)^2(X^2+1)$.

Es gibt also 3 Eigenwerte $\lambda_1 = 2$, $\lambda_2 = i$ und $\lambda_3 = -i$. Hier ist $\hat{A} = T^{-1} \cdot A \cdot T = D + N$ mit

$$D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{bmatrix}, \quad N = \begin{bmatrix} -1/2 & 1/2 & 0 & 0 \\ -1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad T = \begin{bmatrix} -1 & 1 & 2{-}2i & 2{+}2i \\ 1 & 1 & 2 & 2 \\ -1 & -1 & -3{+}i & -3{-}i \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

(Dazu muss man die Prozedur im Beweis zum Struktursatz in Kapitel IV, Satz 21.2, durchgehen.) Wegen $N^2 = 0_{4\times 4}$ ist $\exp(xN) = 1 + xN$: außerdem ist $\exp(xD)$ die Diagonalmatrix mit $\exp(2x), \exp(2x), \exp(xi), \exp(-xi)$ auf der Diagonalen. Daraus erhält man

$$\exp(x\hat{A}) = \begin{bmatrix} (1{-}x/2)\exp(2x) & x\exp(2x)/2 & 0 & 0 \\ -x\exp(2x) & (1{+}x/2)\exp(2x) & 0 & 0 \\ 0 & 0 & \exp(xi) & 0 \\ 0 & 0 & 0 & \exp(-xi) \end{bmatrix}$$

und man kann schließlich auch $\exp(xA) = T \cdot \exp(x\hat{A}) \cdot T^{-1}$ für alle $x \in \mathbb{R}$ berechnen.

Index