Kurzskript zur Vorlesung
Mathematik I und II
für Maschinenbau und Bauingenieurwesen

Prof. Dr. Ulrich Reif
Prof. Dr. Priska Jahnke
PD. Dr. Robert Haller-Dintelmann

Vorbemerkung:

Im vorliegenden Kurzskript werden wesentliche Begriffe, Resultate und Methoden zu den Vorlesungen Mathematik I und II zusammen gestellt. Aufgrund des skizzenhaften Charakters kann es weder den Besuch der Vorlesung noch der Übungen ersetzen.

Korrekturen senden Sie bitte per Email an `reif@mathematik.tu-darmstadt.de`.

Inhaltsverzeichnis

1 Grundbegriffe und Bezeichnungen 4

2 Vektorrechnung 16

3 Komplexe Zahlen 29

4 Lineare Gleichungssysteme 34

5 Matrizenrechnung 47

6 Lineare Abbildungen 56

7 Eigenwerte und -vektoren 65

8 Folgen 76

9 Reihen 82

10 Funktionengrenzwert und Stetigkeit 87

11 Differenziation 90

12 Integration 97

13 Taylor-Reihen 110

14 Fourier-Reihen 119

15 Funktionen mehrerer Veränderlicher 124

16 Differenziation in mehreren Veränderlichen 132

17 Taylor-Reihen in mehreren Veränderlichen 144

18 Extrema 148

19 Parameterintegrale 157

20 Kurvenintegrale 162

21 Integrale im $\mathbb{R}^n$ 174

22 Vektoranalysis 184

1 Grundbegriffe und Bezeichnungen

1.1 Mengen: Die mathematisch korrekte Definition des Mengenbegriffs ist eine überraschend komplizierte Angelegenheit. Für unsere Zwecke genügt aber eine „naive“ Beschreibung, die auf Georg Cantor zurückgeht. Demnach ist eine *Menge* eine Zusammenfassung bestimmter, wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens zu einem Ganzen. Die Objekte der Menge heißen *Elemente* der Menge. Wenn m ein Element der Menge M ist, dann schreiben wir $m \in M$ und anderenfalls $m \notin M$. Mengen können auf unterschiedliche Weise angegeben werden:

- vollständige Aufzählung, z.B. $P = \{\text{Hund, Biene, Spinne}\}$
- unvollständige Aufzählung, z.B. $Q = \{2, 4, 6, 8, \dots\}$
- Angabe von Eigenschaften, z.B. $R = \{x \in \mathbb{N} : |x - 5| < 3\}$, lies *„$R$ ist die Menge aller natürlichen Zahlen x mit der Eigenschaft, dass sich x von der Zahl 5 um weniger als 3 unterscheidet“.*

Die Menge, die keine Elemente enthält, heißt *leere Menge* und wird mit $\emptyset$ bezeichnet.

1.2 Mengen von Zahlen: Wir verwenden folgende Bezeichnungen:

$\mathbb{N} = \{1, 2, 3, \dots\}$	natürliche Zahlen
$\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$	natürliche Zahlen mit 0
$\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$	ganze Zahlen
$\mathbb{Q} = \{p/q : p \in \mathbb{Z} \text{ und } q \in \mathbb{N}\}$	rationale Zahlen
$\mathbb{R} =$ Menge aller Dezimalzahlen	reelle Zahlen
$\mathbb{R}_{>0} = \{x \in \mathbb{R} : x > 0\}$	positive reelle Zahlen
$\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$	nichtnegative reelle Zahlen
$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$	abgeschlossenes Intervall
$(a, b) = \{x \in \mathbb{R} : a < x < b\}$	offenes Intervall
$[a, b) = \{x \in \mathbb{R} : a \leq x < b\}$	halboffenes Intervall
$(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$	halboffenes Intervall
$\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$	Menge der komplexen Zahlen

1.3 Beziehungen zwischen Mengen:

- *Gleichheit:* Zwei Mengen stimmen genau dann überein, wenn Sie dieselben Elemente haben. Dabei ist die Reihenfolge unerheblich, z.B. gilt im Beispiel oben $R = \{7, 4, 3, 6, 5\} = \{3, 4, 5, 7, 6\}$.
- *Teilmenge:* Wenn jedes Element einer Menge A auch in der Menge B enthalten ist, dann heißt A *Teilmenge von* B und B heißt *Obermenge* von A. Wir schreiben dann

$$A \subset B \quad \text{oder auch} \quad B \supset A.$$

1.4 Mengenoperationen: Seien A und B zwei Mengen, dann definieren wir:

- *Vereinigung:* Die Vereinigung von A und B enthält alle Elemente, die in A oder B enthalten sind,
$$A \cup B = \{x : x \in A \text{ oder } x \in B\}.$$

- *Schnitt:* Der Schnitt von A und B enthält alle Elemente, die in A und B enthalten sind,
$$A \cap B = \{x : x \in A \text{ und } x \in B\}.$$

- *Differenz:* Die Differenz von A und B enthält alle Elemente, die in A, aber nicht in B enthalten sind,
$$A \setminus B = \{x : x \in A \text{ und } x \notin B\}.$$

- *Kartesisches Produkt:* Das kartesische Produkt von A und B ist die Menge aller geordneten Paare, deren erstes Element in A und deren zweites Element in B liegt,
$$A \times B = \{(x, y) : x \in A \text{ und } y \in B\}.$$
Für das kartesische Produkt einer Menge mit sich selbst schreibt man auch $A^2 = A \times A$.

1.5 Beispiel: Sei $A = [-2, 4]$ und $B = (1, 5]$. Dann gilt
$$A \cup B = [-2, 5], \quad A \cap B = (1, 4], \quad A \setminus B = [-2, 1].$$
Das kartesische Produkt $A \times B$ bildet ein Rechteck in der Ebene.

1.6 Funktionen: Eine Vorschrift f, die *jedem* Element x einer Menge D genau ein Element y einer anderen Menge Z zuordnet, heißt *Funktion*. Wir schreiben dann
$$f : D \to Z, \quad f(x) = y.$$
Die Menge D heißt *Definitionsmenge* (oder auch Definitionsgebiet) von f, die Menge Z heißt *Zielmenge* von f. In der Beziehung $f(x) = y$ nennt man x das *Argument* von f und y das *Bild* von x. Umgekehrt nennt man x das *Urbild* von y. Man beachte, dass das Bild von x stets eindeutig bestimmt ist, während es zu einem gegebenen $y \in Z$ auch mehrere oder gar keine Urbilder geben kann. Die Menge aller Bilder $B = \{f(x) : x \in D\}$ ist eine Teilmenge von Z und wird als *Bildmenge* von f bezeichnet. Die Menge aller Urbilder zu einer gegebenen Teilmenge $Y \subset Z$ ist eine Teilmenge von D und wird in der Form $f^{-1}(Y) = \{x \in D : f(x) \in Y\}$ geschrieben.
Um auszudrücken, dass die Mengen D, Z, B zu einer gegebenen Funktion f gehören, schreiben wir gelegentlich auch D_f, Z_f, B_f.

1.7 Beispiel: Sei $D = (-1, 2], Z = \mathbb{R}$ und $f : D \to Z, f(x) = x^2 - 1$. Dann ist $B = [-1, 3], f(1) = 0$ und $f^{-1}(\{0\}) = \{-1, 1\}$.

1.8 Beispiel:

- Seien P und Q die Mengen aus Abschnitt 1.1 und $f : P \to Q^2$ die Funktion, die jedem Tier die Zahl seiner Beine und die Zahl seiner Augen zuordnet. Dann ist

$$f(\text{Hund}) = (4, 2), \quad f(\text{Biene}) = (6, 2), \quad f(\text{Spinne}) = (8, 8),$$

  und die Bildmenge ist $B = \{(4, 2), (6, 2), (8, 8)\}$.

- Sei $g : \mathbb{R}^2 \to \mathbb{R}$ die Funktion, die jedem Punkt der Ebene seinen Abstand zum Ursprung zuordnet, dann ist

$$g(x, y) = \sqrt{x^2 + y^2}$$

  und die Bildmenge ist $B = \mathbb{R}_{\geq 0}$.

- Die Vorschrift, die jeder natürlichen Zahl ihre Teiler zuweist, ist keine Funktion von $\mathbb{N}$ nach $\mathbb{N}$, da es Argumente gibt, denen mehr als ein Element der Zielmenge zugewiesen wird.

- Die Vorschrift, die jeder reellen Zahl ihren Kehrwert zuweist, ist keine Funktion von $\mathbb{R}$ nach $\mathbb{R}$, da der Kehrwert von 0 nicht definiert ist.

1.9 Eigenschaften von Funktionen: Eine Funktion $f : D \to Z$ heißt

- *injektiv*, wenn voneinander verschiedene Argumente auch voneinenader verschiedene Bilder besitzen,

$$x_1 \neq x_2 \quad \Rightarrow \quad f(x_1) \neq f(x_2),$$

- *bijektiv*, wenn sie injektiv ist und die Bildmenge mit der Zielmenge übereinstimmt.

1.10 Beispiel:

- Die Funktion $f_1 : \mathbb{R} \to \mathbb{R}, f_1(x) = x^2$ ist weder injektiv noch bijektiv (nicht injektiv, da z.B. $f_1(1) = f_1(-1)$).

- Die Funktion $f_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}, f_2(x) = x^2$ ist injektiv, aber nicht bijektiv (Bildmenge ist $B = \mathbb{R}_{\geq 0}$, Zielmenge ist $\mathbb{R}$).

- Die Funktion $f_3 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, f_3(x) = x^2$ ist bijektiv.

Man beachte, dass alle drei Funktion trotz der übereinstimmenden Funkionsvorschrift voneinander verschieden sind, da die Definitions- und Zielmengen nicht gleich sind.

1.11 Umkehrfunktion: Sei $f : D_f \to Z_f$ eine bijektive Funktion, dann gibt es zu jedem Element der Zielmenge Z_f genau ein Urbild in D_f. Die Vorschrift, die jedem $y \in Z_f$ das Urbild $x \in D_f$ zuweist, ist eine Funktion von Z nach D und wird *Umkehrfunktion* genannt und mit f^{-1} bezeichnet:

$$f^{-1} : Z_f \to D_f, \quad f^{-1}(y) = x \text{ für dasjenige } x \in D_f \text{ mit } f(x) = y.$$

Es gilt dann

- f^{-1} ist injektiv.
- $D_{f^{-1}} = Z_f = B_f$ und $Z_{f^{-1}} = B_{f^{-1}} = D_f$.
- $f^{-1}(f(x)) = x$ für alle $x \in D_f$.
- $f(f^{-1}(y)) = y$ für alle $y \in Z_f$.
- Sind sowohl Definitsmenge als Zielmenge eine Teilmenge von $\mathbb{R}$, dann erhält man das Schaubild von f^{-1} aus dem Schaubild von f durch Spiegelung an der ersten Winkelhalbierenden.

1.12 Beispiel [→ 1.10]: Die Funktion f_3 ist bijektiv und ihre Umkehrfunktion hat die Form

$$f_3^{-1} : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}, \quad f_3^{-1}(y) = \sqrt{y}.$$

1.13 Verkettung: Seien $f : D_f \to Z_f$ und $g : D_g \to Z_g$ zwei Funktionen. Wenn $B_f \subset D_g$ gilt, dann ist die *Verkettung* $h = g \circ f$ (lies "g nach f") definiert durch

$$h : D_f \to Z_g, \quad h(x) = g(f(x)).$$

1.14 Beispiel [→ 1.8]: Die Verkettung $h = g \circ f$ ist definiert, da die Bildmenge Q^2 von f eine Teilmenge der Definitionsmenge $\mathbb{R}^2$ von g ist. h ist eine Funktion von P nach $\mathbb{R}$ mit

$$h(\text{Hund}) = 2\sqrt{5}, \quad h(\text{Biene}) = 2\sqrt{10}, \quad h(\text{Spinne}) = 8\sqrt{2}.$$

1.15 Beispiel: Sei $f(x) := \frac{1}{1-x}$, $D_f := (-\infty, 1)$ und $g(x) := \cos x$.

- Für $D_g = \mathbb{R}$ ist $B_g = [-1, 1] \not\subset D_f$. Die Verkettung $f \circ g$ ist also nicht definiert.
- Für $D_g = (0, 6)$ ist $B_g = [-1, 1) \subset D_f$ und damit

$$h(x) = f(g(x)) = \frac{1}{1 - \cos x}, \quad x \in (0, 6).$$

1.16 Eigenschaften reeller Funktionen: Eine Funktion $f : D \to Z$ mit $D, Z \subseteq \mathbb{R}$ wird auch *reelle Funktion* genannt. Zusätzlich zu den oben schon genannten spielen hier noch die folgende Eigenschaften eine wichtige Rolle: Eine reelle Funktion f heißt

- *monoton wachsend/fallend*, wenn $f(x_1) \leq f(x_2)$ bzw. $f(x_1) \geq f(x_2)$ für alle $x_1, x_2 \in D$ mit $x_1 < x_2$.
- *streng monoton wachsend/fallend*, wenn $f(x_1) < f(x_2)$ bzw. $f(x_1) > f(x_2)$ für alle $x_1, x_2 \in D$ mit $x_1 < x_2$.
- *beschränkt*, wenn es eine Zahl $c \in \mathbb{R}$ gibt, sodass $|f(x)| \leq c$ für alle $x \in D$.
- *gerade/ungerade*, wenn $f(x) = f(-x)$ bzw. $f(x) = -f(-x)$ für alle $x \in D$.
- *periodisch* mit *Periode* L, bzw. *L-periodisch*, wenn $D = \mathbb{R}$ und $f(x + L) = f(x)$ für alle $x \in \mathbb{R}$ gilt.

1.17 Beispiel: Sei $f(x) := 1/(1+x^2)$, $x \in D$.

- Für $D = \mathbb{R}_{\geq 0}$ ist f injektiv und $B = (0, 1]$.
- Für $D = [-1, 2)$ ist f nicht injektiv und $B = (1/5, 1]$.

1.18 Beispiel: Sei $f(x) := x + |x|$, $x \in D$.

- Für $D = \mathbb{R}$ ist f monoton wachsend und unbeschränkt.
- Für $D = [0, 3]$ ist f streng monoton wachsend und beschränkt.
- Das Monom x^n ist eine gerade Funktion, wenn n gerade ist und eine ungerade Funktion, wenn n ungerade ist.

1.19 Regeln:

- Die Verkettung zweier injektiver Funktionen ist injektiv.
- Die Verkettung zweier bijektiver Funktionen f und g ist bijektiv und für die Umkehrfunktionen gilt die Formel

$$(f \circ g)^{-1} = g^{-1} \circ f^{-1}.$$

- Aus strenger Monotonie folgt Injektivität.
- Die Verkettung zweier monoton wachsender Funktionen ist monoton wachsend.
- Die Verkettung zweier monoton fallender Funktionen ist monoton wachsend.
- Summe und Produkt zweier gerader Funktionen sind gerade.
- Die Summe zweier ungerader Funktionen ist ungerade, aber ihr Produkt ist gerade.
- Die Verkettung zweier gerader Funktionen ist gerade.
- Die Verkettung zweier ungerader Funktionen ist ungerade.

1.20 Trigonometrische Funktionen: In der Mathematik ist es üblich Winkel nicht in Grad sondern in *Radiant*, dem sogenannten *Bogenmaß*, anzugeben. Ein Winkel wird im Bogenmaß beschrieben durch die die Länge des Bogens des Einheitskreises, der durch diesen Winkel gebildet wird, vgl. Abbildung 1.
Dabei wird, wie in der Mathematik allgemein üblich, der Einheitskreis *gegen* den Uhrzeigersinn durchlaufen. Ein Winkel, der im Uhrzeigersinn gemessen werden soll, wird entsprechend durch ein negatives Bogenmaß ausgedrückt.
Da die gesamte Kreislinie des Einheitskreises die Länge 2π hat, ergibt sich damit beispielsweise für den rechten Winkel ein Bogenmaß von $\pi/2$ rad. Hier ist „rad“ als Abkürzung von „Radiant“ die Maßeinheit.
In den meisten mathematischen Anwendungen ist die Winkelmessung in Bogenmaß die einzig sinnvolle Wahl. Sofern nichts anderes gesagt wird, ist bei der Spezifikation von

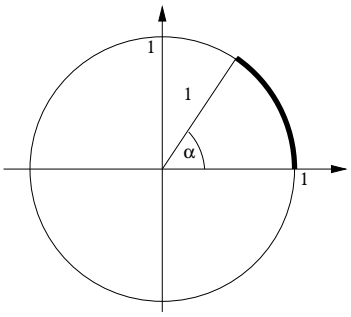


Abbildung 1: Das Bogenmaß eines Winkels

Winkeln also stets davon auszugehen, dass es sich um eine Angabe in Radiant handelt. Insbesondere wird die Maßeinheit rad in der Regel weggelassen. Man schreibt also beispielsweise

$$30^\circ = \pi/6, \quad 45^\circ = \pi/4, \quad 60^\circ = \pi/3, \quad 180^\circ = \pi, \quad 360^\circ = 2\pi.$$

Für $t \in \mathbb{R}$ betrachten wir nun einen Strahl, der vom Ursprung ausgeht und mit der positiven x-Achse einen Winkel mit Bogenmaß t bildet. Den Schnittpunkt dieses Strahls mit dem Einheitskreis bezeichnen wir dann mit P_t, vgl. Abbildung 2.

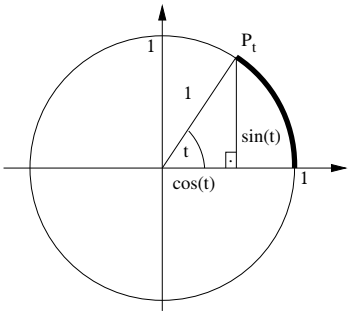


Abbildung 2: Sinus und Cosinus als Koordinaten eines Punktes auf dem Einheitskreis

Man bezeichnet die y-Koordinate des Punktes P_t als *Sinus* von t und die x-Koordinate als *Cosinus* von t und schreibt[1]

$$P_t = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}.$$

Damit sind Sinus und Cosinus reelle Funktionen, die für alle $t \in \mathbb{R}$ definiert sind. Da der Punkt P_t auf dem Einheitskreis liegt, liegen die Werte von Sinus und Cosinus jeweils im Intervall $[-1, 1]$, also

$$D_{\sin} = D_{\cos} = \mathbb{R}, \quad B_{\sin} = B_{\cos} = [-1, 1].$$

Der Quotient von Sinus und Cosinus heißt *Tangens*,

$$\tan t = \frac{\sin t}{\cos t}.$$

[1]Die Klammer um das Argument kann auch weggelassen werden, wenn Verwechslungen ausgeschlossen sind, also z.B. $\sin t$.

Die Nullstellen der Cosinusfunktion führen auf Definitionslücken, der Bildbereich ist die gesamte reelle Achse,

$$D_{\tan} = \mathbb{R} \setminus \{(k + 1/2)\pi : k \in \mathbb{Z}\}, \quad B_{\tan} = \mathbb{R}.$$

Die Graphen von Sinus, Cosinus und Tangens sind in Abbildung 3 dargestellt.

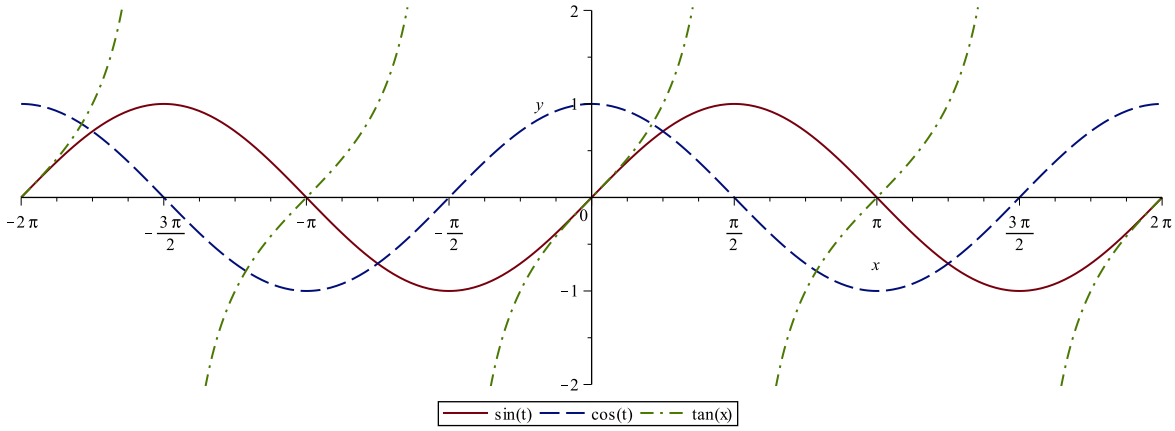


Abbildung 3: Die Graphen von Sinus, Cosinus und Tangens

Die Cotangensfunktion $\cot t = \cos t / \sin t$ findet vergleichsweise selten Anwendung.

1.21 Eigenschaften:

- Die folgenden speziellen Werte sind zu merken:

t	$\sin t$	$\cos t$
0	0	1
$\pi/6$	$1/2$	$\sqrt{3}/2$
$\pi/4$	$\sqrt{2}/2$	$\sqrt{2}/2$
$\pi/3$	$\sqrt{3}/2$	$1/2$
$\pi/2$	1	0

- Sinus und Tangens sind ungerade Funktionen, der Cosinus ist eine gerade Funktion,

$$\sin(t) = -\sin(-t), \quad \cos(-t) = \cos(t).$$

- Addiert man zu einem gegebenen Wert t die Länge 2π des Vollkreises, so stimmen die zugehörigen Punkte überein, also $P_t = P_{t+2\pi}$. Daraus folgt, dass Sinus und Cosinus 2π-periodische Funktionen sind, d.h.,

$$\sin(t) = \sin(t + 2\pi), \quad \cos(t) = \cos(t + 2\pi).$$

- Tangens und Cotangens sind ungerade, π-periodische Funktionen,

$$\tan(t) = \tan(t + \pi), \quad \cot(t) = \cot(t + \pi).$$

- Es gibt darüber hinaus eine Fülle von weiteren Identitäten. Zu merken sind unter anderem:

$$\begin{aligned}
&\sin(t) = \sin(\pi - t) = \cos(\pi/2 - t)\\
&\cos(t) = -\cos(\pi - t) = \sin(\pi/2 + t)\\
&\sin^2(t) + \cos^2(t) = 1\\
&\sin(2t) = 2\sin(t)\cos(t)\\
&\sin(\pi/2 + k\pi) = \cos(k\pi) = (-1)^k, \quad k \in \mathbb{Z}
\end{aligned}$$

- Formeln für $\sin(t/2), \cos(t/2), \sin(s \pm t), \cos(s \pm t), \sin^2 t, \cos^2 t$ usw. können bei Bedarf im Internet oder in Formelsammlungen gefunden werden.

1.22 Arcusfunktionen: Da die trigonometrischen Funktionen nicht injektiv sind, kann man nicht unmittelbar eine Umkehrfunktion angeben. Dazu muss man Definitions- und Zielmengen jeweils geeignet einschränken.

- Für die Sinus-Funktion $f(x) = \sin(x)$ verwendet man $D_f = [-\pi/2, \pi/2]$ und $Z_f = [-1, 1]$. Die nun definierte Umkehrfunktion wird als *Arcussinus* bezeichent und man schreibt dafür

$$\arcsin : [-1, 1] \to [-\pi/2, \pi/2].$$

  Der Arcussinus ordnet also jeder Zahl zwischen -1 und 1 einen Winkel aus dem Intervall $[-\pi/2, \pi/2]$ zu, dessen Sinus dem gegebenen Argument entspricht, vgl. Abbildung 4. Es gilt beispielsweise

$$\arcsin(0) = 0, \quad \arcsin(1/2) = \pi/6, \quad \arcsin(-1) = -\pi/2,$$

  denn

$$\sin(0) = 0, \quad \sin(\pi/6) = 1/2, \quad \sin(-\pi/2) = -1.$$

- Für die Cosinus-Funktion $f(x) = \cos(x)$ verwendet man $D_f = [0, \pi]$ und $Z_f = [-1, 1]$. Die nun definierte Umkehrfunktion wird als *Arcuscosinus* bezeichent und man schreibt dafür

$$\arccos : [-1, 1] \to [0, \pi].$$

  Der Arcuscosinus ordnet also jeder Zahl zwischen -1 und 1 einen Winkel aus dem Intervall $[0, \pi]$ zu, dessen Coinus dem gegebenen Argument entspricht, vgl. Abbildung 4. Es gilt beispielsweise

$$\arccos(1) = 0, \quad \arccos(1/2) = \pi/3, \quad \arccos(-1) = \pi,$$

  denn

$$\cos(0) = 1, \quad \cos(\pi/3) = 1/2, \quad \cos(\pi) = -1.$$

- Für die Tangens-Funktion $f(x) = \tan(x)$ verwendet man $D_f = (-\pi/2, \pi/2)$ und $Z_f = \mathbb{R}$. Die nun definierte Umkehrfunktion wird als *Arcustangens* bezeichent und man schreibt dafür

$$\arctan : \mathbb{R} \to (-\pi/2, \pi/2).$$

Der Arcustangens ordnet also jeder reellen Zahl einen Winkel aus dem Intervall $(-\pi/2, \pi/2)$ zu, dessen Tangens dem gegebenen Argument entspricht, vgl. Abbildung 4. Es gilt beispielsweise

$$\arctan(0) = 0, \quad \arctan(1) = \pi/4, \quad \arctan(-\sqrt{3}) = -\pi/3,$$

denn

$$\tan(0) = 0, \quad \tan(\pi/4) = 1, \quad \tan(-\pi/3) = -\sqrt{3}.$$

Da die trigonometrischen Funktionen periodisch sind, gibt es natürlich zu jedem Wert einer Arcusfunktion jeweils unendlich viele weitere Winkel, die auf dasselbe Ergebnis führen. Beispielswiese ist $\arccos(1) = 0$, aber $\cos(2k\pi) = 1$ für alle $k \in \mathbb{Z}$. Beim Rückschluss von dem Wert einer Winkelfunktion auf den Winkel ist also stets Vorsicht geboten.

1.23 Exponentialfunktion und natürlicher Logarithmus: Für eine beliebige positive Zahl $b > 0$ definiert man die *Exponentialfunktion zur Basis b* durch

$$\exp_b(x) = b^x, \quad D_{\exp} = \mathbb{R}, \; Z_{\exp} = \mathbb{R}_{>0}.$$

Die Exponentialfunktion ist bijektiv und ihre Umkehrfunktion ist der *Logarithmus zur Basis b*,

$$\log_b : \mathbb{R}_{>0} \to \mathbb{R}.$$

Von besonderem Interesse ist die Exponentialfunktion zur Basis e, wobei $e \approx 2,71828$ die sogenannte *Eulersche Zahl* ist. Sie spielt in der Mathematik eine außerordentlich wichtige Rolle. Genauso wie die Kreiszahl π lässt sich e weder durch Brüche noch durch Wurzeln exakt ausdrücken. Auf eine genaue Definition kommen wir später zurück. Wählen wir e als Basis, ergibt sich die *e-Funktion* (oder auch natürliche Exponentialfunktion),

$$\exp(x) = e^x, \quad D_{\exp} = \mathbb{R}, \; Z_{\exp} = \mathbb{R}_{>0}.$$

Die Umkehrfunktion ist der *natürliche Logarithmus*,

$$\ln : \mathbb{R}_{>0} \to \mathbb{R},$$

vgl. Abbildung 4. Es gilt beispielsweise

$$\ln(1) = 0, \quad \ln(e) = 1, \quad \ln(1/e^2) = -2,$$

denn

$$e^0 = 1, \quad e^1 = e, \quad e^{-2} = 1/e^2.$$

1.24 Regeln: Die wichtigsten Rechenregeln für Exponentialfunktion und Logarithmus sind

$e^{a+b} = e^a e^b$	$\ln(ab) = \ln(a) + \ln(b)$
$(e^a)^b = e^{ab}$	$\ln(a^b) = b \ln(a)$
$e^{-a} = 1/e^a$	$\ln(1/a) = -\ln(a)$
$\exp_b a = e^{a \ln(b)}$	$\log_b(a) = \frac{\ln(a)}{\ln(b)}$

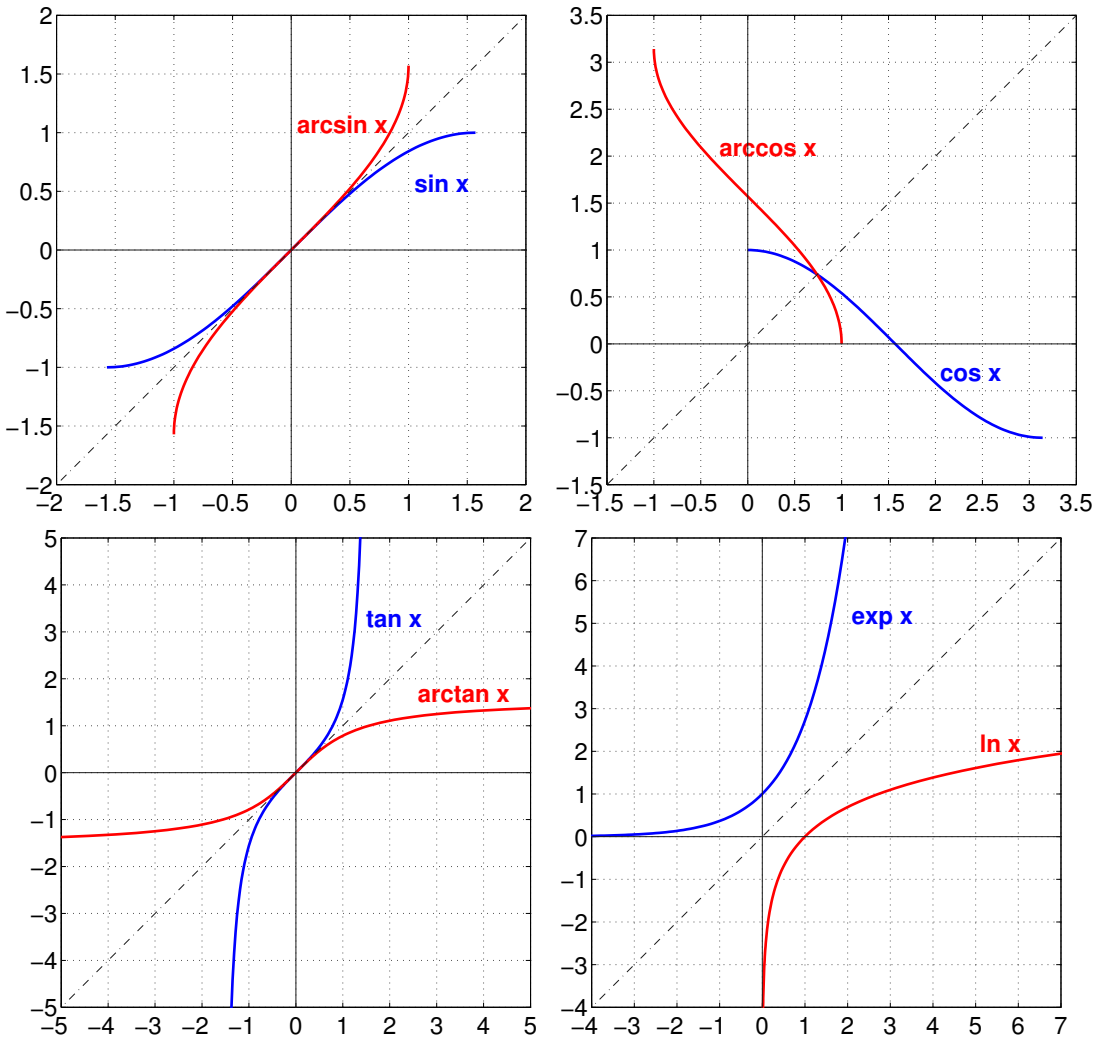


Abbildung 4: Die Graphen der Arcusfunktionen und des natürlichen Logarithmus

1.25 Hyperbelfunktionen: Verschiedentlich sind auch die *Hyperbelfunktionen* von Interesse. Die wichtigsten sind

$$\sinh : \mathbb{R} \to \mathbb{R} \text{ mit } \sinh(x) = \frac{e^x - e^{-x}}{2}, \quad \text{Sinus hyperbolicus,}$$

$$\cosh : \mathbb{R} \to \mathbb{R} \text{ mit } \cosh(x) = \frac{e^x + e^{-x}}{2}, \quad \text{Cosinus hyperbolicus,}$$

$$\tanh : \mathbb{R} \to \mathbb{R} \text{ mit } \tanh(x) = \frac{\sinh(x)}{\cosh(x)}, \quad \text{Tangens hyperbolicus.}$$

Die Graphen dieser Funktionen sind in Abbildung 5 dargestellt. Die zugehörigen Umkehrfunktionen heißen *Areafunktionen* und werden mit arsinh, arcosh, artanh bezeichnet (lies „Areasinus hyperbolicus, etc.“).

1.26 Summen- und Produktzeichen: Für die Summe einer Menge reeller Zahlen $a_1, a_2, \ldots, a_n \in \mathbb{R}$ schreibt man kurz

$$\sum_{j=1}^{n} a_j = a_1 + a_2 + \cdots + a_n.$$

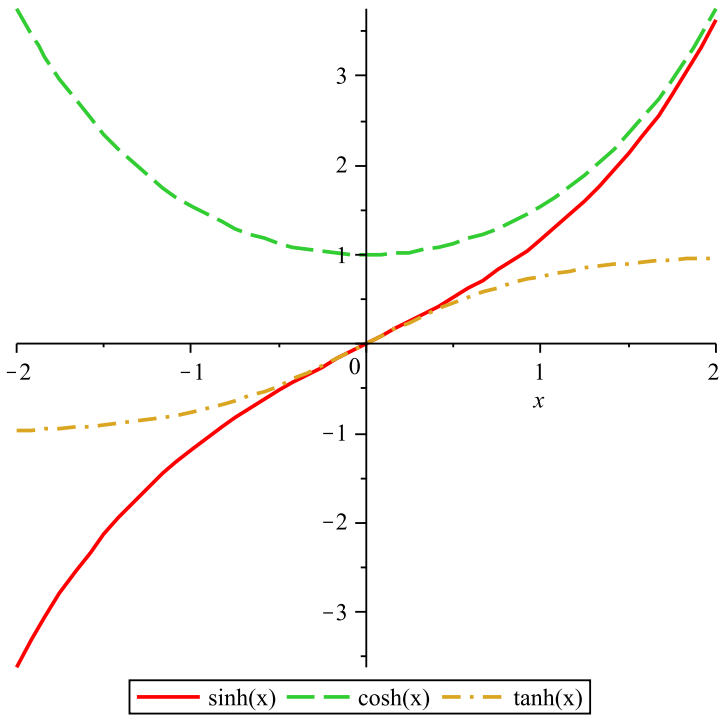


Abbildung 5: Die Graphen der Hyperbelfunktionen sinh, cosh und tanh

Allgemeiner kann die Summenbildung auch bei einem beliebigen Index starten,

$$\sum_{j=m}^{n} a_j = a_m + a_{m+1} + \cdots + a_n.$$

Analog schreibt man für das Produkt

$$\prod_{j=m}^{n} a_j = a_m \cdot a_{m+1} \cdots a_n.$$

Wenn der Startindex m größer ist als der Schlussindex n, dann ist vereinbarungsgemäß der Wert der Summe gleich 0 und der Wert des Produkts gleich 1.

1.27 Beispiel:

- $\displaystyle\sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$
- $\displaystyle\sum_{j=1}^{2} \frac{1}{j} = \frac{1}{1} + \frac{1}{2} = \frac{3}{2}$
- $\displaystyle\sum_{k=5}^{2} \frac{1}{\sqrt{k}} = 0$
- $\displaystyle\prod_{\ell=1}^{5} \ell = 1 \cdot 2 \cdots 5 = 120$

1.28 Rechenregeln für Summen:

- Ausklammern/Ausmultiplizieren:

$$c \sum_{k=m}^{n} a_k = \sum_{k=m}^{n} c a_k$$

- Summe zweier Summen:

$$\sum_{k=m}^{n} a_k + \sum_{k=m}^{n} b_k = \sum_{k=m}^{n} (a_k + b_k)$$

- Indexverschiebung:

$$\sum_{k=m}^{n} a_k = \sum_{k=m-\ell}^{n-\ell} a_{k+\ell}$$

2 Vektorrechnung

2.1 Vektorraum: Eine Menge V nennt man *Vektorraum*, wenn unter anderem[2]

- zwischen je zwei Elementen von V eine Addition sinnvoll definiert ist, d.h.

$$\text{für} \quad v, w \in V \quad \text{ist} \quad v + w \in V,$$

- Elemente von V mit Zahlen multipliziert werden können, d.h.

$$\text{für} \quad v \in V, \alpha \in \mathbb{R} \quad \text{ist} \quad \alpha v \in V,$$

- zwei Distributivgesetze gelten,

$$\alpha(v + w) = \alpha v + \alpha w, \quad (\alpha + \beta)v = \alpha v + \beta v$$

  für alle $\alpha, \beta \in \mathbb{R}$ und alle $v, w \in V$.

Die Elemente eines Vektorraums nennt man *Vektoren*. Beispielsweise bildet die Menge aller Polynome einen Vektorraum, denn Summen und Vielfache von Polynomen sind wieder Polynome; außerdem gelten die genannten Rechenregeln. In diesem Sinne stellen Polynome also einen spezielle Form von Vektoren dar.

2.2 Vektoren in $\mathbb{R}^n$: Das n-fache kartesische Produkt von $\mathbb{R}$ mit sich selbst wird mit $\mathbb{R}^n$ bezeichnet. Speziell erhält man für $n = 1$ die Zahlengerade $\mathbb{R}$, für $n = 2$ die Ebene $\mathbb{R}^2$ und für $n = 3$ den Raum $\mathbb{R}^3$. Die Elemente von $\mathbb{R}^n$ sind also Objekte, die aus n reellen Zahlen bestehen. Wir schreiben sie in der Form

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = [x_1, x_2, \ldots, x_n]^T$$

mit *Koordinaten* $x_i \in \mathbb{R}$. Für zwei beliebige Elemente $\vec{x} = [x_1, \ldots, x_n]^T$ und $\vec{y} = [y_1, \ldots, y_n]^T$ von $\mathbb{R}^n$ definieren wir in naheliegender Weise die Rechenoperationen

- *Addition:*

$$\vec{x} + \vec{y} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

- *Skalarmultiplikation:*

$$\alpha\vec{x} = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}, \quad \alpha \in \mathbb{R},$$

[2]Es fehlen hier streng genommen einige weitere technische Eigenschaften.

Insbesondere ist $1\vec{x} = \vec{x}$, $(-1)\vec{x} = -\vec{x}$ und $0\vec{x} = \vec{0}$. Weitehin rechnet man leicht nach, dass die Distributivgesetzte

$$(\alpha + \beta)\vec{x} = \alpha\vec{x} + \beta\vec{x}, \quad \alpha(\vec{x} + \vec{y}) = \alpha\vec{x} + \alpha\vec{y}, \quad \alpha, \beta \in \mathbb{R},$$

gelten. Somit ist $\mathbb{R}^n$ ein Vektorraum und seine Elemente sind Vektoren. Für Vektoren in $\mathbb{R}^n$ gibt es (mindestens) zwei verschiedene geometrische Interpretationen:

- Wenn ein Koordinatensystem gegeben ist, kann man durch den Vektor $\vec{x}$ die Position eines *Punktes* beschreiben. Die Koordinaten geben dann an, wie weit man ausgehend vom Ursprung in Richtung der Koordinatenachsen jeweils gehen muss, um zu diesem Punkt zu gelangen. Da sich $\vec{x}$ hier auf einen festen Bezugsrahmen bezieht, spricht man auch von einem *gebundenen* Vektor.

- Ebenso kann man den Vektor $\vec{x}$ als Angabe verstehen, die die relative Lage zweier Punkte zueinander beschreibt. Der Vektor charakterisiert also hier die *Differenz* zweier Punkte. Typischerweise veranschaulicht man solch einen *freien* Vektor durch einen Pfeil, der von einem beliebig gewählten Anfangspunkt ausgehend zu einem Endpunkt zeigt. Die Addition zweier Vektoren lässt sich in diesem Modell so veranschaulichen, dass man den Endpunkt des ersten Vektors als Anfangspunkt des zweiten Vektors verwendet.

2.3 Beispiel: Die Unterscheidung zwischen freien und gebundenen Vektoren ist etwas subtil, wird aber im folgenden Beispiel ganz deutlich. Zur Angabe von Zeit*punkten* vewenden wir die Maßeinheit *Uhr*, wir sagen also z.B. *Der Film beginnt um* 20 *Uhr*, wobei sich diese Angabe verinbarungsgemäß auf den Ursprung $Mitternacht = 0\ Uhr$ der Zeitachse bezieht. Wenn der Film um 23 *Uhr* endet, ergibt die Differenz der beiden Zeitpunkte die Dauer des Films, also $23\ Uhr - 20\ Uhr = 3\ Stunden$. Interessanterweise werden Dauern, also Zeit*differenzen*, mit einer anderen Maßeinheit, nämlich *Stunden* gemessen. Wenn wir wissen, dass ein Film 3 *Stunden* dauert, so stellt diese Information einen Zusammenhang zwischen Anfangs- und Endzeitpunkt her, unabhängig davon, wann diese konkret sind. In unserem Beispiel entsprechen also Zeitpunkte den gebunden Vektoren und Dauern den freien Vektoren der Zeit. Zu unterscheiden sind sie an der verwendeten Maßeinheit.

2.4 Norm eines Vektors: Die *euklidische Norm* des Vektors $\vec{x} = [x_1, \ldots, x_n]^T$ ist durch

$$\|\vec{x}\| := \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

definiert. Sie gibt die Länge des Vektors im geometrischen Sinne an und hat folgende Eigenschaften:

- *Positive Definitheit:*

$$\|\vec{x}\| > 0 \quad \text{für} \quad \vec{x} \neq \vec{0}$$
$$\|\vec{x}\| = 0 \quad \text{für} \quad \vec{x} = \vec{0}$$

- *Homogenität:*

$$\|\alpha\vec{x}\| = |\alpha|\,\|\vec{x}\|, \quad \alpha \in \mathbb{R}$$

- *Dreiecksungleichung:*

$$\|\vec{x}+\vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$$
$$\|\vec{x}-\vec{y}\| \geq |\|\vec{x}\| - \|\vec{y}\||$$

2.5 Normierung: Sei $\vec{x} \neq \vec{0}$ ein Vektor, dann erhält man durch die *Normierung*

$$\vec{x}_0 := \frac{\vec{x}}{\|\vec{x}\|}$$

einen Vektor $\vec{x}_0$ der Länge 1, der dieselbe Richtung wie $\vec{x}$ besitzt.

2.6 Skalarprodukt: Das *Skalarprodukt* zweier Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^n$ ist definiert durch

$$\langle \vec{x}, \vec{y} \rangle := x_1 y_1 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i.$$

Die Vektoren $\vec{x}, \vec{y}$ heißen *orthogonal*, wenn

$$\langle \vec{x}, \vec{y} \rangle = 0.$$

Die Vektoren $\vec{x}$ und $\vec{y}$ stehen dann im geometrischen Sinne senkrecht aufeinander. Der Nullvektor ist orthogonal zu allen anderen Vektoren. Skalarprodukt und Norm sind durch die Formeln

$$\|\vec{x}\|^2 = \langle \vec{x}, \vec{x} \rangle$$
$$\langle \vec{x}, \vec{y} \rangle = \frac{1}{4}\big(\|\vec{x}+\vec{y}\|^2 - \|\vec{x}-\vec{y}\|^2\big)$$

verknüpft.

2.7 Eigenschaften:

- *Symmetrie:*

$$\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$$

- *Linearität*

$$\alpha\langle \vec{x}, \vec{y} \rangle = \langle \alpha\vec{x}, \vec{y} \rangle = \langle \vec{x}, \alpha\vec{y} \rangle, \quad \alpha \in \mathbb{R}$$
$$\langle \vec{x}_1 + \vec{x}_2, \vec{y} \rangle = \langle \vec{x}_1, \vec{y} \rangle + \langle \vec{x}_2, \vec{y} \rangle$$
$$\langle \vec{x}, \vec{y}_1 + \vec{y}_2 \rangle = \langle \vec{x}, \vec{y}_1 \rangle + \langle \vec{x}, \vec{y}_2 \rangle$$

- *Binomische Formel:*

$$\langle \vec{x}+\vec{y}, \vec{x}+\vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle + 2\langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{y} \rangle$$

- Sei φ der Winkel zwischen $\vec{x}$ und $\vec{y}$, dann gilt

$$\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\| \cdot \|\vec{y}\| \cos\varphi.$$

- *Cauchy-Schwarz-Ungleichung:*

$$\langle \vec{x}, \vec{y} \rangle \leq \|\vec{x}\| \cdot \|\vec{y}\|$$

2.8 Vektorprodukt: Das *Vektorprodukt* (auch Kreuzprodukt genannt) zweier Vektoren $\vec{x}, \vec{y} \in \mathbb{R}^3$ ist definiert durch

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{bmatrix}.$$

Das Ergebnis $\vec{z} = \vec{x} \times \vec{y}$ ist also wieder ein Vektor in $\mathbb{R}^3$.

2.9 Eigenschaften:

- *Antisymmetrie:*
$$\vec{x} \times \vec{y} = -(\vec{y} \times \vec{x})$$
Daraus folgt insbesondere
$$\vec{x} \times \vec{x} = \vec{0}.$$

- *Linearität:*
$$\begin{aligned} \alpha(\vec{x} \times \vec{y}) = (\alpha\vec{x}) \times \vec{y} = \vec{x} \times (\alpha\vec{y}), \quad \alpha \in \mathbb{R} \\ (\vec{x}_1 + \vec{x}_2) \times \vec{y} = \vec{x}_1 \times \vec{y} + \vec{x}_2 \times \vec{y} \\ \vec{x} \times (\vec{y}_1 + \vec{y}_2) = \vec{x} \times \vec{y}_1 + \vec{x} \times \vec{y}_2 \end{aligned}$$

- *Orthogonalität:*
$$\langle \vec{x} \times \vec{y}, \vec{x} \rangle = \langle \vec{x} \times \vec{y}, \vec{y} \rangle = 0$$

- Sei φ der Winkel zwischen $\vec{x}$ und $\vec{y}$, dann gilt
$$\|\vec{x} \times \vec{y}\| = \|\vec{x}\| \cdot \|\vec{y}\| \cdot |\sin \varphi|.$$
$\vec{x} \times \vec{y}$ ist also ein Vektor, der senkrecht auf $\vec{x}$ und $\vec{y}$ steht und als Länge den Flächeninhalt des von $\vec{x}$ und $\vec{y}$ aufgespannten Parallelogramms besitzt.

- Die drei Vektoren $\vec{x}, \vec{y}, \vec{x} \times \vec{y}$ bilden in dieser Reihenfolge ein *Rechtssystem*. Das heißt anschaulich gesprochen folgendes: Wenn der Daumen und der Zeigefinger der rechten Hand in Richtung $\vec{x}$ und $\vec{y}$ zeigen, dann zeigt der Mittelfinger in Richtung $\vec{x} \times \vec{y}$.

- *Spatprodukt:* Ist $\vec{w}$ noch ein weiterer Vektor im $\mathbb{R}^3$, so beschreibt der Betrag von $\langle \vec{w}, \vec{x} \times \vec{y} \rangle$ das Volumen des durch $\vec{w}, \vec{x}, \vec{y}$ aufgespannten *Parallelepipeds* (oder *Spat*), vgl. Abbildung 6.
Das Vorzeichen von $\langle \vec{w}, \vec{x} \times \vec{y} \rangle$ gibt an, ob $\vec{w}, \vec{x}, \vec{y}$ ein Rechtssystem (positiv) oder ein Linkssystem (negativ) ist. Es gilt:
$$\langle \vec{w}, \vec{x} \times \vec{y} \rangle = \langle \vec{y}, \vec{w} \times \vec{x} \rangle = \langle \vec{x}, \vec{y} \times \vec{w} \rangle.$$

2.10 Geraden in $\mathbb{R}^n$: Seien $\vec{p}$ und $\vec{r}$ Vektoren in $\mathbb{R}^n$ und $\vec{r} \neq \vec{0}$. Die Gleichung
$$g \ : \ \vec{x} = \vec{p} + \lambda\vec{r}, \quad \lambda \in \mathbb{R},$$
beschreibt eine *Gerade in $\mathbb{R}^n$ in parametrisierter Form*. Man bezeichnet $\vec{p}$ als *Aufpunkt*, $\vec{r}$ als *Richtungsvektor* und λ als *Parameter* der Geraden g.

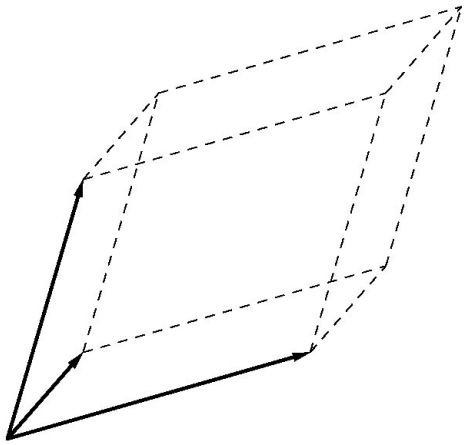


Abbildung 6: Ein Parallelepiped oder Spat

2.11 Beispiel: In Abbildung 7 ist $\vec{p} = [5, 1]^T$ und $\vec{r} = [-2, 1]^T$, also

$$g \ : \ \vec{x} = \begin{bmatrix} 5 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad \lambda \in \mathbb{R}.$$

Für $\lambda = 3$ erhält man den Punkt $\vec{x} = [-1, 4]^T$ und für $\lambda = -1$ den Punkt $\vec{x} = [7, 0]^T$.

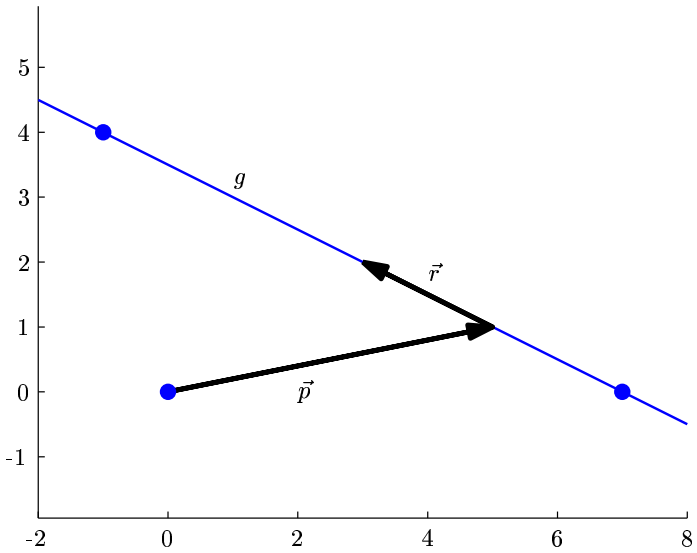


Abbildung 7: Die Gerade mit Aufpunkt $\vec{p}$ und Richtungsvektor $\vec{r}$

2.12 Abstand Punkt-Gerade: Der *Abstand* $d(\vec{q}, g)$ eines Punktes $\vec{q}$ von der Geraden $g : \vec{x} = \vec{p} + \lambda\vec{r}$ ist definiert als

$$d(\vec{q}, g) := \min_{\vec{x} \in g} \|\vec{x} - \vec{q}\|.$$

Dies ist also der kleinste Abstand, den ein Punkt auf der Geraden von $\vec{q}$ haben kann. Der Punkt $\vec{x}^* = \vec{p} + \lambda^* \vec{r}$, für den dieses Minimum angenommen wird, ist dadurch gekennzeichnet, dass der Verbindungsvektor zum Punkt $\vec{q}$ senkrecht zum Richtungsvektor $\vec{r}$ der Geraden ist,

$$\langle \vec{x}^* - \vec{q}, \vec{r} \rangle = \langle \vec{p} - \vec{q} + \lambda^* \vec{r}, \vec{r} \rangle = 0.$$

Löst man diese Gleichung nach λ^* auf, so erhält man

$$\lambda^* = \frac{\langle \vec{q} - \vec{p}, \vec{r} \rangle}{\langle \vec{r}, \vec{r} \rangle}$$

und damit $\vec{x}^*$. Schließlich ist

$$d(\vec{q}, g) = \|\vec{x}^* - \vec{q}\|.$$

2.13 Beispiel [→ 2.11]: Für $g : \vec{x} = [5, 1]^T + \lambda[-2, 1]^T$ und $\vec{q} = [2, 5]^T$ ist

$$\lambda^* = 2 \quad \text{und} \quad \vec{x}^* = \begin{bmatrix} 1 \\ 3 \end{bmatrix},$$

vgl. Abbildung 8.
Damit erhält man

$$d(\vec{q}, g) = \|\vec{x}^* - \vec{q}\| = \sqrt{5}.$$

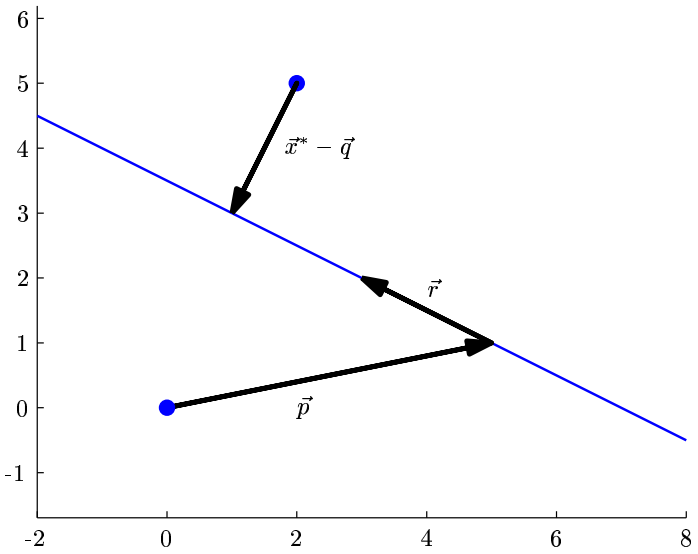


Abbildung 8: Bestimmung des Abstands von $\vec{q}$ zur Geraden g in Beispiel 2.13

2.14 Implizite Form von Geraden in $\mathbb{R}^2$: Sei

$$g \ : \ \vec{x} = \vec{p} + \lambda \vec{r}, \quad \lambda \in \mathbb{R},$$

eine Gerade in $\mathbb{R}^2$ und $\vec{n} \neq \vec{0}$ ein *Normalenvektor*. Dies ist ein Vektor, der senkrecht auf $\vec{r}$ steht, also $\langle \vec{r}, \vec{n} \rangle = 0$. Multipliziert man die Gleichung der Geraden skalar mit $\vec{n}$, dann erhält man die *implizite Form*

$$g \ : \ \langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle.$$

Die Gerade g ist also die Menge aller Punkte $\vec{x} \in \mathbb{R}^2$, die diese Gleichung erfüllen. Auf der linken Seite steht eine Linearkombination der Komponenten von $\vec{x}$, und auf der rechten Seite steht die Konstante $d := \langle \vec{p}, \vec{n} \rangle \in \mathbb{R}$. Einen Normalenvektor $\vec{n}$ erhält man beispielsweise gemäß

$$\vec{r} = \begin{bmatrix} a \\ b \end{bmatrix} \quad \Rightarrow \quad \vec{n} := \begin{bmatrix} b \\ -a \end{bmatrix}.$$

Damit lautet die implizite Form

$$g \ : \ bx - ay = d.$$

2.15 Hessesche Normalform: Normiert man speziell den Normalenvektor $\vec{n}$ auf Länge 1 [→ 2.5], das heißt

$$\vec{n}_0 := \frac{\vec{n}}{\|\vec{n}\|},$$

dann lautet die implizite Form einer Geraden g in $\mathbb{R}^2$

$$g \ : \ \langle \vec{x}, \vec{n}_0 \rangle = d_0, \quad \text{wobei} \quad d_0 := \langle \vec{p}, \vec{n}_0 \rangle = \frac{d}{\|\vec{n}\|}.$$

Diese bezeichnet man als die *Hessesche Normalform* der Geraden g. Sie ist dadurch ausgezeichnet, dass der Betrag der Konstanten d_0 den Abstand der Geraden vom Ursprung angibt, also

$$d(\vec{0}, g) = |d_0|.$$

Der Abstand eines beliebigen Punktes $\vec{q} \in \mathbb{R}^2$ von der Geraden ist durch

$$d(\vec{q}, g) = |d_0 - \langle \vec{q}, \vec{n}_0 \rangle|$$

gegeben.

2.16 Beispiel [→ 2.13]: Sei $g : \vec{x} = [5, 1]^T + \lambda[-2, 1]^T$ und $\vec{q} = [2, 5]^T$. Man erhält

$$\vec{r} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \vec{n} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \Rightarrow \quad d = \langle \vec{p}, \vec{n} \rangle = 7$$

und damit die implizite Form

$$g \ : \ x + 2y = 7.$$

Beispielsweise erfüllen die Punkte $x = 7, y = 0$ und $x = -1, y = 4$ diese Gleichung [→ 2.11]. Die Normierung

$$\|\vec{n}\| = \sqrt{5} \quad \Rightarrow \quad \vec{n}_0 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \Rightarrow \quad d_0 = \langle \vec{p}, \vec{n}_0 \rangle = \frac{7}{\sqrt{5}}$$

liefert die Hessesche Normalform

$$g \ : \ \frac{1}{\sqrt{5}}\, x + \frac{2}{\sqrt{5}}\, y = \frac{7}{\sqrt{5}}.$$

Der Abstand der Geraden vom Ursprung ist also $d_0 = 7/\sqrt{5}$. Der Abstand des Punktes $\vec{q} = [2, 5]^T$ ist wie zuvor

$$d(\vec{q}, g) = |7/\sqrt{5} - \langle [2, 5]^T, [1, 2]^T \rangle / \sqrt{5}| = \sqrt{5}.$$

2.17 Ebenen in $\mathbb{R}^3$: Seien $\vec{p}, \vec{r}_1, \vec{r}_2$ Vektoren in $\mathbb{R}^3$ und $\vec{n} := \vec{r}_1 \times \vec{r}_2 \neq \vec{0}$. Die Gleichung

$$E \ : \ \vec{x} = \vec{p} + \lambda_1 \vec{r}_1 + \lambda_2 \vec{r}_2, \quad \lambda_1, \lambda_2 \in \mathbb{R},$$

beschreibt eine *Ebene in $\mathbb{R}^3$ in parametrisierter Form.* Man bezeichnet $\vec{p}$ als *Aufpunkt*, $\vec{r}_1, \vec{r}_2$ als *Richtungsvektoren*, $\vec{n}$ als *Normalenvektor* und λ_1, λ_2 als *Parameter* der Ebene.

2.18 Abstand Punkt-Ebene: Der *Abstand* $d(\vec{q}, E)$ eines Punktes $\vec{q}$ von der Ebene $\vec{x} = \vec{p} + \lambda_1 \vec{r}_1 + \lambda_2 \vec{r}_2$ ist definiert als

$$d(\vec{q}, E) := \min_{\vec{x} \in E} \|\vec{x} - \vec{q}\|.$$

Dies ist also der kleinste Abstand, den ein Punkt auf der Ebene von $\vec{q}$ haben kann. Der Punkt $\vec{x}^* = \vec{p} + \lambda_1^* \vec{r}_1 + \lambda_2^* \vec{r}_2$, für den dieses Minimum angenommen wird, ist dadurch gekennzeichnet, dass der Verbindungsvektor zum Punkt $\vec{q}$ senkrecht zu beiden Richtungsvektoren $\vec{r}_1, \vec{r}_2$ der Ebene ist, d.h.,

$$\vec{x}^* - \vec{q} = \mu \vec{n}.$$

Multipliziert man diese Gleichung skalar mit $\vec{n}$, dann erhält man

$$\langle \vec{p} - \vec{q}, \vec{n} \rangle = \mu \langle \vec{n}, \vec{n} \rangle \quad \Rightarrow \quad \mu = \frac{\langle \vec{p} - \vec{q}, \vec{n} \rangle}{\|\vec{n}\|^2},$$

da $\langle \vec{r}_1, \vec{n} \rangle = \langle \vec{r}_2, \vec{n} \rangle = 0$. Der Abstand ist also

$$d(\vec{q}, E) = \|\mu \vec{n}\| = |\mu| \cdot \|\vec{n}\| = \frac{|\langle \vec{p} - \vec{q}, \vec{n} \rangle|}{\|\vec{n}\|}.$$

2.19 Beispiel: Sei $\vec{p} = [1, 1, 5]^T$, $\vec{r}_1 = [3, 0, 1]^T$ und $\vec{r}_2 = [1, 2, -1]^T$, also

$$E \ : \ \vec{x} = \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix} + \lambda_1 \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R}.$$

Dann ist der Normalenvektor gegeben durch

$$\vec{n} = \vec{r}_1 \times \vec{r}_2 = \begin{bmatrix} -2 \\ 4 \\ 6 \end{bmatrix} \quad \Rightarrow \quad \|\vec{n}\| = \sqrt{56}.$$

Der Abstand des Punktes $\vec{q} = [1, 0, 7]^T$ von der Ebene ist

$$d(\vec{q}, E) = \frac{|-8|}{\sqrt{56}} = \frac{4}{\sqrt{14}}.$$

Dabei ist $\mu = -1/7$ und $\vec{x}^* = [9/7, -4/7, 43/7]^T$.

2.20 Implizite Form von Ebenen in $\mathbb{R}^3$: Sei

$$E \ : \ \vec{x} = \vec{p} + \lambda_1 \vec{r}_1 + \lambda_2 \vec{r}_2, \quad \lambda_1, \lambda_2 \in \mathbb{R},$$

eine Ebene in $\mathbb{R}^3$ mit Normalenvektor $\vec{n}$. Multipliziert man die Gleichung der Ebene skalar mit $\vec{n}$, dann erhält man die *implizite Form*

$$E \ : \ \langle \vec{x}, \vec{n} \rangle = \langle \vec{p}, \vec{n} \rangle.$$

Die Ebene E ist also die Menge aller Punkte $\vec{x} \in \mathbb{R}^3$, die diese Gleichung erfüllen. Auf der linken Seite steht eine Linearkombination der Komponenten von $\vec{x}$, und auf der rechten Seite steht die Konstante $d := \langle \vec{p}, \vec{n} \rangle \in \mathbb{R}$.

2.21 Hessesche Normalform: Normiert man speziell den Normalenvektor $\vec{n}$ auf Länge 1, d.h.,

$$\vec{n}_0 := \frac{\vec{n}}{\|\vec{n}\|},$$

dann lautet die implizite Form

$$E \;:\; \langle \vec{x}, \vec{n}_0 \rangle = d_0, \quad \text{wobei} \quad d_0 := \langle \vec{p}, \vec{n}_0 \rangle = \frac{d}{\|\vec{n}\|}.$$

Diese bezeichnet man als die *Hessesche Normalform* der Ebene E. Sie ist dadurch ausgezeichnet, dass der Betrag der Konstanten d_0 den Abstand der Ebene vom Ursprung angibt, also

$$d(\vec{0}, E) = |d_0|.$$

Der Abstand eines beliebigen Punktes $\vec{q} \in \mathbb{R}^3$ von der Ebene ist durch

$$d(\vec{q}, E) = |d_0 - \langle \vec{q}, \vec{n}_0 \rangle|$$

gegeben.

2.22 Beispiel [→ 2.19]: Mit $\vec{p} = [1, 1, 5]^T$, $\vec{n} = [-2, 4, 6]^T$ und $\vec{x} = [x, y, z]^T$ erhält man $d = \langle \vec{p}, \vec{n} \rangle = 32$ und damit die implizite Form

$$E \;:\; -2x + 4y + 6z = 32.$$

Die Normierung

$$\|\vec{n}\| = \sqrt{56} = 2\sqrt{14} \quad \Rightarrow \quad \vec{n}_0 = \frac{1}{\sqrt{14}} \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix} \quad \Rightarrow \quad d_0 = \langle \vec{p}, \vec{n}_0 \rangle = \frac{16}{\sqrt{14}}$$

liefert die Hessesche Normalform

$$E \;:\; \frac{-1}{\sqrt{14}}\, x + \frac{2}{\sqrt{14}}\, y + \frac{3}{\sqrt{14}}\, z = \frac{16}{\sqrt{14}}.$$

Der Abstand der Ebene vom Ursprung ist also $d_0 = 16/\sqrt{14}$, und der Abstand des Punktes $\vec{q} = [1, 0, 7]^T$ von der Ebene ist wie zuvor

$$d(\vec{q}, E) = \left| 16/\sqrt{14} - 20/\sqrt{14} \right| = 4/\sqrt{14}.$$

2.23 Schnitt Ebene-Gerade: Zur Berechnung des Schnittpunkts $\vec{x}^*$ einer Ebene E mit einer Geraden g in $\mathbb{R}^3$ verwendet man zweckmäßigerweise für die Ebene die implizite und für die Gerade die parametrische Form,

$$\begin{aligned} E &: \langle \vec{x}, \vec{n} \rangle = d \\ g &: \vec{x} = \vec{p} + \lambda \vec{r}, \quad \lambda \in \mathbb{R}. \end{aligned}$$

Setzt man die Geradengleichung in die Ebenengleichung ein, so erhält man die Bedingung

$$\langle \vec{p}, \vec{n} \rangle + \lambda^* \langle \vec{r}, \vec{n} \rangle = d \tag{2.1}$$

für den Parameter λ^* des Schnittpunkts. Nun sind folgende Fälle zu unterscheiden:

- Falls $\langle \vec{r}, \vec{n} \rangle \neq 0$, dann ist der Schnittpunkt eindeutig bestimmt und durch

$$\vec{x}^* = \vec{p} + \frac{d - \langle \vec{p}, \vec{n} \rangle}{\langle \vec{r}, \vec{n} \rangle} \vec{r}$$

  gegeben.

- Falls $\langle \vec{r}, \vec{n} \rangle = 0$ und $\langle \vec{p}, \vec{n} \rangle = d$, dann ist Gleichung (2.1) für alle $\lambda^* \in \mathbb{R}$ erfüllt; es gibt also unendlich viele Lösungen. Dies bedeutet, dass die Gerade parallel zur Ebene ist und vollständig in dieser liegt.

- Falls $\langle \vec{r}, \vec{n} \rangle = 0$ und $\langle \vec{p}, \vec{n} \rangle \neq d$, dann ist Gleichung (2.1) für kein $\lambda^* \in \mathbb{R}$ erfüllt; es gibt also keinen Schnittpunkt. Dies bedeutet, dass die Gerade parallel zur Ebene ist und nicht in dieser liegt.

2.24 Lineare Teilräume: Eine nichtleere Teilmenge $L \subset \mathbb{R}^n$ heißt *linearer Teilraum*, wenn für $\vec{x}, \vec{y} \in L$ und $\alpha \in \mathbb{R}$ stets auch

$$\vec{x} + \vec{y} \in L, \qquad \alpha \vec{x} \in L.$$

Ein linearer Teilraum muss also mit zwei Vektoren stets deren Summe und mit jedem Vektor dessen Vielfache enthalten. Wegen $L \neq \emptyset$ ist stets $\vec{0} = 0\vec{x} \in L$.

2.25 Beispiel:

- $\mathbb{R}^n$ und $\{\vec{0}\}$ sind lineare Teilräume von $\mathbb{R}^n$.

- $L_1 = \{[x_1, x_2]^T : x_1 + 3x_2 = 0\}$ ist ein linearer Teilraum von $\mathbb{R}^2$.

- $L_2 = \{[x_1, x_2]^T : x_2 \geq 0\}$ ist kein linearer Teilraum von $\mathbb{R}^2$, da zum Beispiel $[-1, 2]^T \in L$, aber $(-1) \cdot [-1, 2]^T = [1, -2]^T \notin L_2$.

- Die linearen Teilräume der Ebene $\mathbb{R}^2$ sind die Ursprungsgeraden sowie $\mathbb{R}^2$ und $\{\vec{0}\}$.

- Die linearen Teilräume von $\mathbb{R}^3$ sind die Ursprungsgeraden, die Ursprungsebenen sowie $\mathbb{R}^3$ und $\{\vec{0}\}$.

2.26 Linearkombination und lineare Hülle: Seien $\vec{p}_1, \ldots, \vec{p}_m$ Vektoren in $\mathbb{R}^n$ und $\lambda_1, \ldots, \lambda_m$ reelle Zahlen. Dann nennt man den Vektor

$$\vec{x} = \lambda_1 \vec{p}_1 + \cdots + \lambda_m \vec{p}_m = \sum_{i=1}^{m} \lambda_i \vec{p}_i$$

eine *Linearkombination* der Vektoren $\vec{p}_i$ mit Koeffizienten λ_i. Betrachtet man die Menge aller möglichen Linearkombinationen von $\vec{p}_1, \ldots, \vec{p}_m$, so erhält man einen linearen Teilraum von $\mathbb{R}^n$. Dieser wird *lineare Hülle* der Vektoren genannt und mit $\mathrm{Lin}(\vec{p}_1, \ldots, \vec{p}_m)$ bezeichnet:

$$\mathrm{Lin}(\vec{p}_1, \ldots, \vec{p}_m) = \{\lambda_1 \vec{p}_1 + \cdots + \lambda_m \vec{p}_m : \lambda_1, \ldots, \lambda_m \in \mathbb{R}\}.$$

Ist $\vec{p}_1 \neq \vec{0}$, so ist $\mathrm{Lin}(\vec{p}_1)$ die Gerade in Richtung $\vec{p}_1$ durch den Ursprung. Sind $\vec{p}_1, \vec{p}_2$ keine Vielfachen voneinander, insbesondere also keiner $= \vec{0}$, so ist $\mathrm{Lin}(\vec{p}_1, \vec{p}_2)$ die durch $\vec{p}_1, \vec{p}_2$ aufgespannte Ebene durch den Ursprung.

2.27 Lineare Unabhängigkeit: Die Vekoren $\vec{p}_1, \ldots, \vec{p}_m \in \mathbb{R}^n$ heißen *linear unabhängig*, wenn man den Nullvektor nur dann als Linearkombination erhält, wenn alle Koeffizienten Null sind. Es muss also gelten:

$$\sum_{i=1}^{m} \lambda_i \vec{p}_i = \vec{0} \quad \Rightarrow \quad \lambda_1 = \cdots = \lambda_m = 0.$$

Kann der Nullvektor auch anders gebildet werden, also mit mindestens einem $\lambda_i \neq 0$, so heißen die Vektoren $\vec{p}_1, \ldots, \vec{p}_m$ *linear abhängig*.

2.28 Beispiel:

- Die *Einheitsvektoren*

$$\vec{e}_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_3 := \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ \ldots \ , \vec{e}_n := \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

  sind linear unabhängig, denn aus

$$\sum_{i=1}^{n} \lambda_i \vec{e}_i = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} = \vec{0}$$

  folgt $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

- Ebenfalls linear unabhängig sind

$$\vec{p}_1 := \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{p}_2 := \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

  Dagegen sind

$$\vec{q}_1 := \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{q}_2 := \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

  linear abhängig, denn $2\vec{q}_1 + \vec{q}_2 = \vec{0}$.

2.29 Basis und Dimension: Sei $L \subset \mathbb{R}^n$ ein linearer Teilraum und $\vec{p}_1, \ldots, \vec{p}_m \in L$. Wenn $\vec{p}_1, \ldots, \vec{p}_m$ linear unabhängig sind und $L = \mathrm{Lin}(\vec{p}_1, \ldots, \vec{p}_m)$ gilt, so bilden $\vec{p}_1, \ldots, \vec{p}_m$ eine *Basis von L*. Die Zahl m der linear unabhängigen Vektoren heißt *Dimension von L*, wir schreiben dafür kurz $\dim L = m$. Die Vektoren $\vec{p}_1, \ldots, \vec{p}_m$ können wir zu einem neuen Objekt zusammengefasst werden, das wir mit dem entsprechenden Großbuchstaben bezeichnen; wir schreiben

$$P = [\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_m].$$

2.30 Beispiel [→ 2.28]:

- Die Einheitsvektoren $\vec{e}_1, \ldots, \vec{e}_n$ bilden eine Basis des $\mathbb{R}^n$. Insbesondere ist
$$\dim \mathbb{R}^n = n.$$
Dagegen ist $\vec{e}_2, \ldots, \vec{e}_n$ keine Basis des $\mathbb{R}^n$, denn $\vec{e}_1 \notin \mathrm{Lin}(\vec{e}_2, \ldots, \vec{e}_n)$.
- $P = [\vec{p}_1, \vec{p}_2]$ ist eine Basis von $\mathbb{R}^2$, aber $Q = [\vec{q}_1, \vec{q}_2]$ nicht.
- Ein Sonderfall: Der Nullraum $L = \{\vec{0}\}$, der nur den Nullvektor enthält, hat Dimension $\dim L = 0$.

2.31 Eigenschaften: Ist L ein linearer Teilrraum von $\mathbb{R}^n$, so gilt:

- Es gibt eine Basis von L.
- Zwei verschiedene Basen von L bestehen immer aus gleich vielen Vektoren.
- $\dim L \leq n$.

2.32 Koordinaten bezüglich einer Basis: Sei $P = [\vec{p}_1, \ldots, \vec{p}_n]$ eine Basis von $\mathbb{R}^n$, dann gibt es zu jedem $\vec{x} \in \mathbb{R}^n$ eindeutig bestimmte reelle Zahlen $\alpha_1, \ldots, \alpha_n$, genannt *$P$-Koordinaten von $\vec{x}$* oder auch *Koordinaten von $\vec{x}$ bezüglich P*, so dass
$$\vec{x} = \alpha_1 \vec{p}_1 + \cdots + \alpha_n \vec{p}_n.$$
Der Vektor $\vec{x}_P = [\alpha_1, \ldots, \alpha_n]^T$ heißt *P-Koordinatenvektor von $\vec{x}$*.

2.33 Beispiel [→ 2.28]:

- Die Koordinaten von $\vec{x} = [x_1, x_2, \ldots, x_n]^T$ bzgl. der der Basis $E = [\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n]$ des $\mathbb{R}^n$ sind $x_1, x_2, \ldots, x_n$, und der zugehörige E-Koordinatenvektor ist $\vec{x}_E = \mathbf{x} = [x_1, x_2, \ldots, x_n]^T$.
- Die Koordinaten von $\vec{x} = [-1, 5]^T$ bzgl. der Basis $P = [\vec{p}_1, \vec{p}_2]$ von $\mathbb{R}^2$ sind $\alpha_1 = 3$ und $\alpha_2 = -2$, denn
$$\vec{x} = \begin{bmatrix} -1 \\ 5 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 3\vec{x}_1 - 2\vec{x}_2.$$
Der P-Koordinatenvektor von $\vec{x}$ ist also $\vec{x}_P = [3, -2]^T$.
- Q-Koordinaten machen keine Sinn, da die Vektoren $\vec{q}_1, \vec{q}_2$ keine Basis bilden.

2.34 Orthogonalbasis und Orthonormalbasis: Die Basis $P = [\vec{p}_1, \ldots, \vec{p}_n]$ von $\mathbb{R}^n$ heißt

- *Orthogonalbasis (OB)*, falls $\langle \vec{p}_i, \vec{p}_j \rangle = 0$ für alle $i \neq j$.
- *Orthonormalbasis (ONB)*, falls $\langle \vec{p}_i, \vec{p}_j \rangle = 0$ für alle $i \neq j$ gilt und zusätzlich $\langle \vec{p}_i, \vec{p}_i \rangle = 1$ für alle i. Diese Bedingungen kann man auch kurz wie folgt schreiben:
$$\langle \vec{p}_i, \vec{p}_j \rangle = \delta_{ij} = \begin{cases} 0 & : \text{ wenn } i \neq j, \\ 1 & : \text{ wenn } i = j. \end{cases}$$
Der Ausdruck δ_{ij} wird *Kronecker–Delta* genannt.

  Jede OB kann durch Normierung [→ 2.5] in eine ONB umgewandelt werden.

Orthogonal- und Orthonormalbasen sind unter anderem deshalb besonders wichtig, weil sich hier die zugehörigen Koordinaten sehr einfach berechnen lassen. Wenn P eine OB ist, dann sind die P-Koordinaten $\alpha_1, \ldots, \alpha_n$ des Vektors $\vec{x}$ gegeben durch

$$\alpha_j = \frac{\langle \vec{x}, \vec{p}_j \rangle}{\langle \vec{p}_j, \vec{p}_j \rangle}, \quad j = 1, \ldots, n.$$

Wenn P sogar eine ONB ist, vereinfacht sich dieser Ausdruck weiter zu

$$\alpha_j = \langle \vec{x}, \vec{p}_j \rangle, \quad j = 1, \ldots, n.$$

Es gilt also

$$\vec{x} = \sum_{j=1}^{n} \langle \vec{x}, \vec{p}_j \rangle \vec{p}_j.$$

2.35 Beispiel [→ 2.28]:

- Die Basis $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ des $\mathbb{R}^n$ ist eine Orthonormalbasis.
- Die Basis $\vec{p}_1, \vec{p}_2$ des $\mathbb{R}^2$ ist keine Orthogonalbasis, da $\langle \vec{p}_1, \vec{p}_2 \rangle \neq 0$.
- Die Vektoren $\vec{u}_1 = [1, 1, 1]^T, \vec{u}_2 = [1, -1, 0]^T, \vec{u}_3 = [1, 1, -2]^T$ bilden eine OB des $\mathbb{R}^3$. Sie bilden aber keine ONB, da z.B. $\langle \vec{u}_1, \vec{u}_1 \rangle = 3$. Die U-Koordinaten des Vektors $\vec{x} = [2, 0, -1]^T$ sind gegeben durch $\vec{x}_U = [2, 0, -1]^T$.
- Durch Normierung der Vektoren $\vec{u}_1, \vec{u}_2, \vec{u}_3$ erhalten wir die ONB

$$\vec{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}.$$

  Die V-Koordinaten des Vektors $\vec{x} = [2, 0, -1]^T$ sind $\vec{x}_V = [2/\sqrt{3}, 0, -1/\sqrt{6}]^T$.

3 Komplexe Zahlen

3.1 Komplexe Multiplikation: Für zwei Vektoren

$$\vec{z}_1 = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}, \quad \vec{z}_2 = \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$$

in $\mathbb{R}^2$ wird neben der üblichen Addition die *komplexe Multiplikation*

$$\vec{z}_1 * \vec{z}_2 := \begin{bmatrix} a_1a_2 - b_1b_2 \\ a_1b_2 + b_1a_2 \end{bmatrix}$$

definiert. Sie hat folgende Eigenschaften:

- *Kommutativgesetz:*
$$\vec{z}_1 * \vec{z}_2 = \vec{z}_2 * \vec{z}_1$$
- *Assoziativgesetz:*
$$\vec{z}_1 * (\vec{z}_2 * \vec{z}_3) = (\vec{z}_1 * \vec{z}_2) * \vec{z}_3$$
- *Distributivgesetz:*
$$\vec{z}_1 * (\vec{z}_2 + \vec{z}_3) = \vec{z}_1 * \vec{z}_2 + \vec{z}_1 * \vec{z}_3$$

Der Raum $\mathbb{R}^2$ versehen mit der komplexen Multiplikation wird *komplexe Zahlenebene* genannt und mit $\mathbb{C}$ bezeichnet. Die Elemente von $\mathbb{C}$ heißen *komplexe Zahlen*.

3.2 Beispiel:

$$\vec{z}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \vec{z}_2 = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \quad \Rightarrow \quad \vec{z}_1 * \vec{z}_2 = \vec{z}_2 * \vec{z}_1 = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$$

3.3 Einheiten: Die Einheitsvektoren $\vec{e}_1$ und $\vec{e}_2$ in der komplexen Zahlenebene werden mit

$$\vec{1} := \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{und} \quad \vec{i} := \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

bezeichnet und *reelle Einheit* bzw. *imaginäre Einheit* genannt. Es gilt also

$$\vec{z} = \begin{bmatrix} a \\ b \end{bmatrix} = a\vec{1} + b\vec{i}.$$

Man bezeichnet a als *Realteil* und b als *Imaginärteil* der komplexen Zahl $\vec{z}$ und schreibt dafür

$$a = \operatorname{Re} \vec{z}, \quad b = \operatorname{Im} \vec{z}.$$

Für eine beliebige komplexe Zahl $\vec{z} = [a, b]^T$ gilt

$$\vec{1} * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \vec{i} * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -b \\ a \end{bmatrix}.$$

$\vec{1}$ ist also das *neutrale Element* der komplexen Multiplikation, während Multiplikation mit $\vec{i}$ den gegebenen Vektor $\vec{z}$ um den Winkel $\pi/2$ gegen den Uhrzeigersinn dreht. Insbesondere gilt

$$\vec{i} * \vec{i} = -\vec{1}.$$

Damit erhält die Multiplikationsregel nach dem Distributivgesetz die Form

$$\begin{aligned}(a_1\vec{1} + b_1\vec{i}) * (a_2\vec{1} + b_2\vec{i}) &= a_1a_2(\vec{1} * \vec{1}) + a_1b_2(\vec{1} * \vec{i}) + b_1a_2(\vec{i} * \vec{1}) + b_1b_2(\vec{i} * \vec{i}) \\ &= (a_1a_2 - b_1b_2)\vec{1} + (a_1b_2 + b_1a_2)\vec{i}.\end{aligned}$$

3.4 Notation: Beim Rechnen mit komplexen Zahlen ist es üblich, die Vektorpfeile wegzulassen. Man schreibt also

$$z = a1 + bi \quad \text{für} \quad \vec{z} = a\vec{1} + b\vec{i}.$$

Weiterhin wird die Notation der reellen Einheit weggelassen. Man schreibt also

$$z = a + bi \quad \text{für} \quad z = a1 + bi.$$

Schließlich wird auch der Mal-Punkt nicht mit einem speziellen Symbol notiert. Man schreibt also

$$z_1 z_2 \text{ oder } z_1 \cdot z_2 \quad \text{für} \quad z_1 * z_2.$$

Die Rechenregeln lauten nun

$$\begin{aligned}(a_1 + b_1 i) + (a_2 + b_2 i) &= (a_1 + a_2) + (b_1 + b_2)i \\ (a_1 + b_1 i) \cdot (a_2 + b_2 i) &= (a_1 a_2 - b_1 b_2) + (a_1 b_2 + b_1 a_2)i.\end{aligned}$$

Das Rechnen mit komplexen Zahlen folgt also denselben Gesetzen wie das Rechnen mit reellen Zahlen. Es ist lediglich die Regel

$$i \cdot i = i^2 = -1$$

zu beachten.

3.5 Beispiel [→ 3.2]:

$$(2 + i) + (4 + 3i) = 6 + 4i, \quad (2 + i) \cdot (4 + 3i) = 5 + 10i$$

3.6 Polarkoordinaten: Der Punkt $z = a + bi$ kann entweder durch seine kartesischen Koordinaten (a, b) oder durch seine *Polarkoordinaten* (r, φ) definiert werden. Dabei ist r der Abstand vom Ursprung und φ der Winkel zur reellen Einheit 1,

$$\begin{bmatrix} a \\ b \end{bmatrix} = r \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix} \quad \text{bzw.} \quad a + bi = r(\cos\varphi + i \sin\varphi).$$

Man nennt r den *Betrag* und φ das *Argument* von z und schreibt

$$r = |z|, \quad \varphi = \arg z.$$

Das Argument φ wird im *mathematisch psoitiven Sinn*, also gegen den Uhrzeigersinn, gemessen. Weiterhin ist zu beachten, dass φ nur bis auf Vielfache von 2π bestimmt ist. Typischerweise wählt man $\varphi \in [0, 2\pi)$ oder $\varphi \in (-\pi, \pi]$.

- Umrechnung von Polarkoordinaten (r, φ) in kartesische Koordinaten (a, b):

$$a = r\cos\varphi, \qquad b = r\sin\varphi.$$

- Umrechnung von kartesischen Koordinaten (a, b) in Polarkoordinaten (r, φ):

$$r = \sqrt{a^2 + b^2}, \qquad \varphi = \begin{cases} \arctan(b/a), & a > 0, \\ \arctan(b/a) + \pi, & a < 0 \text{ und } b \geq 0, \\ \arctan(b/a) - \pi, & \text{falls} \quad a < 0 \text{ und } b < 0, \\ \pi/2, & a = 0 \text{ und } b > 0, \\ -\pi/2, & a = 0 \text{ und } b < 0. \end{cases}$$

3.7 Beispiel: Für $z = 1 - \sqrt{3}i$ ist

$$|z| = r = \sqrt{1^2 + (-\sqrt{3})^2} = 2$$
$$\cos\varphi = a/r = 1/2,\ \sin\varphi = b/r = -\sqrt{3}/2 \ \Rightarrow\ \arg z = \varphi = 5\pi/3.$$

Es gilt also

$$1 - i\sqrt{3} = 2(\cos 5\pi/3 + i \sin 5\pi/3).$$

Man könnte anstelle von $\arg z = 5\pi/3$ auch $\arg z = -\pi/3$ oder $\arg z = 11\pi/3$ wählen.

3.8 Konjugation: Die Spiegelung einer komplexen Zahl an der reellen Achse bezeichnet man als *Konjugation* und schreibt dafür

$$\overline{z} = \overline{a + bi} = a - bi.$$

$\overline{z}$ wird die zu z *konjugiert komplexe Zahl* genannt. Es gilt

- $$\overline{\overline{z}} = z$$
- $$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$$
- $$z \cdot \overline{z} = (a + bi) \cdot (a - bi) = a^2 + b^2 = |z|^2$$
- $$\operatorname{Re} z = \frac{z + \overline{z}}{2}, \quad \operatorname{Im} z = \frac{z - \overline{z}}{2i}$$

3.9 Division: Man berechnet den Quotienten zweier komplexer Zahlen, indem man mit dem konjugiert Komplexen des Nenners erweitert,

$$\frac{z_1}{z_2} = \frac{z_1 \overline{z_2}}{|z_2|^2}, \quad z_2 \neq 0.$$

Damit ist der Nenner reell und die Division problemlos möglich. Division durch Null ist wie üblich ausgeschlossen.

3.10 Beispiel:

$$\frac{4 + i}{2 - 3i} = \frac{(4 + i)(2 + 3i)}{(2 - 3i)(2 + 3i)} = \frac{5 + 14i}{13} = \frac{5}{13} + \frac{14}{13}\, i$$

3.11 Geometrische Deutung der Multiplikation: Gegeben seien zwei komplexe Zahlen mit Polarkoordinaten (r_1, φ_1) und (r_2, φ_2), also

$$z_1 = r_1(\cos\varphi_1 + i \sin\varphi_1)$$
$$z_2 = r_2(\cos\varphi_2 + i \sin\varphi_2)$$

Berechnet man das Produkt, so erhält man

$$z_1 z_2 = r_1 r_2\big((\cos\varphi_1 \cos\varphi_2 - \sin\varphi_1 \sin\varphi_2) + i(\cos\varphi_1 \sin\varphi_2 - \sin\varphi_1 \cos\varphi_2)\big)$$

Gemäß der Additionstheoreme für Winkelfunktionen lässt sich dies einfacher schreiben als

$$z_1 z_2 = r_1 r_2 \big(\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)\big).$$

Die Multiplikation zweier komplexer Zahlen entspricht also einer *Multiplikation der Beträge* und einer *Addition der Argumente*. Genauso entspricht die Division zweier komplexer Zahlen einer *Division der Beträge* und einer *Subtraktion der Argumente*,

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} \big(\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)\big).$$

3.12 Komplexe e-Funktion: Man definiert für rein imaginäre Argumente die e-Funktion durch

$$e^{i\varphi} := \cos\varphi + i \sin\varphi, \quad \varphi \in \mathbb{R}.$$

Dieser Zusammenhang wird auch *Eulersche Formel* genannt. Es gilt $|e^{i\varphi}| = 1$. Das heißt, $e^{i\varphi}$ ist eine Zahl auf dem komplexen Einheitskreis, die durch den Winkel φ bestimmt ist. Damit hat eine komplexe Zahl mit Polarkoordinaten (r, φ) die Darstellung

$$z = re^{i\varphi}$$

und die Multiplikation bekommt die einfache Form

$$z_1 \cdot z_2 = \big(r_1 e^{i\varphi_1}\big) \cdot \big(r_2 e^{i\varphi_2}\big) = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}.$$

Definiert man die e-Funktion für beliebige komplexe Argumente $z = a + ib$ durch

$$e^{a+ib} := e^a \cdot e^{ib} = e^a(\cos b + i \sin b),$$

dann gilt allgemein

$$e^{z_1} \cdot e^{z_2} = e^{z_1 + z_2}, \quad e^{z_1}/e^{z_2} = e^{z_1 - z_2}.$$

Die Bildmenge umfasst alle komplexen Zahlen mit Ausnahme der 0.

3.13 Beispiel:

$$e^0 = e^{2\pi i} = 1, \quad e^{i\pi/2} = i, \quad e^{i\pi} = -1, \quad e^{-i\pi/2} = -i, \quad e^{2+i\pi} = -e^2$$

3.14 Komplexer Logarithmus: Die Umkehrfunktion der e-Funktion wird als *natürlicher Logarithmus* bezeichnet. Sei $w \neq 0$ gegeben, dann muss für $\ln w = z = a + bi$ gelten

$$e^z = e^a e^{ib} = w.$$

Hieraus folgt

$$|w| = e^a, \quad \arg w = b$$

und

$$\ln w = \ln |w| + i \arg w.$$

Es gelten die bekannten Regeln

$$\ln(w_1 \cdot w_2) = \ln w_1 + \ln w_2, \quad \ln(w_1/w_2) = \ln w_1 - \ln w_2.$$

Man beachte, dass der Logarithmus der Zahl 0 nicht definiert ist und dass sich die Mehrdeutigkeit des Arguments auf die Logarithmus-Funktion überträgt. Eindeutigkeit erhält man, indem man wieder $\arg w \in [0, 2\pi)$ fordert.

3.15 Beispiel:

$$\ln(-1) = i\pi, \quad \ln(-e) = 1 + i\pi, \quad \ln(i) = i\pi/2, \quad \ln(3+4i) = \ln 5 + i \arctan(4/3)$$

3.16 Komplexe Potenz-Funktion: Für komplexe Zahlen $x \neq 0$ und y definiert man die *Potenz* x^y durch

$$x^y := e^{y \ln x}.$$

Die komplexe *Wurzelfunktion* ist definiert durch

$$\sqrt[n]{z} := z^{1/n} = e^{(\ln z)/n}, \quad n \in \mathbb{N}.$$

3.17 Beispiel:

$$i^i = e^{i \ln i} = e^{-\pi/2}, \quad \sqrt{-1} = e^{(\ln(-1))/2} = e^{i\pi/2} = i, \quad \sqrt{-16} = 4i$$

3.18 Nullstellen von Polynomen: Die Nullstellen eines quadratischen Polynoms

$$p(z) = az^2 + bz + c, \quad a \neq 0,$$

sind durch die Formel

$$z_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

gegeben. Dabei ist die Wurzel im komplexen Sinn zu verstehen und deshalb stets definiert. Es gibt also immer zwei (unter Umständen zusammenfallende) Lösungen einer quadratischen Gleichung im Komplexen.
Allgemein gilt der *Fundamentalsatz der Algebra*: Das Polynom

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0, \quad a_n \neq 0$$

hat stets n komplexe Nullstellen.

3.19 Beispiel:

$$z^2 - 6z + 13 = 0 \quad \Rightarrow \quad z_{1,2} = \frac{6 \pm \sqrt{-16}}{2} = 3 \pm 2i$$

3.20 Beispiel: Zur Lösung der Gleichung

$$z^3 - 8i = 0$$

setzt man $z = re^{i\varphi}$ und erhält

$$r^3 e^{3i\varphi} = 8e^{i\pi/2}.$$

Ein Vergleich der Beträge liefert $r = 2$, und für das Argument erhält man

$$3\varphi = \pi/2 + 2k\pi, \quad k \in \mathbb{Z}.$$

Dabei wurde verwendet, dass die e-Funktionen übereinstimmen, wenn sich die Argumente um ein ganzzahliges Vielfaches von 2π unterscheiden. Man erhält eine Folge von Argumenten

$$\varphi_k = \frac{\pi}{6} + \frac{2k\pi}{3}, \quad k \in \mathbb{Z},$$

aber nur drei davon führen auf verschiedene Lösungen,

$$\varphi_0 = \frac{\pi}{6}, \quad \varphi_1 = \frac{5\pi}{6}, \quad \varphi_2 = \frac{3\pi}{2}.$$

Alle anderen Werte unterscheiden sich von den gegebenen drei um ein ganzzahliges Vielfaches von 2π und liefern deshalb keine weiteren Lösungen. Man erhält schließlich

$$z_0 = 2e^{i\pi/6} = \sqrt{3} + i, \quad z_1 = 2e^{5i\pi/6} = -\sqrt{3} + i, \quad z_2 = 2e^{3i\pi/2} = -2i.$$

4 Lineare Gleichungssysteme

4.1 Beispiel: Man berechne den Schnittpunkt der drei Ebenen

$$\begin{aligned} E_1 &: 2x + y + z = 1 \\ E_2 &: 3x + y + z = 2 \\ E_3 &: 4x + 2y + 3z = 0, \end{aligned}$$

vgl. Abbildung 9.

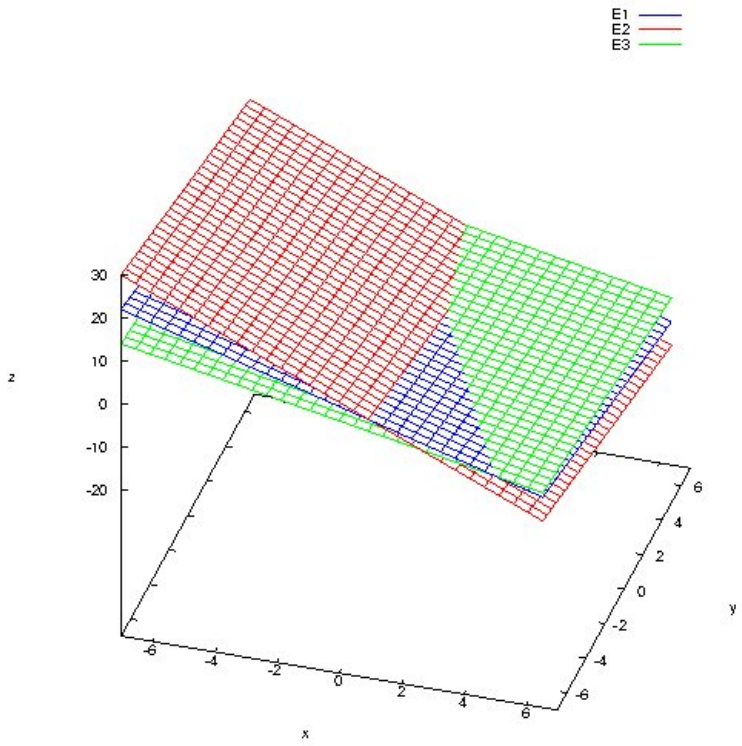


Abbildung 9: Die drei Ebenen aus Beispiel 4.1

Subtrahiert man das dreifache der ersten Zeile vom doppelten der zweiten Zeile, so erhält man die Bedingung

$$-y - z = 1.$$

Subtrahiert man das doppelte der ersten Zeile von der dritten Zeile, so erhält man die Bedingung

$$z = -2.$$

Setzt man dies in die vorherige Gleichung ein, so erhält man die Bedingung

$$-y + 2 = 1$$

und damit den Wert $y = 1$. Setzt man dies in die erste Gleichung ein, so erhält man die Bedingung

$$2x + 1 - 2 = 1$$

und damit den Wert $x = 1$. Der Schnittpunkt ist also $\vec{x} = [1, 1, -2]^T$.

4.2 Beispiel [→ **4.1**]**:** Es soll die Schnittmenge der Ebenen E_1 und E_2, siehe Abbildung 10, berechnet werden.

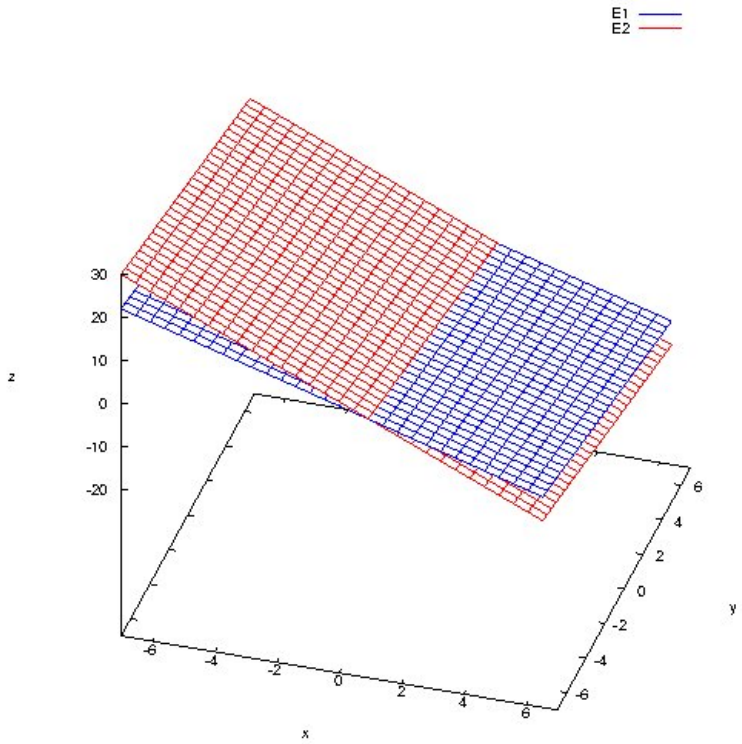


Abbildung 10: Die zwei Ebenen aus Beispiel 4.2

Wie zuvor erhält man aus E_1 und E_2 die Bedingung

$$-y - z = 1.$$

Da keine weiteren Bedingungen vorhanden sind, kann man beispielweise der Variablen z einen beliebigen Wert zuordnen, sagen wir

$$z = t, \quad t \in \mathbb{R}.$$

Damit ergibt sich

$$-y - t = 1 \quad \Rightarrow \quad y = -1 - t.$$

Setzt man dies in die Gleichung von E_1 ein, so erhält man die Bedingung

$$2x + (-t - 1) + t = 1$$

und damit $x = 1$. Die Menge aller Schnittpunkte ist also gegeben durch

$$\vec{x} = \begin{bmatrix} 1 \\ -t-1 \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

Die Lösungsmenge ist also eine Gerade.

4.3 Beispiel [→ 4.1]: Gesucht ist der Schnittpunkt der Ebenen E_1, E_2 und E_4, wobei

$$E_4 \ : \ x + y + z = 3,$$

vgl. Abbildung 11.

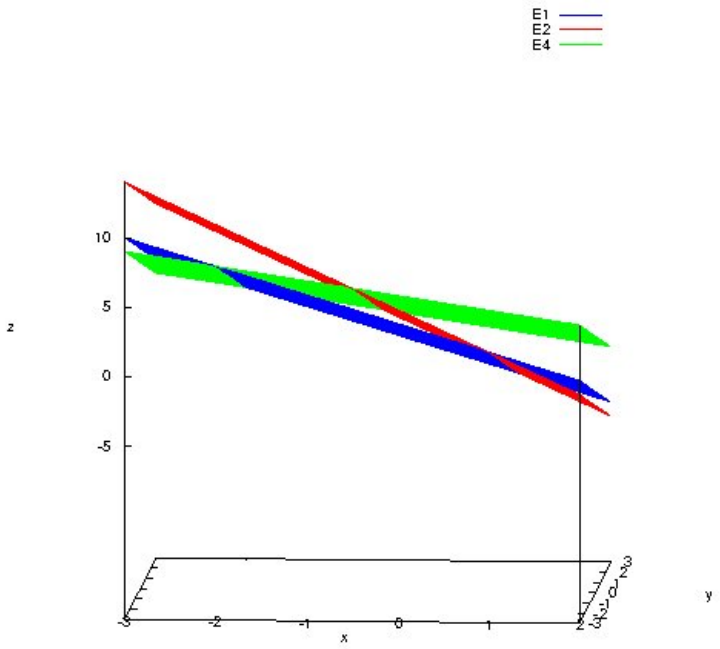


Abbildung 11: Die drei Ebenen aus Beispiel 4.3

Subtrahiert man vom doppelten dieser Gleichung die Gleichung von E_1, so erhält man die Bedingung

$$y + z = 5.$$

Außerdem folgt aus den Gleichungen für E_1 und E_2 wie zuvor

$$-y - z = 1.$$

Addiert man die beiden letzten Gleichungen, so erhält man den Widerspruch

$$0 = 6.$$

Es gibt also keinen Schnittpunkt.

4.4 Lineares Gleichungssystem: Ein *lineares Gleichungssystem (LGS)* mit m Gleichungen für den Vektor $\vec{x} = [x_1, x_2, \ldots, x_n]^T$ der Unbekannten hat die Form

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n &= b_2 \\ \vdots \qquad\qquad \vdots \qquad\qquad & \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= b_m \end{aligned}$$

wobei die Koeffizienten $a_{i,j}$ und die Werte b_i vorgegebene reelle Zahlen sind. Gesucht ist die Menge aller Vektoren $\vec{x}$, für die alle Gleichungen erfüllt sind. Die Koeffizienten $a_{i,j}$ auf der linken Seite kann man zu einem Zahlenschema der Form

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

zusammenfassen. Man nennt A die *Matrix* des Gleichungssystems. Die Werte auf der rechten Seite $b_1, b_2, \ldots, b_m$ kann man zu einem Vektor $\vec{b} = [b_1, \ldots, b_m]^T$ zusammenfassen und man schreibt für das LGS dann auch kurz

$$A\vec{x} = \vec{b}.$$

Man nennt das LGS

- *unterbestimmt*, falls $m < n$,
- *quadratisch*, falls $m = n$,
- *überbestimmt*, falls $m > n$.

4.5 Beispiel:

- In Bsp. 4.1 ergibt sich ein quadratisches Gleichungssystem mit

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 1 & 1 \\ 4 & 2 & 3 \end{bmatrix} \quad \text{und} \quad \vec{b} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}.$$

- In Bsp. 4.2 ergibt sich ein unterbestimmtes Gleichungssystem mit

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 3 & 1 & 1 \end{bmatrix} \quad \text{und} \quad \vec{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

4.6 Elementare Umformungen: Es ist zweckmäßig, das LGS $A\vec{x} = \vec{b}$ in folgendem Schema zu notieren:

	x_1	x_2	$\cdots$	x_n	$\vec{b}$
$\boxed{1}$:	$a_{1,1}$	$a_{1,2}$	$\cdots$	$a_{1,n}$	b_1
$\boxed{2}$:	$a_{2,1}$	$a_{2,2}$	$\cdots$	$a_{2,n}$	b_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\boxed{m}$:	$a_{m,1}$	$a_{m,2}$	$\cdots$	$a_{m,n}$	b_m

Die Zahlen in den Kästchen enthalten fortlaufende Zeilennummern, die nur der Kennzeichnung dienen. In dem Schema sind die folgenden *elementaren Umformungen* erlaubt:

- *Zeilenvertauschung:* Zwei Zeilen

$$\boxed{i} \quad \leftrightarrow \quad \boxed{j}$$

  dürfen vertauscht werden.

- *Spaltenvertauschung:* Zwei Spalten

$$x_i \quad \leftrightarrow \quad x_j$$

dürfen vertauscht werden. Dabei ist zu beachten, dass auch die Einträge in der Kopfzeile vertauscht werden.

- *Linearkombination:* Die i-te Zeile darf durch die Linearkombination

$$\boxed{i} \quad \leftarrow \quad p \times \boxed{i} - q \times \boxed{j}$$

ersetzt werden, sofern $p \neq 0$. Insbesondere kann man $q = 0$ wählen und so eine Skalierung der i-ten Zeile erreichen.

Durch geeignete elementare Umformungen kann man ein gegebenes LGS in eine einfachere Form überführen, deren Lösung sich unmittelbar ablesen lässt.

4.7 Beispiel [→ 4.1]: Das Schema zu dem angegebenen LGS hat die Form

	x	y	z	$\vec{b}$
$\boxed{1}$:	2	1	1	1
$\boxed{2}$:	3	1	1	2
$\boxed{3}$:	4	2	3	0

Durch Linearkombination können die jeweils ersten Koeffizienten der zweiten und der dritten Zeile zu Null gemacht werden:

	x	y	z	$\vec{b}$
$\boxed{1}$:	2	1	1	1
$2 \times \boxed{2} - 3 \times \boxed{1} = \boxed{4}$:	0	−1	−1	1
$1 \times \boxed{3} - 2 \times \boxed{1} = \boxed{5}$:	0	0	1	−2

Das Schema hat nun gestaffelte Form und kann schrittweise aufgelöst werden:

$$\begin{aligned}
\boxed{5} &: \quad z = -2 \\
\boxed{4} &: \quad -y - z = 1 \quad \Rightarrow \quad -y + 2 = 1 \quad \Rightarrow \quad y = 1 \\
\boxed{1} &: \quad 2x + y + z = 1 \quad \Rightarrow \quad 2x + 1 - 2 = 1 \quad \Rightarrow \quad x = 1
\end{aligned}$$

Die Lösung ist also $\vec{x} = [1, 1, -2]^T$.

4.8 Beispiel [→ 4.2]: Das Schema zu dem angegebenen LGS hat die Form

	x	y	z	$\vec{b}$
$\boxed{1}$:	2	1	1	1
$\boxed{2}$:	3	1	1	2

Durch Linearkombination kann der erste Koeffizient der zweiten Zeile zu Null gemacht werden:

	x	y	z	$\vec{b}$
$\boxed{1}$:	2	1	1	1
$2 \times \boxed{2} - 3 \times \boxed{1} = \boxed{3}$:	0	−1	−1	1

In der letzten Zeile kann entweder der Wert von y oder der Wert von z frei gewählt werden. Wir setzen z.B. $z = t$ für eine beliebige Zahl $t \in \mathbb{R}$ und erhalten damit

$$\boxed{3}: \quad -y - z = 1 \quad \Rightarrow \quad -y - t = 1 \quad \Rightarrow \quad y = -1 - t$$
$$\boxed{1}: \quad 2x + y + z = 1 \quad \Rightarrow \quad 2x + (-1 - t) + t = 1 \quad \Rightarrow \quad x = 1$$

Die Lösungsmenge ist also die Gerade

$$\vec{x} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \quad t \in \mathbb{R}.$$

4.9 Beispiel [→ 4.3]: Das LGS hat hier die Form

	x	y	z	$\vec{b}$
$\boxed{1}$:	2	1	1	1
$\boxed{2}$:	3	1	1	2
$\boxed{3}$:	1	1	1	3

Elimination der Einträge in der ersten Spalte mittels Linearkombination ergibt

	x	y	z	$\vec{b}$
$\boxed{1}$:	2	1	1	1
$2 \times \boxed{2} - 3 \times \boxed{1} = \boxed{4}$:	0	−1	−1	1
$2 \times \boxed{3} - 1 \times \boxed{1} = \boxed{5}$:	0	1	1	5

Elimination in der zweiten Spalte ergibt die gestaffelte Form

	x	y	z	$\vec{b}$
$\boxed{1}$:	2	1	1	1
$\boxed{4}$:	0	−1	−1	1
$\boxed{5} + \boxed{4} = \boxed{6}$:	0	0	0	6

Aus der letzten Zeile ergibt sich der Widerspurch

$$0x + 0y + 0z = 6.$$

Es existiert also keine Lösung.

4.10 Gestaffelte Form: Wie in den Beispielen zuvor gesehen, lässt sich die Lösung eines LGS einfach bestimmen, indem man es durch elementare Umformungen auf *gestaffelte Form* bringt:

$\tilde{x}_1$	$\tilde{x}_2$	$\cdots$	$\tilde{x}_r$	$\tilde{x}_{r+1}$	$\cdots$	$\tilde{x}_n$	$\vec{b}$
•	$*$	$\cdots$	$*$	$*$	$\cdots$	$*$	$*$
0	•	$\cdots$	$*$	$*$	$\cdots$	$*$	$*$
0	0	$\cdots$	$*$	$*$	$\cdots$	$*$	$*$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
0	0	$\cdots$	•	$*$	$\cdots$	$*$	$*$
0	0	$\cdots$	0	0	$\cdots$	0	$\times$
$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
0	0	$\cdots$	0	0	$\cdots$	0	$\times$

Dabei sind

- $\tilde{x}_1, \ldots, \tilde{x}_n$ eine Umordnung der gesuchten Komponenten $x_1, \ldots, x_n$, die durch Spaltenvertauschungen entsteht,
- alle mit $\bullet$ markierten Einträge von Null verschieden,
- alle mit $*$ markierten Einträge beliebig,
- alle mit $\times$ markierten Einträge beliebig.

Die Existenz von Lösungen hängt von den mit $\times$ markierten Einträgen ab:

- Wenn ein einziger dieser Einträge von Null verschieden ist, dann besitzt das LGS keine Lösung.
- Wenn es keine Nullzeilen und damit keine derartigen Einträge gibt oder wenn alle diese Einträge gleich Null sind, dann exisitieren Lösungen. Diese sind wie folgt gegegeben: Die Werte von $\tilde{x}_{r+1}, \ldots, \tilde{x}_n$ können beliebig vorgegeben werden,

  $$\tilde{x}_{r+1} = t_1, \quad \ldots, \quad \tilde{x}_n = t_{n-r}, \quad t_1, \ldots, t_{n-r} \in \mathbb{R}.$$

  Davon ausgehend können der Reihe nach die Werte von $\tilde{x}_r, \tilde{x}_{r-1}, \ldots, \tilde{x}_1$ bestimmt werden.

4.11 Gauss-Algorithmus: Der *Gauss-Algorithmus* gibt Regeln an, mit denen ein LGS auf gestaffelte Form gebracht werden kann:

1. Suche ein Element $a_{i,j} \neq 0$. Vertausche die erste mit der j-ten Spalte und vertausche die erste mit der i-ten Zeile.
2. Ersetze alle Zeilen mit Index $i \geq 2$ durch die Linearkombination

   $$a_{1,1} \times \boxed{i} - a_{i,1} \times \boxed{1}.$$

Damit haben die erste Zeile und die erste Spalte die gewünschte Form. Sie werden im weiteren Verlauf des Algorithmus nicht mehr verändert. Nun wendet man das Verfahren analog auf die zweite Zeile und die zweite Spalte an, wobei zu beachten ist, dass die erste Zeile nicht mehr für Zeilenvertauschungen verwendet werden darf. So verfährt man weiter, bis die gestaffelte Form erreicht ist.

4.12 Beispiel: Für einen reellen Parameter $\alpha \in \mathbb{R}$ sei das folgende LGS gegeben:

	x_1	x_2	x_3	x_4	x_5	$\vec{b}$
$\boxed{1}$:	1	1	3	-2	4	-1
$\boxed{2}$:	-1	-1	-3	2	-4	α
$\boxed{3}$:	0	0	-2	1	0	3
$\boxed{4}$:	0	0	4	-2	0	-6

Elimination der Einträge in der ersten Spalte mittels Linearkombination ergibt

	x_1	x_2	x_3	x_4	x_5	$\vec{b}$
$\boxed{1}$:	1	1	3	−2	4	−1
$\boxed{5}$:	0	0	0	0	0	$\alpha - 1$
$\boxed{3}$:	0	0	−2	1	0	3
$\boxed{4}$:	0	0	4	−2	0	−6

Um die zweite Zeile in die gewünschte Form zu bringen, wird in den Zeilen $\boxed{5}, \boxed{3}, \boxed{4}$ ein von Null verschiedener Eintrag gesucht. Wir wählen z.B. den Eintrag $a_{3,4} = 1$. Vertauschung der zweiten und der vierten Spalte sowie der Zeilen $\boxed{5}$ und $\boxed{3}$ ergibt

	x_1	x_4	x_3	x_2	x_5	$\vec{b}$
$\boxed{1}$:	1	−2	3	1	4	−1
$\boxed{3}$:	0	1	−2	0	0	3
$\boxed{5}$:	0	0	0	0	0	$\alpha - 1$
$\boxed{4}$:	0	−2	4	0	0	−6

Nun werden die Einträge der Zeilen $\boxed{5}, \boxed{4}$ in der zweiten Spalte zu Null gemacht:

	x_1	x_4	x_3	x_2	x_5	$\vec{b}$
$\boxed{1}$:	1	−2	3	1	4	−1
$\boxed{3}$:	0	1	−2	0	0	3
$\boxed{5}$:	0	0	0	0	0	$\alpha - 1$
$\boxed{6}$:	0	0	0	0	0	0

Damit ist die gestaffelte Form erreicht. Es ist $r = 2$, und die Umordnung der Lösungskomponenten ist hier

$$\tilde{x}_1 = x_1, \quad \tilde{x}_2 = x_4, \quad \tilde{x}_3 = x_3, \quad \tilde{x}_4 = x_2, \quad \tilde{x}_5 = x_5.$$

Nun sind zwei Fälle zu unterscheiden:

- Wenn $\alpha \neq 1$, dann gibt es keine Lösung.
- Wenn $\alpha = 1$, dann gibt es einen Lösungsraum mit $n - r = 3$ freien Parametern,

$$x_3 = t_1, \quad x_2 = t_2, \quad x_5 = t_3.$$

Durch Einsetzen in die Zeilen $\boxed{3}$ und $\boxed{1}$ erhält man schließlich die Lösung

$$\vec{x} = \begin{bmatrix} 5 \\ 0 \\ 0 \\ 3 \\ 0 \end{bmatrix} + t_1 \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix} + t_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t_3 \begin{bmatrix} -4 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad t_1, t_2, t_3 \in \mathbb{R}.$$

4.13 Homogene LGS: Ein LGS $A\vec{x} = \vec{b}$ heißt *homogen*, wenn die rechte Seite der Nullvektor ist und anderenfalls *inhomogen*. Ein homogenes LGS besitzt stets mindestens eine Lösung, nämlich den Nullvektor. Betrachtet man die gestaffelte Form, dann sind alle mit $\times$ markierten Einträge Null. Man kann also die Werte

$$\tilde{x}_{r+1} = t_1, \quad , \ldots , \quad \tilde{x}_n = t_{n-r}, \quad t_1, \ldots, t_{n-r} \in \mathbb{R}$$

beliebig vorgeben und erhält somit eine Lösungsmenge mit $(n-r)$ freien Parametern. Diese entsprechen $(n-r)$ genau linear unabhängigen Lösungen. Die Lösungsmenge bezeichnet man als *Kern von* A und schreibt dafür

$$\ker A := \{\vec{x} : A\vec{x} = \vec{0}\}.$$

Der Kern von A ist ein linearer Teilraum des $\mathbb{R}^n$ mit $\dim \ker A := n - r$ (vgl. 5.11 im nächsten Kapitel). Die Zahl r, also die Anzahl der von Null verschiedenen Zeilen in der gestaffelten Form wird als *Rang* von A bezeichnet und man schreibt $\operatorname{rang} A := r$. Es gilt also

$$\dim \ker A + \operatorname{rang} A = n,$$

d.h., die Dimension des Kerns und der Rang der Matrix ergeben zusammen die Spaltenzahl (*Dimensionsformel*).

4.14 Beispiel [→ 4.12]: Sei

$$A = \begin{bmatrix} 1 & 1 & 3 & -2 & 4 \\ -1 & -1 & -3 & 2 & -4 \\ 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 4 & -2 & 0 \end{bmatrix},$$

dann erhält man für das homogene LGS $A\vec{x} = \vec{0}$ die gestaffelte Form

	x_1	x_4	x_3	x_2	x_5	$\vec{b}$
$\boxed{1}$:	1	−2	3	1	4	0
$\boxed{3}$:	0	1	−2	0	0	0
$\boxed{5}$:	0	0	0	0	0	0
$\boxed{6}$:	0	0	0	0	0	0

.

Hier ist wir zuvor $r = 2$ und damit

$$\operatorname{rang} A = 2 \quad \text{und} \quad \dim \ker A = 3.$$

Mit

$$\vec{x}_1 := \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \quad \vec{x}_2 := \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{x}_3 := \begin{bmatrix} -4 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

ist der Kern von A gegeben durch

$$\ker A = \{t_1\vec{x}_1 + t_2\vec{x}_2 + t_3\vec{x}_3,\ t_1, t_2, t_3 \in \mathbb{R}\} = \operatorname{Lin}(\vec{x}_1, \vec{x}_2, \vec{x}_3).$$

Es ist also $\vec{x}_1, \vec{x}_2, \vec{x}_3$ eine Basis von $\ker A$.

4.15 Superpositionsprinzip: Sei $\vec{x}_s$ eine Lösung des LGS $A\vec{x} = \vec{b}$ und $\vec{x}_h \in \ker A$ eine Lösung des zugehörigen homogenen Systems, dann ist auch $\vec{x} := \vec{x}_s + \vec{x}_h$ eine Lösung. Sind umgekert $\vec{x}$ und $\vec{x}_s$ Lösungen von $A\vec{x} = \vec{b}$, dann ist $\vec{x}_h := \vec{x} - \vec{x}_s \in \ker A$ eine Lösung des homogenen Systsms. Man kann also jede Lösung von $A\vec{x} = \vec{b}$ in der Form

$$\vec{x} = \vec{x}_s + \vec{x}_h, \quad \vec{x}_h \in \ker A,$$

darstellen. Mit anderen Worten gilt: Die allgemeine Lösung eines inhomogenen Systems erhält man als Summe einer speziellen Lösung dieses Systems und der allgemeinen Lösung des zugehörigen homogenen Systems. Dieser grundlegende Sachverhalt wird als *Superpositionsprinzip* bezeichnet.
Beachte: Für $\vec{b} \neq \vec{0}$ ist der Lösungsraum des inhomogenen Systems $A\vec{x} = \vec{b}$ kein linearer Teilraum, da Null keine Lösung liefert.

4.16 Beispiel [→ 4.12]: Sei speziell $\alpha = 1$. Man rechnet leicht nach, dass z.B.

$$\vec{x}_s := \begin{bmatrix} 0 \\ 0 \\ -5 \\ -7 \\ 0 \end{bmatrix}$$

das gegebene inhomogene LGS

$$\begin{bmatrix} 1 & 1 & 3 & -2 & 4 \\ -1 & -1 & -3 & 2 & -4 \\ 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 4 & -2 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} -1 \\ 1 \\ 3 \\ -6 \end{bmatrix}$$

löst. Zusammen mit dem in Beispiel 4.14 bestimmten Kern von A erhält man somit die Lösungsmenge

$$\vec{x} = \vec{x}_s + t_1\vec{x}_1 + t_2\vec{x}_2 + t_3\vec{x}_3, \quad t_1, t_2, t_3 \in \mathbb{R}.$$

Diese Darstellung unterscheidet sich von der in Beispiel 4.12 angegeben Form. Die Gesamtheit der Lösungen ist aber in beiden Fällen genau dieselbe. Dies sieht man, indem man in der hier angegebenen Lösung den freien Parameter t_1 durch $t_1 + 5$ ersetzt.

4.17 Determinante: Sei A eine $(n \times n)$-Matrix. Dann kann man die eindeutige Lösbarkeit des LGS $A\vec{x} = \vec{b}$ mit Hilfe der *Determinante* von A entscheiden. Die Determinante ist eine reelle Zahl, die wie folgt definiert ist: Wenn A eine (1×1)-Matrix ist, dann ist $\det A := a_{1,1}$. Anderenfalls gilt

$$\det A := \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det A_{i,j}.$$

Dabei ist i ein beliebiger Zeilenindex und $A_{i,j}$ eine $(n-1) \times (n-1)$-Matrix, die durch Streichen der i-ten Zeile und der j-ten Spalte entsteht. Damit ist die Berechnung der Determinante auf ein Problem niedrigerer Dimension zurückgeführt und wiederholte Anwendung führt schließlich auf Determinanten von Matrizen der Dimension (1×1).

Anstelle der oben angebenen Formel, die man auch *Entwicklung nach der i-ten Zeile* nennt, kann man auch nach der j-ten Spalte entwickeln,

$$\det A := \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det A_{i,j}.$$

Es gilt: Das quadratische *LGS* $A\vec{x} = \vec{b}$ ist genau dann eindeutig lösbar, wenn $\det A \neq 0$. Äquivalent hierzu sind die Aussagen

$$\det A \neq 0 \;\Leftrightarrow\; \dim \ker A = 0 \;\Leftrightarrow\; \operatorname{rang} A = n.$$

4.18 Spezialfälle:

- $n = 2$:
$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \det A = ad - bc.$$

- $n = 3$: Entwicklung nach der ersten Zeile ergibt
$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}, \quad \det A = a(ei - hf) - b(di - gf) + c(dh - ge).$$
Alternativ verwendet man die *Regel von Sarrus*. Achtung, diese Regel ist nicht für höherdimensionale Matrizen gültig.

- Wenn A eine obere oder untere *Dreiecksmatrix* ist, also
$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ 0 & 0 & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n,n} \end{bmatrix} \quad \text{oder} \quad A = \begin{bmatrix} a_{1,1} & 0 & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & 0 & \cdots & 0 \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \cdots & a_{n,n} \end{bmatrix},$$
dann ist $\det A$ das Produkt der Diagonalelemente,
$$\det A = a_{1,1} a_{2,2} \cdots a_{n,n}.$$

4.19 Beispiel:

- $$\det \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix} = 2$$

- $$\det \begin{bmatrix} 1 & 3 & 2 \\ 2 & 0 & 1 \\ 3 & 2 & 1 \end{bmatrix} = 9$$

- $$\det \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 0 & 2 \end{bmatrix} = 1 \cdot 3 \cdot 4 \cdot 2 = 24$$

4.20 Geometrische Bedeutung:

- Für $n = 2$ ist der Betrag von $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ gleich dem Flächeninhalt des von den Vektoren $\begin{bmatrix} a \\ c \end{bmatrix}$ und $\begin{bmatrix} b \\ d \end{bmatrix}$ aufgespannten Parallelogramms, vgl. Abbildung 12.

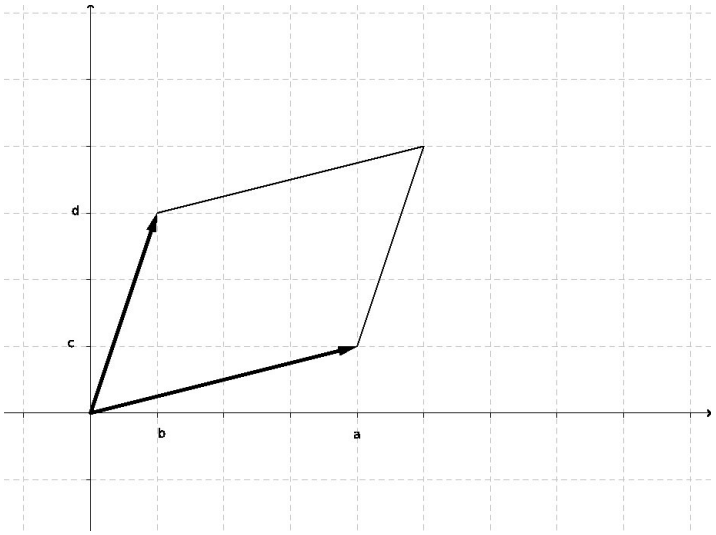


Abbildung 12: Das von den Vektoren $[a, c]^T$ und $[b, d]^T$ aufgespannte Parallelogramm

- Für $n = 3$ gilt:

$$\det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = \det(\vec{a}\ \vec{b}\ \vec{c}) = \langle \vec{a}, \vec{b} \times \vec{c} \rangle.$$

  Der Betrag ist nach 2.9 das Volumen des Spats, der von den Vektoren $\vec{a}$, $\vec{b}$, $\vec{c}$ aufgespannt wird.

4.21 Determinanten und elementare Umformungen: Elementare Umformungen ändern die Determinante einer $n \times n$-Matrix A:

- Vertauscht man zwei Spalten oder zwei Zeilen in A, so ändert sich das Vorzeichen. Das heißt, wenn A_1 die neue Matrix bezeichnet, so gilt $\det A_1 = -\det A$.
- Ersetzt man die i-te Zeile in A durch $p\cdot(i\text{-te Zeile})+q\cdot(j\text{–te Zeile})$, so multipliziert sich die Determinante mit p. Das heißt, wenn A_1 die neue Matrix bezeichnet, so gilt $\det A_1 = p \cdot \det A$. *Achtung:* Im Fall $p = 0$ erhlt man das Ergebnis $\det A_1 = 0$ und kann somit keine Rückschlsse mehr auf $\det A$ ziehen.

4.22 Beispiel: Sei $A = \begin{bmatrix} 0 & 3 & 0 \\ 2 & 1 & 1 \\ -1 & 2 & 0 \end{bmatrix}$. Vertausche die ersten zwei Spalten:

$$A_1 = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 2 & 1 \\ 2 & -1 & 0 \end{bmatrix}.$$

Es gilt $\det A_1 = -\det A$. Ersetze in A_1 nun 2. Zeile durch $3\cdot(2\text{–te Zeile})-1\cdot(1\text{–te Zeile})$. Danach in A_2 die 3. Zeile durch $3 \cdot (3\text{–te Zeile}) - 2 \cdot (1\text{–te Zeile})$:

$$A_2 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 3 \\ 2 & -1 & 0 \end{bmatrix}. \quad A_3 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 3 \\ 0 & -3 & 0 \end{bmatrix}.$$

Es gilt $\det A_2 = 3\det A_1$, $\det A_3 = 3\det A_2$. Insgesamt also $\det A_3 = 3\det A_2 = 9\det A_1 = -9\det A$.
Ersetze in A_3 die 3. Zeile durch $2 \cdot (3\text{–te Zeile}) + (2\text{–te Zeile})$:

$$A_4 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 3 \\ 0 & 0 & 3 \end{bmatrix}.$$

Es gilt $\det A_4 = 2\det A_3$. Insgesamt also $\det A_4 = -18\det A$. Wegen $\det A_4 = 3\cdot6\cdot3 = 54$ also $\det A = -3$.

5 Matrizenrechnung

5.1 Transponierter Vektor: Die Notation $\vec{x} \in \mathbb{R}^n$ bezieht sich per Definition[3] immer auf einen stehenden Vektor,

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Der *transponierte Vektor* $\vec{x}^T$ ist das zugehörige liegende Zahlenschema

$$\vec{x}^T := [x_1, x_2, \ldots, x_n],$$

das man auch als *liegenden Vektor* bezeichnet. Die Transposition eines liegenden Vektors ergibt wieder einen stehenden Vektor,

$$\vec{x} = (\vec{x}^T)^T = [x_1, x_2, \ldots, x_n]^T.$$

5.2 Matrix: Ein Zahlenschema der Form

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

heißt $(m \times n)$-*Matrix*. Im Fall $n = m$ heißt die Matrix *quadratisch*. Die Einträge $a_{i,j} \in \mathbb{R}$ heißen *Elemente* der Matrix. Analog zur Schreibweise $\vec{x} \in \mathbb{R}^n$ für Vektoren schreiben wir $A \in \mathbb{R}^{m \times n}$ für $(m \times n)$-Matrizen. Die *Spalten* der Matrix A sind Vektoren in $\mathbb{R}^m$,

$$A = [\vec{a}^1, \vec{a}^2, \ldots, \vec{a}^n], \quad \vec{a}^j := \begin{bmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{bmatrix}.$$

Die *Zeilen* der Matrix A sind liegende Vektoren in $\mathbb{R}^n$,

$$A = \begin{bmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vdots \\ \vec{a}_m^T \end{bmatrix}, \quad \vec{a}_i^T := [a_{i,1}, a_{i,2}, \ldots, a_{i,n}].$$

5.3 Beispiel: Die (3×4)-Matrix $A \in \mathbb{R}^{3 \times 4}$ sei gegeben durch

$$A = \begin{bmatrix} 3 & 1 & 4 & 5 \\ 1 & 5 & 2 & 0 \\ 0 & 7 & 3 & 2 \end{bmatrix}.$$

Dann ist

$$a_{3,2} = 7, \quad \vec{a}^3 = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{a}_2^T = [1, 5, 2, 0], \quad \vec{a}_2 = \begin{bmatrix} 1 \\ 5 \\ 2 \\ 0 \end{bmatrix}.$$

[3] Diese Definition bezieht sich auf das vorliegende Skript und ist keineswegs allgemeingültig.

5.4 Addition und Skalarmultiplikation: Seien $A, B \in \mathbb{R}^{m\times n}$ zwei Matrizen gleichen Formats. Dann gilt Folgendes:

- *Addition, Subtraktion:* $C := A \pm B$ ist eine $(m \times n)$-Matrix mit Elementen

$$c_{i,j} = a_{i,j} \pm b_{i,j}.$$

- *Skalarmultiplikation:* Für $\alpha \in \mathbb{R}$ ist $C := \alpha A$ eine $(m \times n)$-Matrix mit Elementen

$$c_{i,j} = \alpha a_{i,j}.$$

  Insbesondere ist $1A = A$, $(-1)A = -A$ und

$$0A = 0_{m,n} := \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

  die $(m \times n)$-*Nullmatrix*. Wenn aus dem Zusammenhang klar ist, welches Format die Nullmatrix hat, schreiben wir anstelle von $0_{m,n}$ auch einfach 0.

- *Distributivgesetze:* Für $\alpha, \beta \in \mathbb{R}$ gilt

$$(\alpha + \beta)A = \alpha A + \beta A, \quad \alpha(A + B) = \alpha A + \alpha B.$$

5.5 Matrizenprodukt: Sei A eine $(m \times n)$-Matrix und B eine $(n \times k)$-Matrix. Dann ist das *Matrizenprodukt* $C := A \cdot B$ eine $(m \times k)$-Matrix, die durch

$$c_{i,j} = \langle \vec{a}_i, \vec{b}^j \rangle = \sum_{s=1}^{n} a_{i,s} b_{s,j}$$

definiert ist. Das Element $c_{i,j}$ ist also das Skalarprodukt der i-ten Zeile von A mit der j-ten Spalte von B. Insbesondere macht das Matrizenprodukt nur dann Sinn, wenn die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmt, da anderenfalls das Skalarprodukt nicht erklärt ist. Der Mal-Punkt wird meist weggelassen, wenn Verwechslungen ausgeschlossen sind, also $AB = A \cdot B$.

5.6 Beispiel: Für

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 0 & 3 \\ -4 & 1 & 2 \end{bmatrix}$$

ist

$$AB = \begin{bmatrix} -3 & 3 & 14 \\ 6 & -1 & 4 \end{bmatrix}, \quad B \cdot B = B^2 = \begin{bmatrix} 1 & 3 & 7 \\ -10 & 4 & 7 \\ -15 & -2 & 3 \end{bmatrix}.$$

Die Produkte $A \cdot A$ und $B \cdot A$ sind nicht definiert.

5.7 Vektoren als spezielle Matrizen: Ein Vektor $\vec{x} \in \mathbb{R}^n$ kann als $(n \times 1)$-Matrix aufgefasst werden, also als Matrix mit nur einer Spalte. Genauso kann der liegende Vektor $\vec{x}^T$ als $(1 \times n)$-Matrix aufgefasst werden, also als Matrix mit nur einer Zeile. In diesem Sinne ist für $A \in \mathbb{R}^{m \times n}$ das Matrizenprodukt $\vec{b} = A \cdot \vec{x}$ ein Vektor in $\mathbb{R}^m$ mit Komponenten

$$b_i = \langle \vec{a}_i, \vec{x} \rangle = \sum_{j=1}^{n} a_{i,j} x_j.$$

Man kann den Vektor $\vec{b}$ auch als Linearkombination der Spalten von A deuten,

$$\vec{b} = A\vec{x} = \sum_{j=1}^{n} \vec{a}^j x_j.$$

Analog ist für $\vec{y} \in \mathbb{R}^m$ das Matrizenprodukt $\vec{c}^T = \vec{y}^T A$ ein liegender Vektor mit Komponenten

$$c_j = \langle \vec{y}, \vec{a}^j \rangle,$$

der auch als Linearkombination der Zeilen von A gedeutet werden kann,

$$\vec{c}^T = \vec{y}^T A = \sum_{i=1}^{m} \vec{a}_i^T y_i.$$

Sind $\vec{x}, \vec{z} \in \mathbb{R}^n$ zwei Vektoren gleicher Länge, dann ist das Matrizenprodukt

$$\vec{x}^T \cdot \vec{z} = \sum_{j=1}^{n} x_j z_j = \langle \vec{x}, \vec{z} \rangle$$

gerade das Skalarprodukt der beiden Vektoren. Für beliebige Vektoren $\vec{x} \in \mathbb{R}^n$ und $\vec{y} \in \mathbb{R}^m$ ist aber auch das Produkt

$$B = \vec{x} \cdot \vec{y}^T = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \cdots & x_1 y_m \\ x_2 y_1 & x_2 y_2 & \cdots & x_2 y_m \\ \vdots & \vdots & \ddots & \vdots \\ x_n y_1 & x_n y_2 & \cdots & x_n y_m \end{bmatrix}$$

definiert. *Merke:*

- Liegender Vektor mal stehender Vektor ergibt eine reelle Zahl.
- Stehender Vektor mal liegender Vektor ergibt eine Matrix.

5.8 Beispiel: Für

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 2 & -1 \end{bmatrix}, \quad \vec{x} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \vec{z} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

ist

$$A\vec{x} = \begin{bmatrix} 11 \\ 3 \end{bmatrix}, \quad \vec{y}^T A = \begin{bmatrix} 2 & 4 & 5 \end{bmatrix}, \quad \vec{x}^T \vec{z} = 5, \quad \vec{x}\,\vec{y}^T = \begin{bmatrix} 6 & -3 \\ 4 & -2 \\ 2 & -1 \end{bmatrix}.$$

Meist werden die Komponenten liegender Vektoren durch Kommata und Matrizenelemente durch Zwischenraum getrennt. Da einzeilige Matrizen aber liegenden Vektoren entsprechen, werden hier beide Varianten verwendet. Man schreibt also auch $\vec{y}^T A = [2, 4, 5]$. Dies entspricht im Übrigen den Konventionen der Programmiersprache Matlab, bei der Matrizenelemente einer Zeile entweder durch ein Leerzeichen oder durch ein Komma getrennt werden können.

5.9 Rechenregeln:

- Für die Matrizenmultiplikation und -addition gilt das *Distributivgesetz*, d.h., es gilt
$$A \cdot (B + C) = A \cdot B + A \cdot C.$$
Insbesondere gilt $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$.

- Es gilt das *Assoziativgesetz*, d.h., es gilt
$$A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$
Nachdem die Reihenfolge der Berechnung beliebig ist, schreibt man für das Produkt auch kurz ABC. Speziell für die Multiplikation mit einem Skalar $\alpha \in \mathbb{R}$ gilt $A(\alpha B) = (\alpha A)B = \alpha(AB)$.

- Das *Kommutativgesetz* gilt dagegen *nicht*, d.h., im Allgemeinen ist
$$AB \neq BA$$
und genauso
$$(A + B)^2 \neq A^2 + 2AB + B^2.$$

- Aus $AB = 0$ folgt *nicht* $A = 0$ oder $B = 0$.

5.10 Beispiel: Für
$$A = \begin{bmatrix} 2 & -1 \\ -2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$
ist
$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad BA = \begin{bmatrix} -2 & 1 \\ -4 & 2 \end{bmatrix}.$$

5.11 LGS in Matrizenschreibweise: Die Definition der Matrizenmultiplikation erklärtnun insbesondere die im vorigen Kapitel eingeführte Schreibweise für lineare Gleichungssysteme: Der i-te Eintrag von $A\vec{x} = \vec{b}$ liefert
$$a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n = b_i$$
und somit gerade die i-te Zeile des in 4.4 angegebenen lineares Gleichungssystem. Es lässt sich nun mit Hilfe der Rechenregeln für Matrizen leicht überprüfen, dass der Lösungsraum $\ker A$ des homogenen LGS $A\vec{x} = \vec{0}$ tatsächlich ein 5inearer Teilraum von $\mathbb{R}^n$ ist: Für $\vec{x}, \vec{y} \in \ker A$ und $\lambda \in \mathbb{R}$ gilt $A\vec{x} = A\vec{y} = \vec{0}$, also $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y} = \vec{0} + \vec{0} = \vec{0}$ und $A(\lambda\vec{x}) = \lambda(A\vec{x}) = \vec{0}$. Also sind $\vec{x} + \vec{y}$ und $\lambda\vec{x}$ ebenfalls in $\ker A$ enthalten. Genauso lässt sich das Superpositionsprinzip [→ 4.15] leicht nachrechnen.

5.12 Transposition: Die *Transponierte* einer $(m\times n)$-Matrix A ist eine $(n\times m)$-Matrix, die mit A^T bezeichnet wird. Die Spalten von A^T sind die transponierten Zeilen von A,

$$A = \begin{bmatrix} \vec{a}_1^{\,T} \\ \vec{a}_2^{\,T} \\ \vdots \\ \vec{a}_m \end{bmatrix} \quad \Rightarrow \quad A^T = [\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m].$$

Es gilt

- $(A^T)^T = A$.
- $(A+B)^T = A^T + B^T$ und $(AB)^T = B^T A^T$.
- $\langle A\vec{x}, \vec{y}\rangle = (A\vec{x})^T \cdot \vec{y} = \vec{x}^T \cdot (A^T\vec{y}) = \langle \vec{x}, A^T\vec{y}\rangle$.

5.13 Symmetrische Matrizen: Eine quadratische Matrix A heißt *symmetrisch*, wenn $A^T = A$. Es gilt

- Die Summe symmetrischer Matrizen ist symmetrisch.
- Das Produkt symmetrischer Matrizen ist im Allgemeinen *nicht* symmetrisch.
- Für beliebiges A ist sowohl $A^T + A$ als auch $A^T \cdot A$ symmetrisch.

5.14 Beispiel: Für

$$A = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

ist

$$A^T = \begin{bmatrix} 3 & 0 & 1 \\ 2 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}, \quad A^T + A = \begin{bmatrix} 6 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 10 & 6 & 1 \\ 6 & 8 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

Insbesondere sind $A^T + A$ und $A^T A$ symmetrisch.

5.15 Determinante [→ 4.17]: Seien A und B zwei $(n \times n)$-Matrizen, dann gilt

- *Vielfaches:*
$$\det(\alpha A) = \alpha^n \det A, \quad \alpha \in \mathbb{R}$$
Achtung, Exponent von α beachten!
- *Produkt:*
$$\det(AB) = \det A \cdot \det B$$
- *Transponierte:*
$$\det(A^T) = \det A$$

5.16 Matrix-Gleichungssysteme: Ein LGS der Form

$$AX = B$$

heißt auch *Matrix-Gleichungssystem*. Dabei sind $A \in \mathbb{R}^{m\times n}$ und $B \in \mathbb{R}^{m\times k}$ gegeben und $X \in \mathbb{R}^{n\times k}$ gesucht. Die Bestimmung der Lösung X erfolgt vollkommen analog zum Lösen linearer Gleichungssystem gemäß Kapitel 2, indem A auf gestaffelte Form gebracht wird. Nun sind auf der rechten Seite aber alle Spalten von B umzuformen, und die Kopfzeile des Lösungsschemas enthält die Zeilenvektoren von X. Die Kriterien für die Lösbarkeit sind vollkommen analog. Insbesondere ist auch ein Matrix-Gleichungssystem mit einer quadratischen Matrix A genau dann eindeutig lösbar, wenn $\det A \neq 0$.

5.17 Beispiel: Gegeben sei das Matrix-Gleichungssystem $AX = B$ mit

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 7 \\ 3 & 1 \\ 3 & 5 \end{bmatrix}.$$

Die Lösung X ist also eine (3×2)-Matrix. Für das Schema

	$\vec{x}_1^T$	$\vec{x}_2^T$	$\vec{x}_3^T$	$\vec{b}^1$	$\vec{b}^2$
$\boxed{1}$	2	0	1	2	7
$\boxed{2}$	1	2	1	3	1
$\boxed{3}$	2	1	2	3	5

liefert der Gauß-Algorithmus

	$\vec{x}_1^T$	$\vec{x}_2^T$	$\vec{x}_3^T$	$\vec{b}^1$	$\vec{b}^2$
$\boxed{1}$	2	0	1	2	7
$\boxed{4}$	0	4	1	4	-5
$\boxed{5}$	0	1	1	1	-2
$\boxed{1}$	2	0	1	2	7
$\boxed{4}$	0	4	1	4	-5
$\boxed{6}$	0	0	3	0	-3

Damit ergibt sich

$$\begin{aligned} \boxed{6} &: 3\vec{x}_3^T = [0, -3] \quad \Rightarrow \quad \vec{x}_3^T = [0, -1] \\ \boxed{4} &: 4\vec{x}_2^T + \vec{x}_3^T = [4, -5] \quad \Rightarrow \quad \vec{x}_2^T = [1, -1] \\ \boxed{1} &: 2\vec{x}_1^T + \vec{x}_3^T = [2, 7] \quad \Rightarrow \quad \vec{x}_1^T = [1, 4] \end{aligned}$$

und schließlich die Lösung

$$X = \begin{bmatrix} 1 & 4 \\ 1 & -1 \\ 0 & -1 \end{bmatrix}.$$

5.18 Einheitsmatrix: Die aus den Einheitsvektoren $\vec{e}_1, \ldots, \vec{e}_n$ [→ 2.26] gebildete $(n \times n)$-Matrix

$$E_n := \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = [\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n]$$

heißt *Einheitsmatrix*. Wenn das Format der Einheitsmatrix aus dem Zusammenhang klar ist, schreiben wir für E_n auch kurz E. Für eine beliebige $(m \times n)$-Matrix A gilt

$$AE_n = E_m A = A.$$

Für die Determinante gilt [→ 4.18] $\det E = 1$.

5.19 Inverse Matrix: Sei A eine $(n \times n)$-Matrix. mit $\det A \neq 0$. Dann ist die Lösung des Matrix-Gleichungssystems

$$AX = E$$

eindeutig bestimmt. Sie wird *inverse Matrix* oder auch kurz *Inverse* von A genannt und mit A^{-1} bezeichnet. Matrizen mit $\det A = 0$ oder nicht-quadratische Matrizen besitzen keine Inverse. Es gilt

- $AA^{-1} = A^{-1}A = E$
- $(A^{-1})^{-1} = A$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^T)^{-1} = (A^{-1})^T$
- $\det A^{-1} = \frac{1}{\det A}$, da $1 = \det E = \det(AA^{-1}) = \det A \cdot \det A^{-1}$ [→ 5.15].

Sei $AX = B$ ein beliebiges Gleichungssystem, dann erhält man nach Multiplikation von links mit A^{-1} die Lösung X,

$$AX = B \quad \Rightarrow \quad A^{-1}AX = A^{-1}B \quad \Rightarrow \quad EX = A^{-1}B \quad \Rightarrow \quad X = A^{-1}B.$$

Die Berechnung der Inversen lohnt sich immer dann, wenn wiederholt Gleichungssysteme mit derselben Matrix A und verschiedenen rechten Seiten gelöst werden müssen.

5.20 Beispiel: Für $n = 2$ gilt

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Im Nenner steht gerade $\det A = ad - bc$, sodass die angegebene Inverse für $\det A \neq 0$ definiert ist.

5.21 Beispiel: Für

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

liefert der Gauß-Algorithmus

	$\vec{x}_1^T$	$\vec{x}_2^T$	$\vec{x}_3^T$	$\vec{e}_1$	$\vec{e}_2$	$\vec{e}_3$
$\boxed{1}$	1	2	1	1	0	0
$\boxed{2}$	1	1	2	0	1	0
$\boxed{3}$	2	1	1	0	0	1
$\boxed{4}$	0	−1	1	−1	1	0
$\boxed{5}$	0	−3	−1	−2	0	1
$\boxed{6}$	0	0	−4	1	−3	1

und damit

$$A^{-1} = X = \frac{1}{4} \begin{bmatrix} -1 & -1 & 3 \\ 3 & -1 & -1 \\ -1 & 3 & -1 \end{bmatrix}.$$

5.22 Orthogonale Matrizen: Eine $(n \times n)$-Matrix A heißt *orthogonal*, wenn

$$A^T A = E.$$

Für orthogonale Matrizen A und B gilt:

- $A^{-1} = A^T$.
- $AA^T = E$.
- A^T ist orthogonal.
- AB ist orthogonal.
- $\det A = \pm 1$, da $1 = \det E = \det(AA^T) = \det A \cdot \det A^T = (\det A)^2$.

Eine Matrix ist genau dann orthogonal, wenn ihre Spaltenvektoren eine *Orthonormalbasis (ONB)* [→ 2.34] bilden, d.h., wenn $\langle \vec{a}^i, \vec{a}^j \rangle = \delta_{i,j}$. Die Spaltenvektoren bilden genau dann eine ONB, wenn auch die Zeilenvektoren eine ONB bilden, d.h., wenn $\langle \vec{a}_i^T, \vec{a}_j^T \rangle = \delta_{i,j}$.

5.23 Beispiel: Für beliebige Winkel φ ist die Matrix

$$A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$$

orthogonal.

5.24 Beispiel: Die Matrix

$$A = \frac{1}{3} \begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix}$$

ist orthogonal.

5.25 Cramersche Regel: Wir betrachten nochmals das lineare Gleichungssystem $A\vec{x} = \vec{b}$ mit einer $n \times n$ Matrix A. Es gilt stets $A\vec{e}_2 = \vec{a}^2, \ldots, A\vec{e}_n = \vec{a}^n$. Also können wir schreiben

$$A[\vec{x}, \vec{e}_2, \ldots, \vec{e}_n] = A_1 \quad \text{wobei} \quad A_1 := [\vec{b}, \vec{a}^2, \ldots, \vec{a}^n].$$

Die $(n \times n)$-Matrix A_1 auf der rechten Seite entsteht also dadurch, dass man die erste Spalte von A durch $\vec{b}$ ersetzt. Der zweite Faktor auf der linken Seite ist ebenfalls eine $(n \times n)$-Matrix, die untere Dreiecksform hat, siehe 4.18. Ihre Determinante ist gleich dem Produkt der Diagonalelemente, also x_1. Man erhält schließlich für die erste Lösungskomponente mit Hilfe der Produktregel die Formel

$$x_1 = \frac{\det A_1}{\det A},$$

sofern $\det A \neq 0$. Bezeichne allgemein A_k die $(n \times n)$-Matrix, die entseht, wenn man die k-te Spalte von A durch $\vec{b}$ ersetzt, dann gilt analog

$$x_k = \frac{\det A_k}{\det A}.$$

Diese sogenannte *Cramersche Regel* ist in der Regel nur dann effizient, wenn man an einzelnen Lösungskomponenten, nicht aber am kompletten Vektor $\vec{x}$ interessiert ist.

6 Lineare Abbildungen

6.1 Lineare Abbildung: Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}^m$ heißt *lineare Abbildung* von $\mathbb{R}^n$ nach $\mathbb{R}^m$, wenn für alle $\vec{x}_1, \vec{x}_2$ und alle $\alpha \in \mathbb{R}$ gilt

$$f(\alpha\vec{x}_1) = \alpha f(\vec{x}_1)$$
$$f(\vec{x}_1 + \vec{x}_2) = f(\vec{x}_1) + f(\vec{x}_2).$$

Im Fall $f : \mathbb{R}^n \to \mathbb{R}^n$ nennen wir f auch eine lineare Abbildung *in* $\mathbb{R}^n$.

6.2 Beispiel: Für einen gegebenen Vektor $\vec{a} \in \mathbb{R}^n$ ist die Abbildung $f : \mathbb{R}^n \to \mathbb{R}$ mit

$$f(\vec{x}) = \langle \vec{a}, \vec{x} \rangle$$

linear.

6.3 Beispiel: Für eine gegebene Matrix $A \in \mathbb{R}^{m\times n}$ ist die Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ mit

$$f(\vec{x}) = A\vec{x}$$

linear.

6.4 Matrixform: Seien $\vec{e}_1, \ldots, \vec{e}_n$ die Einheitsvektoren in $\mathbb{R}^n$ dann kann jeder Vektor $\vec{x} \in \mathbb{R}^n$ in der Form

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{j=1}^{n} x_j \vec{e}_j$$

dargestellt werden [→ 2.26]. Für den Funktionswert der linearen Abbildung $f : \mathbb{R}^n \to \mathbb{R}^m$ an der Stelle $\vec{x}$ gilt dann

$$f(\vec{x}) = f\left(\sum_{j=1}^{n} x_j \vec{e}_j\right) = \sum_{j=1}^{n} x_j f(\vec{e}_j).$$

Er ist also durch die Funktionwerte $f(\vec{e}_1), \ldots, f(\vec{e}_n)$ der Einheitsvektoren vollständig bestimmt. Verwenden wir diese Funktionswerte als Spaltenvektoren einer $(m\times n)$-Matrix A, d.h.,

$$A = [\vec{a}^1, \vec{a}^2, \ldots, \vec{a}^n], \quad \vec{a}^j := f(\vec{e}_j),$$

dann gilt

$$A\vec{x} = A\sum_{j=1}^{n} x_j \vec{e}_j = \sum_{j=1}^{n} x_j A\vec{e}_j = \sum_{j=1}^{n} x_j \vec{a}^j = \sum_{j=1}^{n} x_j f(\vec{e}_j) = f(\vec{x}).$$

Eine lineare Abbildung f kann also stets in der Matrixform $f(\vec{x}) = A\vec{x}$ geschrieben werden. Die Matrix-Abbildungen gemäß Beispiel 6.3 umfassen also tatsächlich die Menge *aller* linearen Abbildungen.

6.5 Beispiel [→ 6.2]: Sei $\vec{a} = [a_1, \ldots, a_n]^T$, dann ist $\langle \vec{a}, \vec{e}_j \rangle = a_j$. Die Matrix A ist also durch den Zeilenvektor

$$A = [f(\vec{e}_1), \ldots, f(\vec{e}_n)] = [a_1, \ldots, a_n] = \vec{a}^T$$

gegeben, $f(\vec{x}) = \vec{a}^T \vec{x}$.

6.6 Summe und Verkettung:

- Sei $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{m \times n}$ sowie $f(\vec{x}) = A\vec{x}$ und $g(\vec{y}) = B\vec{y}$ die zugehörigen linearen Abbildungen. Dann ist die *Summenabbildung* $h := f + g : \mathbb{R}^n \to \mathbb{R}^m$ gegeben durch

  $$h(\vec{y}) = (f+g)(\vec{y}) = f(\vec{y}) + g(\vec{y}) = A\vec{y} + B\vec{y} = (A+B)\vec{y}.$$

  Die Matrizenaddition entspricht also der Summe der zugehörigen linearen Abbildungen.

- Sei $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{n \times k}$ sowie $f(\vec{x}) = A\vec{x}$ und $g(\vec{y}) = B\vec{y}$. Dann ist die *verkettete Abbildung* $h := f \circ g : \mathbb{R}^k \to \mathbb{R}^m$ gegeben durch

  $$h(\vec{y}) = f(g(\vec{y})) = f(B\vec{y}) = AB\vec{y}.$$

  Die Matrizenmultiplikation entspricht also der Verkettung der zugehörigen linearen Abbildungen.

6.7 Fixpunkt: Sei A eine $(n \times n)$-Matrix. Ein Punkt $\vec{v} \in \mathbb{R}^n$ heißt *Fixpunkt* von A, wenn $A\vec{v} = \vec{v}$. Die Menge aller Fixpunkte ist ein linearer Teilraum von $\mathbb{R}^n$. Er wird mit

$$\operatorname{fix} A := \{\vec{v} \in \mathbb{R}^n : A\vec{v} = \vec{v}\}$$

bezeichnet. Zur Bestimmung von Fixpunkten schreibt man

$$A\vec{v} = \vec{v} \quad \Rightarrow \quad A\vec{v} = E\vec{v} \quad \Rightarrow \quad (A-E)\vec{v} = \vec{0}.$$

Es gilt also

$$\operatorname{fix} A = \ker(A - E).$$

6.8 Beispiel [→ 5.24]: Es gilt

$$A = \frac{1}{3}\begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix} \quad \Rightarrow \quad A - E = \frac{1}{3}\begin{bmatrix} -1 & 2 & -1 \\ -1 & -1 & 2 \\ 2 & -1 & -1 \end{bmatrix}$$

und die Lösung der Fixpunktgleichung $(A-E)\vec{v} = \vec{0}$ ist die Fixpunktgerade

$$\operatorname{fix} A = \left\{ t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}.$$

6.9 Spezielle lineare Abbildungen in $\mathbb{R}^n$:

- Die *Streckung* $f(\vec{x}) = \lambda\vec{x}$ mit Streckfaktor $\lambda \in \mathbb{R}$ ist gegeben durch $A = \lambda E$. Insbesondere erhält man für $\lambda = 1$ die identische Abbildung $f(\vec{x}) = \vec{x}$, für $\lambda = 0$ die Nullabbildung $f(\vec{x}) = \vec{0}$ und für $\lambda = -1$ die Punktspiegelung $f(\vec{x}) = -\vec{x}$.

- Eine lineare Abbildung $f(\vec{x}) = R\vec{x}$ heißt *Drehung*, wenn R orthogonal ist und $\det R = 1$ gilt. Drehungen sind *normerhaltend*, d.h., $\|R\vec{x}\| = \|\vec{x}\|$, denn

$$\|R\vec{x}\|^2 = (R\vec{x})^T \cdot (R\vec{x}) = \vec{x}^T(R^T R)\vec{x} = \vec{x}^T E\vec{x} = \vec{x}^T\vec{x} = \|\vec{x}\|^2.$$

  – In $\mathbb{R}^2$ ist eine Drehung um den Ursprung um den Winkel φ gegeben durch [→ 5.23]

$$R_\varphi := \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}.$$

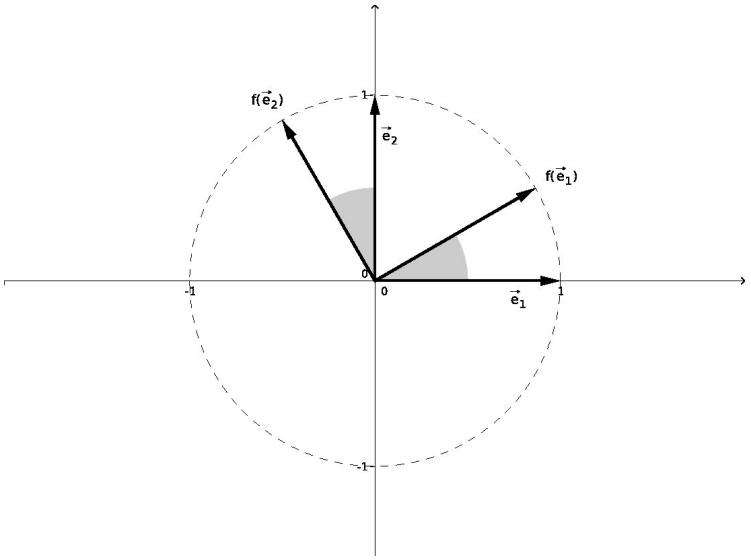


Abbildung 13: Eine Drehung um den (grau schattierten) Winkel φ in $\mathbb{R}^2$

  – In $\mathbb{R}^3$ ist eine Drehung um die z-Achse um den Winkel φ gegeben durch

$$R^z_\varphi := \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

  Für allgemeine Drehmatrizen in $\mathbb{R}^3$ ist die Drehachse durch die Fixpunktgerade gegeben. Der Drehwinkel φ bestimmt sich gemäß der Formel

$$2\cos\varphi + 1 = \operatorname{spur} R,$$

  wobei $\operatorname{spur} R := r_{1,1} + r_{2,2} + r_{3,3}$ die Summe der Diagonalelemente von R ist.

- Eine lineare Abbildung $f(\vec{x}) = P\vec{x}$ heißt *Projektion*, wenn

$$P^2 = P$$

gilt. In diesem Fall ist jeder Bildpunkt $\vec{v} = P\vec{x}$ ein Fixpunkt von P, denn

$$P\vec{v} = P(P\vec{x}) = P^2\vec{x} = P\vec{x} = \vec{v}.$$

Das heißt, jeder Punkt $\vec{x}$ wird durch einmalige Anwendung der Abbildung f auf die Menge $\operatorname{fix} P$ abgebildet und bleibt bei weiteren Anwendungen der Abbildung dann unverändert. Die Projektion heißt *senkrecht*, wenn

$$\langle P\vec{x} - \vec{x}, P\vec{x}\rangle = 0$$

für alle $\vec{x} \in \mathbb{R}^n$. Eine Projektion ist genau dann senkrecht, wenn P symmetrisch ist.

- Die senkrechte Projektion auf die Ursprungsgerade $g : t\vec{v}$, $t \in \mathbb{R}$, ist durch

  $$P_g := \frac{\vec{v} \cdot \vec{v}^T}{\|\vec{v}\|^2}$$

  gegeben. Insbesondere ist $d(\vec{x}, g) = \|P_g\vec{x} - \vec{x}\|$ der Abstand des Punktes $\vec{x}$ von der Geraden g [→ 2.12].
- Sei $M : \langle \vec{x}, \vec{n}\rangle = 0$ eine implizit gegebene Menge, also z.B. eine Gerade in $\mathbb{R}^2$ oder eine Ebene in $\mathbb{R}^3$. Die senkrechte Projektion auf M ist durch

  $$P_M := E - \frac{\vec{n} \cdot \vec{n}^T}{\|\vec{n}\|^2}$$

  gegeben. Insbesondere ist $d(\vec{x}, M) = \|P_M\vec{x} - \vec{x}\|$ der Abstand des Punktes $\vec{x}$ von der Menge M [→ 2.15], [→ 2.21].

- Eine lineare Abbildung $g(\vec{x}) = S\vec{x}$ heißt *Spiegelung*, wenn

  $$S^2 = E$$

  gilt. Zweimaliges Spiegeln führt also auf den Ausgangspunkt zurück. Wenn P eine Projektion ist, dann ist

  $$S := 2P - E$$

  eine Spiegelung an der Menge $\operatorname{fix} P = \operatorname{fix} S$, denn

  $$S^2 = (2P - E) \cdot (2P - E) = 4P^2 - 2PE - 2EP + E^2 = E$$

  und

  $$S\vec{v} = \vec{v} \quad \Leftrightarrow \quad 2P\vec{v} - \vec{v} = \vec{v} \quad \Leftrightarrow \quad P\vec{v} = \vec{v}.$$

  Umgekehrt ist für eine Spielgelung S die Matrix

  $$P := \frac{1}{2}(S + E)$$

  eine Projektion, denn

  $$P^2 = \frac{1}{4}(S + E) \cdot (S + E) = \frac{1}{4}(S^2 + SE + ES + E^2) = \frac{1}{4}(2S + 2E) = P.$$

– Die senkrechte Spiegelung an der Geraden $g : t\vec{v}$, $t \in \mathbb{R}$, ist gegeben durch

$$S_g := 2P_g - E = 2\frac{\vec{v} \cdot \vec{v}^T}{\|\vec{v}\|^2} - E.$$

– Die senkrechte Spiegelung an der Menge $M : \langle \vec{x}, \vec{n} \rangle = 0$ ist gegeben durch

$$S_M := 2P_M - E = E - 2\,\frac{\vec{n} \cdot \vec{n}^T}{\|\vec{n}\|^2}.$$

Diese Abbildung wird auch *Householder-Transformation* genannt.

6.10 Beispiel [→ 5.24]: Die Matrix

$$R = \frac{1}{3}\begin{bmatrix} 2 & 2 & -1 \\ -1 & 2 & 2 \\ 2 & -1 & 2 \end{bmatrix}$$

ist orthogonal und es gilt $\det R = 1$. Also ist R eine Drehung. Gemäß Beispiel 6.8 ist die Drehgerade gegeben durch $g : t[1,1,1]^T$, $t \in \mathbb{R}$, und für den Drehwinkel gilt

$$2\cos\varphi + 1 = 2 \quad \Rightarrow \quad \cos\varphi = 1/2 \quad \Rightarrow \quad \varphi = \pm\pi/3.$$

Das Vorzeichen des Drehwinkels hängt davon ab, aus welcher Richtung man auf die Drechachse schaut.

6.11 Beispiel: Sei $\vec{v} = [3,1]^T$. Die senkrechte Projektion auf die Gerade $g : t\vec{v}, t \in \mathbb{R}$, ist

$$P_g = \frac{1}{10}\begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix}$$

und die senkrechte Spiegelung an der Geraden ist

$$S_g = 2P_g - E = \frac{1}{5}\begin{bmatrix} 4 & 3 \\ 3 & -4 \end{bmatrix}.$$

6.12 Beispiel: Sei $\vec{n} = [1,2,-1]^T$. Die senkrechte Projektion auf die Ebene $M : \langle \vec{x}, \vec{n} \rangle = 0$ ist

$$P_M = \frac{1}{6}\begin{bmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \\ 1 & 2 & 5 \end{bmatrix}$$

und die Spiegelung an der Ebene ist

$$S_M = 2P_M - E = \frac{1}{3}\begin{bmatrix} 2 & -2 & 1 \\ -2 & -1 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

6.13 Abbildungsmatrix bezüglich beliebiger Basis: Es sei $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^n$ eine Basis von $\mathbb{R}^n$. Mit $\vec{x}_V = [\alpha_1, \ldots, \alpha_n]^T$ sei der V–Koordinatenvektor von $\vec{x}$ bezeichnet, d.h. [→ 2.32]

$$\vec{x} = \alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n.$$

Es sei $f : \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung. Schreiben wir die V–Koordinatenvektoren der Funktionswerte $f(\vec{v}_j)$ als Spalten in eine Matrix, so erhalten wir

$$\tilde{A} = [\vec{a}_V^1, \vec{a}_V^2, \ldots, \vec{a}_V^n] \in \mathbb{R}^{n \times n}, \quad \vec{a}_V^j := f(\vec{v}_j)_V.$$

Die Matrix $\tilde{A}$ heißt *Abbildungsmatrix zu f bezüglich* $\vec{v}_1, \ldots, \vec{v}_n$. Im Fall der Einheitsvektoren $\vec{v}_i = \vec{e}_i$ ist $\tilde{A}$ die Matrix A aus [→ 6.4].
Die Abbildungsmatrix $\tilde{A}$ und f stehen in folgendem Zusammenhang:

$$f(\vec{x})_V = \tilde{A}\vec{x}_V.$$

Wenn also $\vec{x}_V$ der Koordinatenvektor von $\vec{x}$ bezüglich $\vec{v}_1, \ldots, \vec{v}_n$ ist, so ist $\tilde{A}\vec{x}_V$ der Koordinatenvektor von $f(\vec{x})$ bezüglich $\vec{v}_1, \ldots, \vec{v}_n$.
Ebenso wie die Matrix A beschreibt auch die Matrix $\tilde{A}$ die Abbildung f eindeutig. Durch geeignete Wahl der Basis $\vec{v}_1, \ldots, \vec{v}_n$ kann man erreichen, dass die Matrix $\tilde{A}$ sehr leicht berechnet werden kann.

6.14 Beispiel: Die Abbildung $f : \mathbb{R}^3 \to \mathbb{R}^3$ beschreibe die Drehung um $\varphi = \frac{\pi}{3}$ um die Achse $t\vec{a}, t \in \mathbb{R}$, wobei $\vec{a} = [-1, 1, 1])$. Die Matrix zu f bezüglich der Standardbasis $\vec{e}_1, \vec{e}_2, \vec{e}_3$ des $\mathbb{R}^3$ ist $A = [f(\vec{e}_1), f(\vec{e}_2), f(\vec{e}_3)]$. Allerdings sind die Vektoren $f(\vec{e}_i)$ nicht leicht zu berechnen.

Eine passendere Basis findet sich wie folgt: Wähle $\vec{v}_1$ auf der Drehachse, wähle $\vec{v}_2$ orthogonal zu $\vec{v}_1$ und bestimme schließlich $\vec{v}_3$ also Kreuzprodukt von $\vec{v}_1$ und $\vec{v}_2$. Normiert man diese Vektoren, dann erhält man die ONB $V = [\vec{v}_1, \vec{v}_2, \vec{v}_3]$. Konkret liefert dieses Vorgehen

$$\vec{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \vec{v}_3 = \vec{v}_1 \times \vec{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix},$$

wobei der Vektor $\mathbf{v}_2$ natuerlich nicht eindeutig festgelegt ist. Es gilt $f(\vec{v}_1) = \vec{v}_1$ und [→ 6.9]

$$f(\vec{v}_2) = \cos(\varphi)\vec{v}_2 + \sin(\varphi)\vec{v}_3 = \frac{1}{2}\vec{v}_2 + \frac{\sqrt{3}}{2}\vec{v}_3$$
$$f(\vec{v}_3) = -\sin(\varphi)\vec{v}_2 + \cos(\varphi)\vec{v}_3 = -\frac{\sqrt{3}}{2}\vec{v}_2 + \frac{1}{2}\vec{v}_3.$$

Die Koordinaten von $f(\vec{v}_j)$ bezüglich der Basis $\vec{v}_1, \vec{v}_2, \vec{v}_3$ können hier abgelesen werden. Sie ergeben die Spalten von $\tilde{A}$:

$$\tilde{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}.$$

Für einen beliebigen Vektor $\vec{x} \in \mathbb{R}^3$ lässt sich $f(\vec{x})$ nun in folgenden Schritten berechnen: Berechne zuerst die V-Koordianten $\vec{x}_V$ von $\vec{x}$. Berechne dann den Bildvektor $\vec{y}_V = \tilde{A}\vec{x}_V$. Dies ist der V-Koordinatenvektor von $\vec{y} = f(\vec{x})$, aus dem dann $\vec{y}$ berechnet werden kann. Dieses Vorgehen wollen wir im nächsten Abschnitt genauer untersuchen.

6.15 Basiswechsel: Sei $V = [\vec{v}_1, \ldots, \vec{v}_n]$ eine Basis von $\mathbb{R}^n$. Wir haben diese Schreibweise schon in 2.32 verwendet, erkennen nun aber V als die $(n \times n)$-Matrix, deren Spalten durch die Vektoren $\vec{v}_j$ gegeben sind. Es gilt

$$\vec{x} = \alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n = V \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = V\vec{x}_V.$$

Die Matrix V rechnet also V-Koordinaten in Standardkoordinaten um. Da V aufgrund der Basiseigenschaft invertierbar ist, kann man diesen Prozess auch umkehren: Man erhält die V-Koordinaten aus den Standardkoordinaten durch Multiplikation mit der Inversen,

$$\vec{x} = V\vec{x}_V \quad \Leftrightarrow \quad \vec{x}_V = V^{-1}\vec{x}_V.$$

Sei nun $f : \mathbb{R}^n \to \mathbb{R}^n$ eine lineare Abbildung, die in Standardkoordinaten durch die Matrix A gegeben ist, also $\vec{y} = f(\vec{x}) = A\vec{x}$. Dann gilt für die V-Koordinaten des Bildpunkts

$$\vec{y}_V = V^{-1}\vec{y} = V^{-1}A\vec{x} = V^{-1}AV\vec{x}_V.$$

In V-Koordinaten wird die Abbildung f deshalb beschrieben durch

$$\vec{y}_V = \tilde{A}\vec{x}_V, \quad \tilde{A} = V^{-1}AV.$$

Kennt man umgekehrt die Matrix $\tilde{A}$, so erhält man die Darstellung von f bezüglich der Standardbasis gemäß

$$\vec{y} = A\vec{x}, \quad A = V\tilde{A}V^{-1}.$$

Einfacher wird die Umrechnung im Fall einer Orthonormalbasis [→ 5.22]: Bilden $v_1, \ldots, v_n$ eine ONB , dann ist die Matrix V orthogonal [→ 5.22] und $V^{-1} = V^T$. Es gilt dann

$$A = V\tilde{A}V^T \quad \text{bzw.} \quad \tilde{A} = V^TAV.$$

6.16 Beispiel [→ 6.14]: Es gilt

$$V = \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{bmatrix}.$$

Da $\vec{v}_1, \vec{v}_2, \vec{v}_3$ eine ON–Basis bilden, ist V orthogonal. Also

$$A = V\tilde{A}V^T = \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ -2 & -1 & 2 \end{bmatrix}.$$

6.17 Beispiel [→ 6.11]: Sei $\vec{v}_1 = [3,1]^T$ und $\vec{v}_2 = [1,2]$, dann ist

$$V = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}, \quad V^{-1} = \frac{1}{5} \begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix}$$

Die V-Koordinaten des Vektors $\vec{x} = [9,8]^T$ sind gegeben durch

$$\vec{x}_V = V^{-1}\vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.$$

Man könnte auch schreiben $[9,8]_V^T = [2,3]^T$. Dieser Zusammenhang wird in Abbildung 14 veranschaulicht. Die Projektion P_g und die Spiegelung S_g haben in V-Koordinaten die

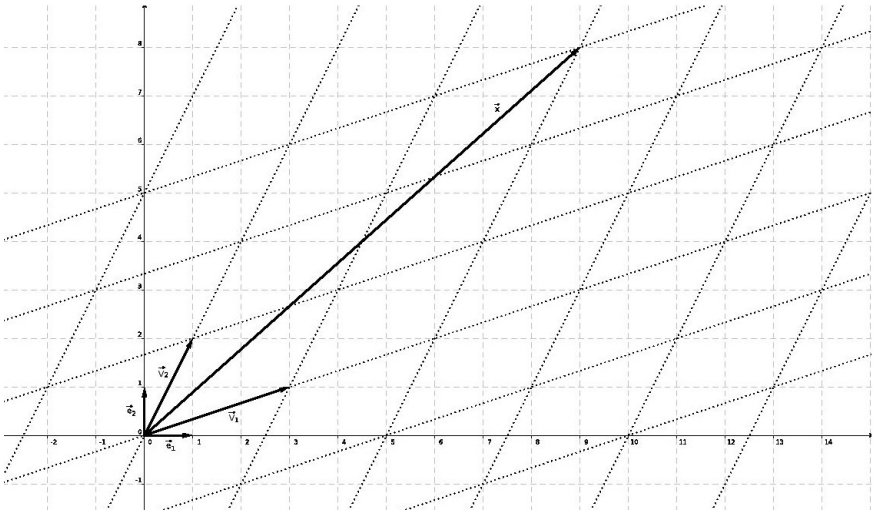


Abbildung 14: Das gestrichelte Raster gehört zur Standardbasis, das gepunktete Raster zur Basis V. Den Punkt $\vec{x} = [9,8]^T$ kann man geometrisch erkennen, in dem man im gepunkteten Raster zwei Schritte in Richtung $\vec{v}_1$ und drei Schritte in Richtung $\vec{v}_2$ geht.

Form

$$\tilde{P}_g = V^{-1}P_gV = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}, \quad \tilde{S}_g = V^{-1}S_gV = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}.$$

6.18 Beispiel [→ 6.10]: Die Matrix

$$V = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & 1 & \sqrt{2} \\ -\sqrt{3} & 1 & \sqrt{2} \\ 0 & -2 & \sqrt{2} \end{bmatrix}$$

ist orthogonal, $VV^T = E$. Die Drehmatrix R geht durch Basiswechsel über in

$$\tilde{R} = V^T RV = \begin{bmatrix} 1/2 & \sqrt{3}/2 & 0 \\ -\sqrt{3}/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Mit $\varphi = -\pi/3$ ist dies gerade die in 6.9 angegebene Form einer Drehung um die dritte Koordinatenachse. Diese ist hier durch den dritten Basisvektor $\vec{v}_3 = \sqrt{3}/3\,[1,1,1]^T$ gegeben und stimmt also mit der zuvor bestimmten Drehachse überein.

6.19 Beispiel [→ 6.18]: Sei V wie zuvor und

$$f(\vec{x}) := \vec{x} \times [1,1,1]^T.$$

Die Matrixform von f ist

$$f(\vec{x}) = [f(\vec{e}_1), f(\vec{e}_2), f(\vec{e}_3)]\vec{x} = A\vec{x}, \quad A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$$

und man erhält

$$\tilde{A} = V^T A V = \begin{bmatrix} 0 & \sqrt{3} & 0 \\ -\sqrt{3} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Diese Matrix lässt sich in der Form $\tilde{A} = \tilde{A}_3 \cdot \tilde{A}_2 \cdot \tilde{A}_1$ in Faktoren zerlegen, wobei

$$\tilde{A}_1 := \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \tilde{A}_2 := \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \tilde{A}_3 := \sqrt{3}E.$$

Die Abbildung $\tilde{A}_1$ ist eine Projektion in die $\vec{v}_1\vec{v}_2$-Ebene, $\tilde{A}_2$ ist eine Drehung um die $\vec{v}_3$-Achse um den Winkel $-\pi/2$, und $\tilde{A}_3$ ist eine Streckung um den Faktor $\sqrt{3}$. Die Abbildung f lässt sich also als Verkettung einer Projektion, einer Drehung und einer Streckung deuten.

7 Eigenwerte und -vektoren

7.1 Fixgeraden: Eine Ursprungsgerade $g : t\vec{v}, t \in \mathbb{R}$, heißt *Fixgerade* der linearen Abbildung $f(\vec{x}) = A\vec{x}$, wenn jeder Punkt auf g wieder auf einen Punkt auf g abgebildet wird. Man beachtge, dass anders als bei einer Fix*punkt*geraden hier Punkte auf der Geraden nicht auf sich selbst abgebildet werden müssen. Betrachten wir das Bild $f(\vec{v}) = A\vec{v}$ des Richtungsvektors. Da es wieder auf der Geraden g liegt, muss es eine Zahl $\lambda \in \mathbb{R}$ geben, sodass

$$A\vec{v} = \lambda\vec{v}.$$

Wenn dies der Fall ist, dann gilt natürlich für jeden anderen Punkt auf der Geraden die analoge Beziehung $A(t\vec{v}) = \lambda(t\vec{v})$. Der Faktor λ gibt also das Streckverhältnis zwischen Bild und Urbild für Punkte auf der Fixgeraden an. Insbesondere handelt es sich um eine Fixpunktgerade, wenn $\lambda = 1$.

7.2 Eigenwert und -vektor: Sei A eine $(n \times n)$-Matrix. Ein Vektor $\vec{v} \neq \vec{0}$ heißt *Eigenvektor* von A zum *Eigenwert* $\lambda \in \mathbb{R}$, wenn

$$A\vec{v} = \lambda\vec{v}.$$

Wenn $\vec{v}$ ein Eigenvektor von A ist, dann ist auch jedes Vielfache $t\vec{v}$ ein Eigenvektor, sofern $t \neq 0$. Zur Bestimmung von Eigenwerten und Vektoren betrachtet man analog zur Bestimmung von Fixpunkten [→ 6.7] die Gleichung

$$(A - \lambda E)\vec{v} = \vec{0}. \tag{7.1}$$

Die Lösbarkeit dieses LGS hängt von der Determinante

$$p(\lambda) := \det(A - \lambda E)$$

ab. Die Funktion p ist ein Polynom vom Grad n in der Variablen λ und wird als das *charakteristische Polynom* von A bezeichnet. Betrachten wir nun einen festen Wert λ.

- Falls $p(\lambda) \neq 0$, dann ist das LGS (7.1) eindeutig lösbar und es folgt $\vec{v} = \vec{0}$. Folglich ist $\vec{v}$ kein Eigenvektor und λ auch kein Eigenwert.
- Falls $p(\lambda) = 0$, dann besitzt (7.1) nichttriviale Lösungen. Es gibt also einen Vektor $\vec{v} \neq \vec{0}$ mit $(A - \lambda E)\vec{v} = \vec{0}$ und dies ist gerade ein Eigenvektor zum Eigenwert λ.

Die Eigenwerte der Matrix A sind also die Nullstellen des charakteristischen Polynoms $p(\lambda) = \det(A - \lambda E)$. Die zugehörigen Eigenvektoren bestimmen sich aus dem LGS $(A - \lambda E)\vec{v} = \vec{0}$.

7.3 Beispiel: Für

$$A = \begin{bmatrix} 5 & -9 & -3 \\ 4 & -9 & -4 \\ -6 & 15 & 8 \end{bmatrix}$$

ist das charakteristische Polynom

$$p(\lambda) = \det \begin{bmatrix} (5-\lambda) & -9 & -3 \\ 4 & (-9-\lambda) & -4 \\ -6 & 15 & (8-\lambda) \end{bmatrix} = -\lambda^3 + 4\lambda^2 - \lambda - 6.$$

Wenn es ganzzahlige Nullstellen gibt, so müssen sie Teiler des konstanten Terms -6 sein. Deshalb überprüft man zunächst die Teiler von -6: Es ist $\lambda = 1$ beispielsweise keine Nullstelle, aber $\lambda_1 := 2$ liefert $p(\lambda_1) = 0$. Nun teilt man p durch den Linearfaktor $(2 - \lambda)$ und erhält

$$q(\lambda) := p(\lambda) : (2 - \lambda) = \lambda^2 - 2\lambda - 3.$$

Die beiden restlichen Eigenwerte von A sind Nullstellen von q,

$$q(\lambda) = 0 \quad \Rightarrow \quad \lambda_2 := 3, \quad \lambda_3 := -1.$$

Insgesamt erhält man also die Faktorisierung

$$p(\lambda) = (2 - \lambda)(3 - \lambda)(-1 - \lambda).$$

Der Eigenvektor zu λ_1 ergibt sich aus

$$(A - \lambda_1 E)\vec{v}_1 = \begin{bmatrix} 3 & -9 & -3 \\ 4 & -11 & -4 \\ -6 & 15 & 6 \end{bmatrix} \vec{v}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \vec{v}_1 := \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Analog erhält man

$$\vec{v}_2 := \begin{bmatrix} 0 \\ 1 \\ -3 \end{bmatrix} \quad \text{und} \quad \vec{v}_3 := \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}.$$

Natürlich können auch beliebige Vielfache der angegebenen Eigenvektoren verwendet werden, z.B. $\vec{v}_2 = [0, -1, 3]^T$ oder $\vec{v}_3 = [7, 7, -7]^T$.

7.4 Regeln:

- Es gibt höchstens n verschiedene reelle Eigenwerte einer $n \times n$-Matrix A; mit Vielfachheiten gezählt gibt es genau n komplexe Eigenwerte von A.

- Wenn $\vec{v}_1$ und $\vec{v}_2$ Eigenvektoren von A zum *selben* Eigenwert λ sind, dann ist auch $t_1\vec{v}_1 + t_2\vec{v}_2$ ein Eigenvektor von A zum Eigenwert λ, sofern $t_1\vec{v}_1 + t_2\vec{v}_2 \neq \vec{0}$. Die Menge aller Eigenvektoren von A zum Eigenwert λ zusammen mit dem Nullvektor bildet einen linearen Teilraum von $\mathbb{R}^n$. Man nennt ihn *Eigenraum*, kurz $\mathrm{Eig}(A, \lambda)$.

- Wenn $\vec{v}$ ein Eigenvektor von A zum Eigenwert λ ist, dann ist $\vec{v}$ ein Eigenvektor von A^k zum Eigenwert λ^k. Falls A invertierbar ist, darf dabei auch $k = -1$ gewählt werden, d.h., die Eigenwerte der Inversen sind die Kehrwerte der Eigenwerte der gegebenen Matrix.

- Die Summe aller (auch der komplexen) Eigenwerte ist gleich der Summe der Diagonalelemente von A,

  $$\mathrm{spur}\, A := a_{1,1} + \cdots + a_{n,n} = \lambda_1 + \cdots + \lambda_n.$$

  Im Beispiel ist $\mathrm{spur}\, A = 5 - 9 + 8 = 4$ und $\lambda_1 + \lambda_2 + \lambda_3 = 2 + 3 - 1 = 4$.

- Das Produkt aller Eigenwerte (auch der komplexen) ist gleich der Determinante von A,

  $$\det A = \lambda_1 \cdots \lambda_n.$$

  Das heißt, eine Matrix ist genau dann invertierbar, wenn alle Eigenwerte von Null verschieden sind. Im Beispiel ist $\det A = -6$ und $\lambda_1 \cdot \lambda_2 \cdot \lambda_3 = 2 \cdot 3 \cdot (-1) = -6$.

- Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.
- Die Eigenwerte (aber nicht die Eigenvektoren!) von A und A^T stimmen überein.

7.5 Spezialfälle:

- Wenn A eine obere oder untere Dreiecksmatrix ist [→ 4.18], dann stimmen die Eigenwerte mit den Diagonalelementen überein,

  $$p(\lambda) = (a_{1,1} - \lambda)(a_{2,2} - \lambda) \cdots (a_{n,n} - \lambda), \quad \lambda_1 = a_{1,1},\ \lambda_2 = a_{2,2}, \ldots, \lambda_n = a_{n,n}.$$

  Insbesondere sind die Eigenwerte der Einheitsmatrix $\lambda_1 = \cdots = \lambda_n = 1$.
- Wenn A symmetrisch ist, dann sind alle Eigenwerte reell und die zugehörigen Eigenvektoren können durch geeignete Normierung so gewählt werden, dass sie eine Orthonormalbasis bilden.
- Wenn ein Eigenwert eine k-fache Nullstelle des charakteristischen Polynoms ist, dann gibt es hierzu mindestens einen und höchstens k linear unabhängige Eigenvektoren.

7.6 Beispiel:

$$A = \begin{bmatrix} 3 & 2 & -8 \\ 0 & 7 & 1 \\ 0 & 0 & -5 \end{bmatrix} \Rightarrow\ p(\lambda) = (3-\lambda)(7-\lambda)(-5-\lambda) \Rightarrow\ \lambda_1 = 3,\ \lambda_2 = 7,\ \lambda_3 = -5.$$

7.7 Beispiel:

$$A = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 2 & -1 \\ -1 & 0 & 1 \end{bmatrix} \Rightarrow\ p(\lambda) = (1-\lambda)(2-\lambda)^2 \Rightarrow\ \lambda_1 = 1,\ \lambda_2 = \lambda_3 = 2.$$

Ein Eigenvektor zu $\lambda_1 = 1$ ist $\vec{v}_1 = [0, 1, 1]^T$. Zu der doppelten Nullstelle $\lambda_2 = \lambda_3 = 2$ ergibt sich das Gleichungssystem

$$\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{bmatrix} \vec{v} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Dieses hat zwei linear unabhängige Lösungen, z.B. $\vec{v}_2 = [-1, 0, 1]^T$ und $\vec{v}_3 = [0, 1, 0]^T$.

7.8 Beispiel:

$$A = \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow\ p(\lambda) = (1-\lambda)^3 \Rightarrow\ \lambda_1 = \lambda_2 = \lambda_3 = 1.$$

Zu der dreifachen Nullstelle $\lambda_1 = \lambda_2 = \lambda_3 = 1$ ergibt sich das Gleichungssystem

$$\begin{bmatrix} 0 & 3 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \vec{v} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Dieses hat nur eine linear unabhängige Lösung, $\vec{v} = [1, 0, 0]^T$. In Fällen wie diesem können anstelle der fehlenden Eigenvektoren sogenannte *Hauptvektoren* bestimmt werden (in der Literatur zu finden unter dem Stichwort *Jordan-Form*).

7.9 Diagonalisierung: Wenn es zu einer $(n\times n)$-Matrix A genau n linear unabhängige Eigenvektoren $\vec{v}_1,\dots,\vec{v}_n$ gibt, dann kann man diese zu einer $(n\times n)$-Matrix V zusammenfassen. Mit der Diagonalmatrix

$$D=\operatorname{diag}(\lambda_1,\dots,\lambda_n)=\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}$$

erhält man dann

$$AV=VD \quad\text{bzw.}\quad D=V^{-1}AV \quad\text{bzw.}\quad A=VDV^{-1}.$$

Man sagt dann, dass A *diagonalisierbar* ist. Bezüglich der Basis aus Eigenvektoren [→ 6.15] ist A also ähnlich zu einer Diagonalmatrix, $\tilde{A}=D$.
Eine weitere Anwendung der Diagonalierung ergibt sich bei dem Problem, Potenzen von Matrizen effizient zu berechnen. Man erhält

$$A^k=A\cdots A=(VDV^{-1})\cdots(VDV^{-1})=VD^kV^{-1},$$

wobei $D^k=\operatorname{diag}(\lambda_1^k,\lambda_2^k,\dots,\lambda_n^k)$.

7.10 Beispiel [→ 7.3]: Für

$$A=\begin{bmatrix}5 & -9 & -3\\ 4 & -9 & -4\\ -6 & 15 & 8\end{bmatrix}$$

erhält man

$$V=\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 1 & -3 & -1\end{bmatrix},\quad V^{-1}=\begin{bmatrix}2 & -3 & -1\\ 1 & -2 & -1\\ -1 & 3 & 1\end{bmatrix},\quad D=\begin{bmatrix}2 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & -1\end{bmatrix}.$$

Damit ist

$$A^5=VD^5V^{-1}=\begin{bmatrix}1 & 0 & 1\\ 0 & 1 & 1\\ 1 & -3 & -1\end{bmatrix}\cdot\begin{bmatrix}2^5 & 0 & 0\\ 0 & 3^5 & 0\\ 0 & 0 & (-1)^5\end{bmatrix}\cdot\begin{bmatrix}2 & -3 & -1\\ 1 & -2 & -1\\ -1 & 3 & 1\end{bmatrix}$$

$$=\begin{bmatrix}65 & -99 & -33\\ 244 & -489 & -244\\ -666 & 1365 & 698\end{bmatrix}.$$

7.11 Quadratische Gleichungen in $\mathbb{R}^n$: Wir wollen nun geometrische Figuren beschreiben wie z.B. Kreise, Kugeln, Ellipsen, Ellipsoide, Paraboloide, Hyperboloide, ...
Diese ergeben sich als Lösungsmengen von *quadratischen* Gleichungen, wohingegen die bisher betrachteten Geraden, Ebenen usw. Lösungsmengen von *linearen* Gleichungen waren.
Die allgemeine quadratische Gleichung in n Variablen lautet

$$\vec{x}^T\cdot A\cdot\vec{x}+\vec{a}^T\cdot\vec{x}+\alpha=0, \tag{Q}$$

und hierbei haben wir die folgenden gegebenen Koeffizienten: eine symmetrische Matrix $A\in\mathbb{R}^{n\times n}$, einen Spaltenvektor $\vec{a}\in\mathbb{R}^n$, und eine reelle Zahl α. Gesucht sind alle Spaltenvektoren $\vec{x}\in\mathbb{R}^n$, die (Q) lösen. Diese Lösungsmenge nennt man dann *Quadrik* oder im ebenen Fall $n=2$ auch *Kegelschnitt*.

7.12 Beispiel: Im $\mathbb{R}^2$ betrachten wir

$$Q(x,y) = x^2 - 2y^2 + 4xy + \sqrt{5}x + 2\sqrt{5}y - \frac{3}{4} = 0.$$

Wir führen die Schreibweise $\vec{x} = \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ ein und können dann die Gleichung $Q(x,y) = 0$ auf folgende Form bringen:

$$\begin{bmatrix} x & y \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \sqrt{5} & 2\sqrt{5} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} - \frac{3}{4} = 0,$$

also bekommen wir die Gleichung (Q) mit

$$A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}, \qquad \vec{a} = \begin{bmatrix} \sqrt{5} \\ 2\sqrt{5} \end{bmatrix}, \qquad \alpha = -\frac{3}{4}.$$

Die uns interessierende Frage ist jetzt: was beschreibt diese Gleichung geometrisch?
Um die geometrische Gestalt einer Quadrik besser verstehen zu können, werden wir zu einem besser geeigneten Koordinatensystem übergehen. Das Verfahren dazu besteht aus mehreren Schritten und verläuft wie folgt:

7.13 Transformation auf Normalform:

Schritt 1: diagonalisiere A (Hauptachsentransformation). Nach Voraussetzung ist die Matrix A symmetrisch. Gemäß 7.5 existiert eine ONB $\vec{v}_1, \ldots, \vec{v}_n$ aus Eigenvektoren von A. Stellt man diese zur Matrix $V = [\vec{v}_1, \ldots, \vec{v}_n]$ zusammen, so ergibt dieses eine orthogonale Matrix und es gilt [→ 7.9]

$$V^T A V = V^{-1} A V = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{bmatrix} =: A',$$

wobei $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A sind.

Schritt 2: substituiere in (Q). Wir führen neue Koordinaten $\vec{x}'$ ein gemäß $\vec{x} = V\vec{x}'$. Wenn wir dies in (Q) einsetzen, dann folgt

$$\begin{aligned} & (V\vec{x}')^T A V \vec{x}' + \vec{a}^T (V\vec{x}') + \alpha = 0 \\ \iff \quad & \vec{x}'^T V^T A V \vec{x}' + (V^T \vec{a})^T \vec{x}' + \alpha = 0 \\ \iff \quad & \vec{x}'^T A' \vec{x}' + \vec{a}'^T \vec{x}' + \alpha = 0, \end{aligned}$$

wobei wir $\vec{a}' := V^T \vec{a}$ gesetzt haben.

Der Nutzen dieses Schrittes besteht darin, dass in dieser quadratischen Gleichung jetzt keine gemischten Produkte mehr enthalten sind, denn A' ist eine Diagonalmatrix.

Schritt 3: verschiebe den Ursprung. Durch quadratische Ergänzung kann nun für alle $i \in \{1, \ldots, n\}$ mit $\lambda_i \neq 0$ der Summand $a'_i x_i$ zum Verschwinden gebracht werden. Das liefert einen erneuten Koordinatenwechsel von $\vec{x}'$ zu $\vec{x}''$. Hierbei wird der Ursprung des Koordinatensystems verschoben.

Das Ergebnis ist dann die *Normalform*

$$\vec{x}''^T A' \vec{x}'' + \vec{a}''^T \vec{x}'' + \alpha' = 0, \quad \text{wobei} \quad a''_i = 0 \quad \text{wann immer} \quad \lambda_i \neq 0.$$

7.14 Beispiel [→ 7.12]: Wir setzen Beispiel 7.12 fort und führen die Transformation auf Normalform durch. Die zu untersuchende Quadrik ist die Lösungsmenge der Gleichung

$$Q(x,y) := \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \sqrt{5} & 2\sqrt{5} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \frac{3}{4} = 0.$$

Schritt 1: Das charakteristische Polynom von A ist

$$\begin{vmatrix} 1-\lambda & 2 \\ 2 & -2-\lambda \end{vmatrix} = \lambda^2 + \lambda - 2 - 4 = \lambda^2 + \lambda - 6 = (\lambda - 2)(\lambda + 3),$$

und es ergeben sich die Eigenwerte $\lambda_1 = 2$ und $\lambda_2 = -3$. Die zugehörigen Eigenvektoren (auf Länge Eins normiert) sind dann

$$\text{zu} \quad \lambda_1 = 2\colon \quad \vec{v}_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix},$$
$$\text{zu} \quad \lambda_2 = -3\colon \quad \vec{v}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

Es ist $\vec{v}_1, \vec{v}_2$ eine ONB. Also ist

$$V := \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix},$$

und diese Matrix beschreibt eine Drehung um $26.56\ldots^\circ$ im Gegenuhrzeigersinn. Mit dieser Matrix V gilt

$$V^T A V = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix} =: A'.$$

Schritt 2: Es ist

$$\vec{a}' = V^T \vec{a} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} \sqrt{5} \\ 2\sqrt{5} \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix},$$

und deshalb gilt: ein Koordinatenvektor $\left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ löst die Gleichung $Q(x,y) = 0$ genau dann, wenn der neue Koordinatenvektor $\left[\begin{smallmatrix} x' \\ y' \end{smallmatrix}\right] := V^T \left[\begin{smallmatrix} x \\ y \end{smallmatrix}\right]$ folgende Gleichung erfüllt:

$$Q'(x',y') = \begin{bmatrix} x' & y' \end{bmatrix} \cdot \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} 4 & 3 \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \end{bmatrix} - \frac{3}{4} = 0.$$

Schritt 3: Wir führen die quadratische Ergänzung durch:

$$\begin{aligned}
& Q'(x',y') = 0 \\
\iff\quad & 2x'^2 - 3y'^2 + 4x' + 3y' - \frac{3}{4} = 0 \\
\iff\quad & 2(x'+1)^2 - 2 - 3(y'^2 - y') - \frac{3}{4} = 0 \\
\iff\quad & 2(x'+1)^2 - 3\Big(\Big(y' - \frac{1}{2}\Big)^2 - \frac{1}{4}\Big) - \frac{11}{4} = 0 \\
\iff\quad & 2(x'+1)^2 - 3\Big(y' - \frac{1}{2}\Big)^2 + \frac{3}{4} - \frac{11}{4} = 0 \\
\iff\quad & 2(x'+1)^2 - 3\Big(y' - \frac{1}{2}\Big)^2 - 2 = 0 \\
\iff\quad & (x'+1)^2 - \frac{(y'-1/2)^2}{2/3} - 1 = 0 \\
\iff\quad & x''^2 - \frac{y''^2}{2/3} - 1 = 0,
\end{aligned}$$

und dies ist die Normalform. Wir fassen die Umrechnungen zusammen:

$$\begin{aligned}
x'' &= x' + 1, & y'' &= y' - \frac{1}{2}, \\
x' &= \frac{2}{\sqrt{5}}x + \frac{1}{\sqrt{5}}y, & y' &= -\frac{1}{\sqrt{5}}x + \frac{2}{\sqrt{5}}y.
\end{aligned}$$

7.15 Normalformen für $n = 2$: Im Falle $n = 2$ haben wir folgende Normalformen:

Fall A: beide Eigenwerte von A sind $\neq 0$:

$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$	Ellipse (evtl. ein Kreis)
$\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$	leere Menge
$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$	Hyperbel
$x^2 + a^2y^2 = 0$	ein Punkt
$x^2 - a^2y^2 = 0$	ein Geradenpaar

Fall B: ein Eigenwert von A ist Null:

$x^2 - 2py = 0$	Parabel
$x^2 - a^2 = 0$	ein Paar paralleler Geraden
$x^2 + a^2 = 0$	leere Menge
$x^2 = 0$	eine Gerade.

Unter Umständen müssen zum Erreichen einer dieser Normalformen noch die x- und y-Variable vertauscht werden.
Im Falle $n = 3$ gibt es bereits 17 verschiedene Normalformen, auf deren Auflistung wir hier verzichten wollen.

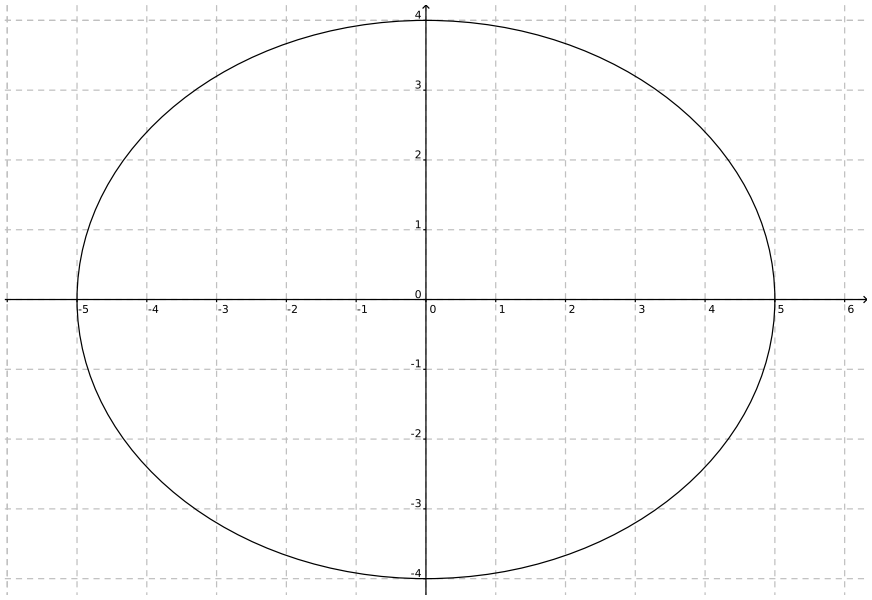


Abbildung 15: Eine Normalform–Ellipse mit Parametern $a = 5$ und $b = 4$.

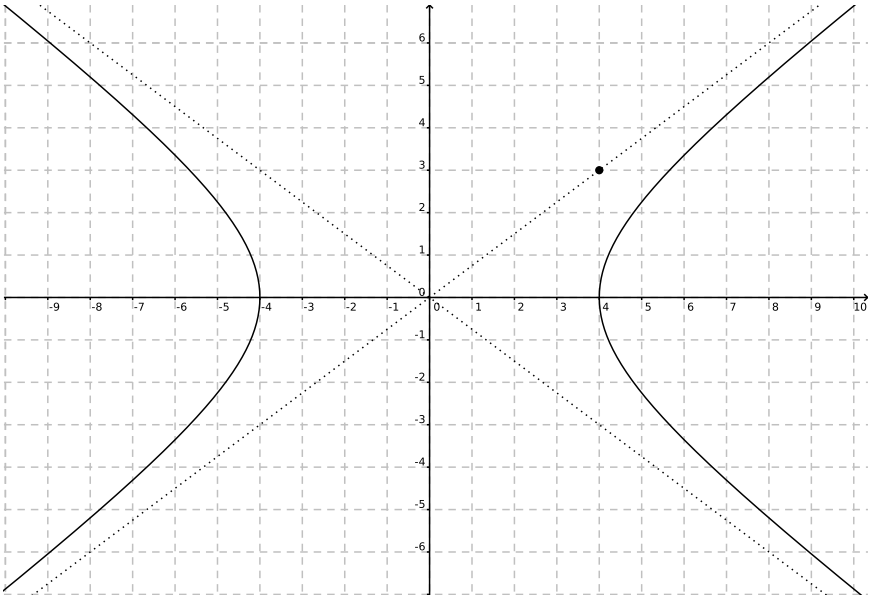


Abbildung 16: Eine Normalform–Hyperbel mit Parametern $a = 4$ und $b = 3$.

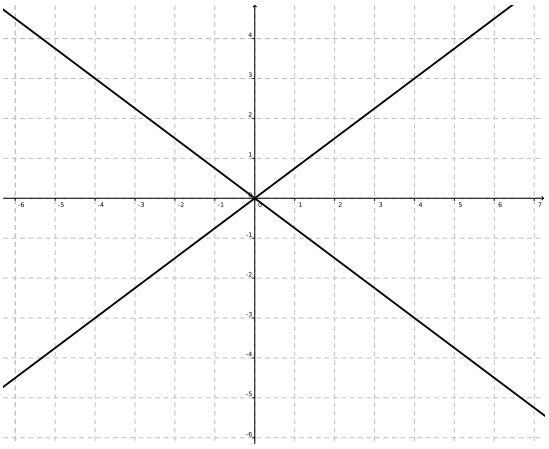


Abbildung 17: Ein Geradenpaar als Lösung der Gleichung $x^2 - a^2y^2 = 1$ mit $a = \frac{4}{3}$.

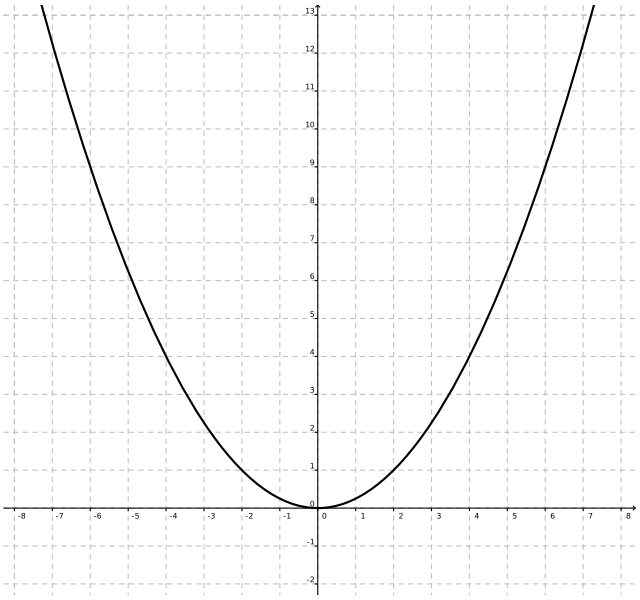


Abbildung 18: Eine Normalform–Parabel als Lösung der Gleichung $x^2 - 2py = 0$ mit $p = 2$.

7.16 Beispiel [→ 7.14]: Wir vollenden Beispiel 7.12 und 7.14. Die erhaltene Normalform ist

$$x''^2 - \frac{y''^2}{\frac{2}{3}} - 1 = 0,$$

also liegt eine Hyperbel vor. Für diese können wir weitere geometrische Informationen angeben: die Halbachsenparameter a und b erfüllen $a^2 = 1$ und $b^2 = \frac{2}{3}$, also ist $a = 1$ und $b = \sqrt{\frac{2}{3}}$.
Die Lage des *Mittelpunkts* der Hyperbel ergibt sich durch die Bedingung $\left[\begin{smallmatrix} x'' \\ y'' \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$. Wir hatten die Umrechnungen $x'' = x' + 1$ und $y'' = y' - \frac{1}{2}$, also gilt für den Mittelpunkt $\left[\begin{smallmatrix} x' \\ y' \end{smallmatrix}\right] = \left[\begin{smallmatrix} -1 \\ 1/2 \end{smallmatrix}\right]$. Im x-y-System bekommen wir wegen $V^{-1} = V^T$ dann für die Koordinaten des Zentrums:

$$\begin{bmatrix} x \\ y \end{bmatrix} = V \begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ \frac{1}{2} \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} -\frac{5}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{\sqrt{5}}{2} \\ 0 \end{bmatrix}.$$

Die Matrix V beschreibt eine Drehung um 26.56°, also ist die x''–Achse gegen die x–Achse um 26.56° gedreht.

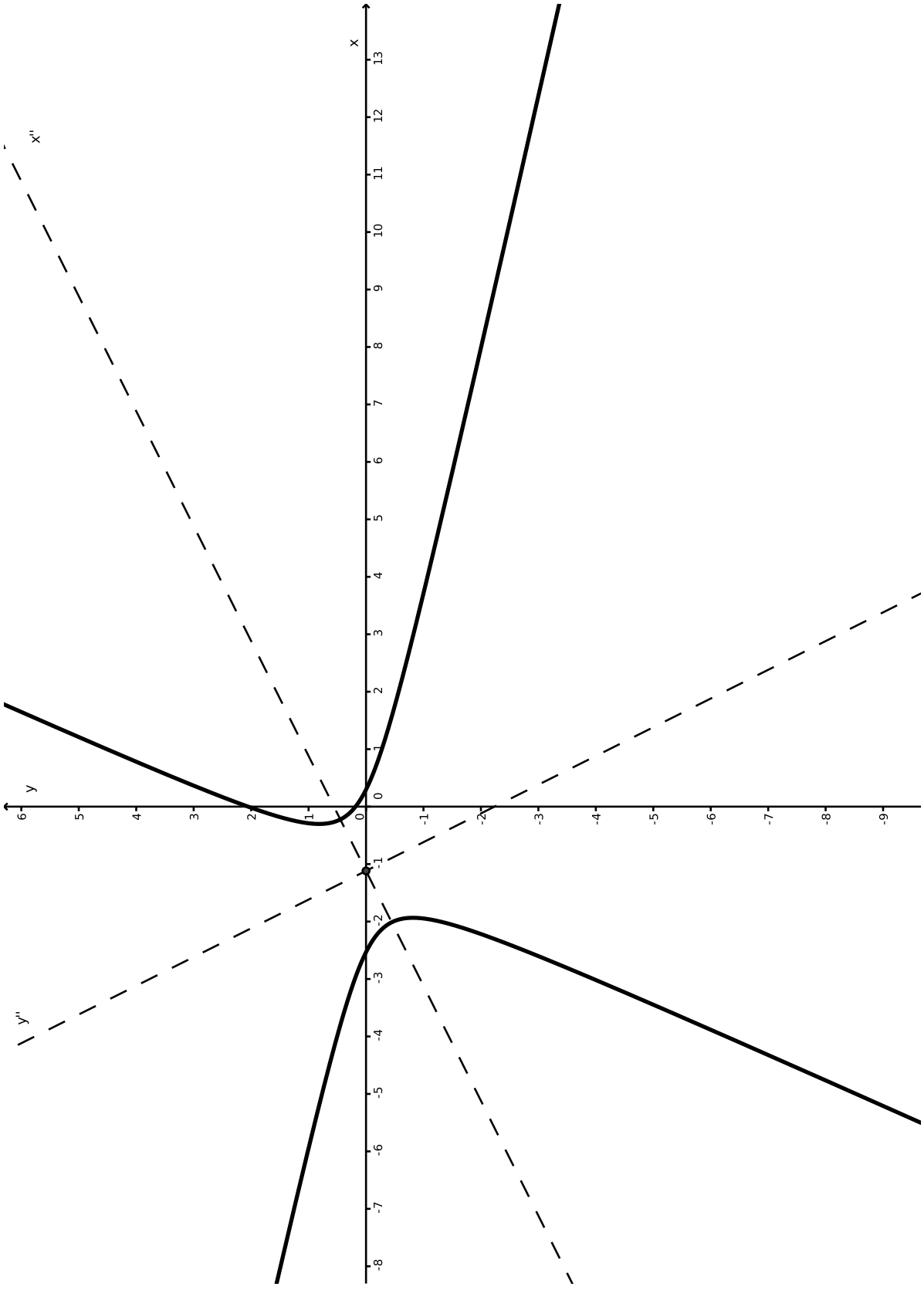


Abbildung 19: Die Quadrik aus Beispiel 7.12 und Beispiel 7.14 (aus Papierformatsgründen ist das Bild um 90° gedreht worden).

8 Folgen

8.1 Erinnerung: Wir wollen kurz an einige Rechenregeln und Begriffe erinnern:

- Für reelle Zahlen x, y gilt die *Dreiecksungleichung* [→ 2.4]

$$|x+y| \leq |x| + |y|, \quad |x-y| \geq ||x| - |y||.$$

  Es folgt, dass für endlich viele reelle Zahlen $x_1, x_2, \ldots, x_n$ gilt:

$$\Big|\sum_{i=1}^{n} x_i\Big| = |x_1 + x_2 + \cdots + x_n| \leq |x_1| + |x_2| + \cdots + |x_n| = \sum_{i=1}^{n} |x_i|.$$

- Die *Fakultät* einer Zahl $n \in \mathbb{N}_0$ ist gegeben durch

$$n! = \prod_{j=1}^{n} j.$$

  Im Fall $n = 0$ gibt es keine Faktoren, sodass das Produkt vereinbarungsgemäß [→ 1.26] den Wert 1 hat, also $0! = 1$. Anderenfalls ist $n! = 1 \cdot 2 \cdots n$, also beispielsweise $1! = 1,\ 3! = 3 \cdot 2 \cdot 1 = 6,\ 5! = 120$.

- Der *Binomialkoeffizient* der Zahlen $n, k \in \mathbb{N}_0$ mit $k \leq n$ ist gegeben durch

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

  Es ist also beispielsweise $\binom{4}{2} = 6$ oder $\binom{n}{0} = 1$. Insbesondere gilt die *binomische Formel*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

  Das *Pascalsche Dreieck* bietet eine effiziente Methode, Binomialkoeffizienten allein durch Additionen zu berechnen. Hier ergibt sich jeder Eintrag im Inneren des Dreiecks als Summe der beiden direkt über ihm stehenden Zahlen,

$$\begin{array}{ccccccccccc}
 & & & & 1 & & & & & & \\
 & & & 1 & & 1 & & & & & \\
 & & 1 & & 2 & & 1 & & & & \\
 & 1 & & 3 & & 3 & & 1 & & & \\
1 & & 4 & & 6 & & 4 & & 1 & & \\
 & & & & \vdots & & & & & &
\end{array}$$

  Es ist also $\binom{n}{k}$ immer eine natürliche Zahl.

8.2 Folge: Wir beschäftigen uns nun mit unendlichen Abfolgen reeller Zahlen, also beispielsweise

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \ldots$$

Man kann diese Zahlen als Funktionswerte einer Funktion a verstehen, die jeder natürlichen Zahl n, dem sogenannten *Index*, eine reelle Zahl zuweist, im Beispiel ist also

$a(1) = 1/2$, $a(2) = 1/4, a(5) = 1/32$, usw. In anderen Fällen kann es zweckmäßig sein, den Bereich der Indizes nicht bei 1, sondern bei einer anderen ganzen Zahl $k \in \mathbb{Z}$ beginnen zu lassen. Das Definitionsgebiet der Funktion ist dann also nicht $\mathbb{N}$, sondern $\mathbb{Z}_{\geq k} = \{n \in \mathbb{Z} : n \geq k\}$. Dies führt auf die folgende Definition:

Eine Funktion $a : \mathbb{Z}_{\geq k} \to \mathbb{R}$, die jeder ganzen Zahl $n \geq k$ eine reelle Zahl zuweist, nennt man *reelle Zahlenfolge* oder kurz *Folge*.

Abweichend von der üblichen Notation bezeichnet man den Funktionswert von a an der Stelle n meist nicht mit $a(n)$ sondern mit a_n. Man nennt a_n das *n-te Folgenglied* oder das *Folgenglied zum Index n*. Ebenfalls abweichend von der Standardnotation für Funktionen gibt es für die Angabe von Folgen verschiedene Schreibweisen: Üblich sind Bezeichnungen wie $(a_n)_{n\in\mathbb{N}}$ oder $(a_n)_{n\geq k}$ oder $a_n, n \in \mathbb{Z}_{\geq k}$. Wenn es nicht relevant ist, wo die Indizierung beginnt, schreibt man auch vereinfachend $(a_n)_n$ oder schlicht a_n. Auch die Folgenglieder können in verschiedener Form angegeben werden:

- **Aufzählung:** Durch Angabe einiger Folgenglieder wird das zugrunde liegende Prinzip verdeutlicht, z.B. im Fall $(a_n)_{n\in\mathbb{N}}$:

$$\begin{array}{rll} 1,2,3,4,5,\ldots, & \text{also} & a_n = n. \\ 1,\frac{1}{2},\frac{1}{3},\frac{1}{4},\frac{1}{5},\ldots, & \text{also} & a_n = \frac{1}{n}. \\ -1,1,-1,1,-1,\ldots, & \text{also} & a_n = (-1)^n. \\ 1,-\frac{1}{2},\frac{1}{4},-\frac{1}{8},\frac{1}{16},\ldots, & \text{also} & a_n = (-1)^{n+1}\frac{1}{2^{n-1}}. \\ 1,0,3,0,5,0,7,0,\ldots, & \text{also} & a_n = \begin{cases} n, & \text{wenn } n \text{ ungerade} \\ 0, & \text{wenn } n \text{ gerade} \end{cases} \end{array}$$

  Dieses Vorgehen ist intuitiv und deshalb gelegentlich zweckmäßig, aber dennoch unpräzise.

- **Explizite Form:** Hier wird eine Formel zur Berechnung von a_n angegeben, z.B.

$$a_n = \sqrt[n]{n!}, \quad n \in \mathbb{N}.$$

- **Rekursion:** Bei dieser Form sind einige Folgenglieder $a_1, a_2, \ldots, a_m$ explizit gegeben. Alle weiteren Folgenglieder werden dann mittels einer Formel aus den vorhergehenden berechnet, z.B.

$$a_1 = 0, \quad a_2 = 1, \quad a_n = a_{n-1} + a_{n-2}, \quad n > 2.$$

  Das Folgenglied zu einem gegebenen Index lässt sich mittels der Rekursionsformel meist einfach bestimmen (insbesonder mit Hilfe eines Computerprogramms), die Herleitung allgemeingültiger Aussagen oder die Angabe einer expliziten Form kann dagegen beliebig schwierig sein.

8.3 Eigenschaften: Eine Folge $a_n, n \geq k$, heißt

- *konstant*, wenn alle Folgenglieder denselben Wert haben.
- *positiv/negativ*, wenn alle Folgenglieder positiv bzw. negativ sind.
- *alternierend*, wenn aufeinanderfolgende Folgenglieder verschiedenes Vorzeichen haben, d.h.,

$$a_{n+1}a_n < 0, \quad n \in \mathbb{N}.$$

- *monoton wachsend/fallend*, wenn

$$a_{n+1} \geq a_n \quad \text{bzw.} \quad a_{n+1} \leq a_n, \quad n \in \mathbb{N}.$$

- *streng monoton wachsend/fallend*, wenn

$$a_{n+1} > a_n \quad \text{bzw.} \quad a_{n+1} < a_n, \quad n \in \mathbb{N}.$$

- *beschränkt*, wenn es eine Zahl $c \in \mathbb{R}$ gibt, sodass

$$|a_n| \leq c, \quad n \in \mathbb{N}.$$

Alle Eigenschaften einer Folge können mit dem Zusatz *„fast überall (f.ü.)“* versehen werden. Dies bedeutet, dass diese Eigenschaft erst ab einem gewissen Index n_0 durchgängig erfüllt ist und für die endlich vielen Indizes unterhalb von n_0 verletzt sein darf. Offensichtlich spielt für Eigenschaften, die f.ü. gelten, der Startindex k keine Rolle.

8.4 Beispiel:

- Die Folge $(a_n)_n$ mit $a_n = n^2 - 15$ ist streng monoton wachsend und f.ü. positiv.
- Die Folge $(a_n)_n$ mit $a_n = (-1)^n/n^2$ ist alternierend und beschränkt, da $|a_n| \leq 1$.
- Eine konstante Folge ist monoton wachsend, monoton fallend und beschränkt.

8.5 Nullfolge: Eine Folge $(a_n)_{n \geq k}$ heißt *Nullfolge*, wenn für jedes beliebig vorgegebene $\varepsilon > 0$ gilt:

$$|a_n| < \varepsilon \quad \text{fast überall.}$$

Es darf also jeweils nur endlich viele Indizes geben, für die $|a_n| \geq \varepsilon$ gilt. Die Eigenschaft, eine Nullfolge zu sein, ist vom Startindex k unabhängig.

8.6 Beispiel:

- Gegeben sei die Folge $(a_n)_n$ mit $a_n = 1/n$. Die Bedingung $|a_n| = \frac{1}{n} < \varepsilon$ ist äquivalent zu $n > 1/\varepsilon$. Unabhängig davon, wie klein man ε nun wählt, ist $1/\varepsilon$ eine feste Zahl. Es gibt folglich immer nur endlich viele Werte von n, die kleiner sind. Also ist $(a_n)_n$ eine Nullfolge.
- Gegeben sei die Folge $(b_n)_n$ mit $b_n = \cos(n\pi) + 1$. Für gerades n ist hier $b_n = 2$, während sich für ungerades n der Wert $b_n = 0$ ergibt. Wählen wir nun $\varepsilon = 1/2$, dann gibt es zwar unendlich viele Folgenglieder mit $|b_n| = 0 < 1/2$, es gibt aber auch unendlich viele Folgenglieder mit $|b_n| = 2 > 1/2$. Also ist $(b_n)_n$ keine Nullfolge.

8.7 Regeln: Sei $(a_n)_n$ eine Nullfolge, dann gilt:

- $(a_n)_n$ ist beschränkt.
- Sei $(b_n)_n$ eine Nullfolge, dann ist auch $(a_n \pm b_n)_n$ eine Nullfolge.
- Sei $c \in \mathbb{R}$ eine Konstante, dann ist $(a_n \cdot c)_n$ eine Nullfolge.
- Sei $(c_n)_n$ eine beschränkte Folge, dann ist $(a_n \cdot c_n)_n$ eine Nullfolge.
- Sei $(d_n)_n$ eine Folge mit $|d_n| \leq |a_n|$ f.ü., dann ist $(d_n)_n$ eine Nullfolge.

8.8 Beispiel: Gegeben sei die Nullfolge $(a_n)_n$ mit $a_n = 1/n$.

- $(6/n)_n$ ist eine Nullfolge.
- Die Folge $(c_n)_n$ mit $c_n = \sin n - \pi$ ist beschränkt durch $|c_n| \leq 5$. Also ist

$$\left(\frac{\sin n - \pi}{n}\right)_n$$

  eine Nullfolge.

- Sei

$$d_n = \frac{(-1)^n 5n}{n^2 - 3},$$

  dann ist $|d_n| < |6/n|$ für $n \geq 5$. Also ist $(d_n)_n$ eine Nulfolge.

- Seien p und q zwei Polynome, dann ist die Folge $(a_n)_n$ mit $a_n = p(n)/q(n)$ genau dann eine Nullfolge, wenn $\text{grad}(p) < \text{grad}(q)$.

8.9 Grenzwert: Eine Folge $(a_n)_n$ heißt *konvergent mit Grenzwert (oder Limes)* a, wenn $(a_n - a)_n$ eine Nullfolge ist. Man schreibt dann

$$\lim_{n\to\infty} a_n = a$$

und liest: „Der Grenzwert von a_n für n gegen Unendlich ist a“. Man beachte, dass das ∞-Symbol hier nur eine Schreibweise ist. In der Definition des Grenzwerts kommen nur *endliche* Werte von n vor. Eine Folge, die nicht konvergent ist, heißt *divergent*.

8.10 Regeln: Seien $(a_n)_n$ und $(b_n)_n$ zwei konvergente Folgen mit Grenzwert a bzw. b, dann gilt:

- Die Folge der Beträge $(|a_n|)_n$ konvergiert gegen $|a|$, also $\lim_{n\to\infty} |a_n| = |a|$.
- Die Folge $(a_n)_n$ ist beschränkt, denn $|a_n - a| \leq c$ [→ 8.7] und

$$|a_n| = |a_n - a + a| \leq |a_n - a| + |a| \leq c + a.$$

- Die Folge $(\alpha a_n + \beta b_n)_n$ ist konvergent mit

$$\lim_{n\to\infty} (\alpha a_n + \beta b_n) = \alpha a + \beta b, \quad \alpha, \beta \in \mathbb{R}.$$

- Die Folge $(a_n \cdot b_n)_n$ ist konvergent mit

$$\lim_{n\to\infty} (a_n \cdot b_n) = a \cdot b.$$

- Die Folge $(a_n/b_n)_n$ ist konvergent mit

$$\lim_{n\to\infty} (a_n/b_n) = a/b,$$

  sofern $b \neq 0$. Die Folge ist divergent, falls $a \neq 0$ und $b = 0$.

8.11 Wichtige Grenzwerte:

$$\lim_{n\to\infty} \frac{p_0 + p_1 n + \cdots + p_k n^k}{q_0 + q_1 n + \cdots + q_k n^k} = \frac{p_k}{q_k}, \quad q_k \neq 0$$

$$\lim_{n\to\infty} \frac{\ln n}{n^\alpha} = 0, \quad \alpha > 0$$

$$\lim_{n\to\infty} n^\alpha q^n = 0, \quad \alpha \in \mathbb{R},\ |q| < 1$$

$$\lim_{n\to\infty} \sqrt[n]{n} = 1$$

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e \approx 2.718$$

$$\lim_{n\to\infty} \frac{n}{\sqrt[n]{n!}} = e$$

8.12 Konvergenzkriterien:

- *Cauchy:* Eine Folge $(a_n)_n$ reeller Zahlen ist genau dann konvergent, wenn es für jedes beliebig vorgegebene $\varepsilon > 0$ eine Zahl n_0 gibt, sodass

$$|a_n - a_m| < \varepsilon \quad \text{für alle} \quad n, m \geq n_0.$$

- *Intervallschachtelung:* Sei $(a_n)_n$ monoton wachsend, $(b_n)_n$ monoton fallend und $(b_n - a_n)_n$ eine Nullfolge. Dann konvergieren beide Folgen gegen denselben Grenzwert und es gilt

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n \in [a_k,\, b_k], \quad k \in \mathbb{N}.$$

- *Einschachtelung:* Seien $(a_n)_n, (b_n)_n$ konvergente Folgen mit demselben Grenzwert a und $(c_n)_n$ eine weitere Folge. Wenn $a_n \leq c_n \leq b_n$ f.ü. gilt, dann ist die Folge $(c_n)_n$ ebenfalls konvergent gegen a.

- Sei $(a_n)_n$ beschränkt und monoton f.ü., dann ist $(a_n)_n$ konvergent.

8.13 Bestimmte Divergenz: Eine Folge $(a_n)_n$ heißt *bestimmt divergent*, wenn sie f.ü. positiv oder f.ü. negativ ist und $(1/a_n)_n$ eine Nullfolge ist. Wir schreiben dann

$$\lim_{n\to\infty} a_n = \infty \quad \text{bzw.} \quad \lim_{n\to\infty} a_n = -\infty.$$

8.14 Beispiel:

- Es gilt

$$\lim_{n\to\infty} e^n/n = \infty, \quad \lim_{n\to\infty} \ln(1/n) = -\infty.$$

- Die Folge $(1 + n + (-1)^n n)_n$ ist positiv und nicht beschränkt. Dennoch ist sie nicht bestimmt divergent, da alle ungeraden Folgenglieder den Wert 1 haben.

8.15 Vektorfolge: Eine *Folge reeller Vektoren* (kurz *Vektorfolge*) ist eine unendliche Abfolge

$$\vec{a}_1, \vec{a}_2, \vec{a}_3, \ldots, \vec{a}_n, \ldots$$

von Vektoren im $\mathbb{R}^d$. Mathematisch präziser ist eine Vektorfolge eine Abbildung $\mathbb{N} \longrightarrow \mathbb{R}$, wobei jeder natürlichen Zahl (Index) ein Vektor zugeordnet wird.
Man schreibt dafür $(\vec{a}_n)_{n\in\mathbb{N}}$ oder auch kurz $(\vec{a}_n)_n$. Die Komponenten $a_{i,n}$ von $\vec{a}_n$ bilden die reellen Zahlenfolgen $(a_i)_n$,

$$\vec{a}_n = \begin{bmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{d,n} \end{bmatrix}, \quad (\vec{a}_n)_n = \begin{bmatrix} (a_1)_n \\ (a_2)_n \\ \vdots \\ (a_d)_n \end{bmatrix}.$$

Die Vektorfolge $(\vec{a}_n)_n$ heißt *konvergent mit Grenzwert* $\vec{a}$, wenn alle Komponentenfolgen konvergieren,

$$\lim_{n\to\infty} \vec{a}_n = \vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{bmatrix}, \quad a_i = \lim_{n\to\infty} a_{i,n}.$$

9 Reihen

9.1 Reihe: Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge. Dann definiert man die zugehörige Folge $(s_m)_{m\in\mathbb{N}}$ gemäß

$$\begin{aligned} s_1 &:= a_1 \\ s_2 &:= a_1 + a_2 \\ s_3 &:= a_1 + a_2 + a_3 \\ &\vdots \\ s_m &:= a_1 + a_2 + \cdots + a_m = \sum_{n=1}^{m} a_n \\ &\vdots \end{aligned}$$

Man bezeichnet die Folge $(s_m)_m$ als die *Reihe* und die Folgenglieder s_m als die *Partialsummen* zur Folge $(a_n)_n$. Wenn die Reihe, also die Folge der Partialsummen, konvergiert, dann schreibt man für den Grenzwert

$$\sum_{n=1}^{\infty} a_n := \lim_{m\to\infty} s_m.$$

Vollkommen analog definiert man zur Folge $(a_n)_{n\geq k}$ die die Reihe $(s_m)_{m\geq k}$ durch die Partialsummen

$$s_m := a_k + a_{k+1} + \cdots + a_m = \sum_{n=k}^{m} a_n, \quad m \geq k,$$

und schreibt im Falle der Konvergenz

$$\sum_{n=k}^{\infty} a_n := \lim_{m\to\infty} s_m.$$

Gemäß den Regeln für Grenzwerte von Folgen gilt

$$\sum_{n=1}^{\infty}(\alpha a_n + \beta b_n) = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n, \quad \alpha, \beta \in \mathbb{R},$$

sofern die Reihen zu den Folgen $(a_n)_n$ und $(b_n)_n$ konvergieren.

9.2 Beispiel: Für ein gegebenes $q \neq 1$ sei die Folge $(a_n)_{n\geq 0}$ gegeben durch $a_n := q^n$. Die zugehörige Folge der Partialsummen

$$s_0 = 1,\ s_1 = 1 + q,\ s_2 = 1 + q + q^2, \ldots$$

wird als *geometrische Reihe* bezeichnet. In diesem Fall lassen sich die Partialsummen explizit angeben,

$$s_m = \frac{1 - q^{m+1}}{1 - q}.$$

Die geometrische Reihe konvergiert also genau dann mit Grenzwert

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1 - q},$$

wenn $|q| < 1$. Anderenfalls ist die geometrische Reihe divergent.

9.3 Notwendige Bedingung: Sei $\sum_{n=1}^{\infty} a_n$ eine konvergente Reihe mit Grenzwert s. Dann gilt

$$a_n = s_n - s_{n-1} = (s_n - s) - (s_{n-1} - s).$$

Beide Klammern auf der rechten Seite sind Nullfolgen, also muss auch $(a_n)_n$ eine Nullfolge sein [→ 8.7]. Für die Konvergenz einer Reihe ist es also notwendig, dass $(a_n)_n$ eine Nullfolge ist. Diese Bedingung ist aber keineswegs hinreichend, d.h., selbst wenn $(a_n)_n$ eine Nullfolge ist, kann die zugehörige Reihe divergent sein.

9.4 Beispiel:

- Sei $a_n := 1/n$. Die zugehörige Folge der Partialsummen

  $$s_1 = 1,\ s_2 = 1 + 1/2,\ s_3 = 1 + 1/2 + 1/3, \ldots$$

  wird als *harmonische Reihe* bezeichnet. Diese ist bestimmt divergent,

  $$\sum_{n=1}^{\infty} \frac{1}{n} = \infty.$$

- Die Reihe zur Folge $a_n = 1/n^{\alpha}$ konvergiert genau dann, wenn $\alpha > 1$. Es ist z.B.

  $$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

9.5 Leibniz-Kriterium: Sei $(a_n)_n$ eine alternierende Nullfolge [→ 8.3] mit der Eigenschaft, dass die Folge der Beträge $(|a_n|)_n$ monoton fällt, dann konvergiert die zugehörige Reihe.

9.6 Beispiel: Sei $a_n := (-1)^{n+1}/n$. Die zugehörige Folge der Partialsummen

$$s_1 = 1,\ s_2 = 1 - 1/2,\ s_3 = 1 - 1/2 + 1/3, \ldots$$

wird als *alternierende harmonische Reihe* bezeichnet. Sie konvergiert nach dem Leibniz-Kriterium und hat den Grenzwert

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln 2.$$

9.7 Sprachgebrauch: In der Praxis wird das Symbol $\sum_{n=1}^{\infty} a_n$, das eigentlich nur für den Grenzwert einer Reihe steht, häufig auch für die Reihe selbst verwendet. Man sagt also:

Die Reihe $\sum_{n=1}^{\infty} a_n$ *...*

und meint damit eigentlich

Die Reihe zur Folge $(a_n)_n$ *...*

Außerdem wird der Grenzwert einer Reihe gelegentlich auch nur als *Wert* der Reihe bezeichnet. Der Satz

Die Reihe $\sum_{n=0}^{\infty}(-1)^n/(2n+1)$ *hat den Wert* $\pi/4$.

ist also zu interpretieren als

Die zur Folge $(a_n)_{n\geq 0}, a_n = (-1)^n/(2n+1)$ *gehörende Reihe hat den Grenzwert* $\pi/4$.

9.8 Vergleichskriterien: Gegeben seien die Folgen $(a_n)_n$, $(b_n)_n$.

- *Konvergente Majorante:* Wenn $\sum_{n=1}^{\infty} a_n$ konvergiert und

$$|b_n| \leq a_n \quad \text{f.ü.},$$

  dann konvergiert auch $\sum_{n=1}^{\infty} b_n$.

- *Divergente Minorante:* Wenn $\sum_{n=1}^{\infty} a_n$ divergiert und

$$0 \leq a_n \leq b_n \quad \text{f.ü.},$$

  dann divergiert auch $\sum_{n=1}^{\infty} b_n$.

9.9 Absolute Konvergenz: Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt *absolut konvergent*, wenn die Reihe $\sum_{n=1}^{\infty} |a_n|$ konvergiert. Eine absolut konvergente Reihe ist nach dem Majorantenkriterium konvergent. Eine Reihe, die konvergent, aber nicht absoulut konvergent ist, heißt *bedingt konvergent*. Bei absoult konvergenten Reihen können die Folgenglieder a_n beliebig umgeordnet werden, ohne dass sich der Wert der Reihe ändert. Hier gilt also in einem gewissen Sinne das Kommutativgesetz. Bei bedingt konvergenten Reihen bewirkt eine Umordnung der Folgenglieder a_n hingegen in der Regel eine Änderung des Reihenwertes oder sogar einen Verlust der Konvergenz.

9.10 Beispiel:

- [→ 9.4] Die Reihe $\sum_{n=1}^{\infty} (-1)^n/n^2$ ist absolut konvergent.
- [→ 9.6] Die Reihe $\sum_{n=1}^{\infty} (-1)^n/n$ ist bedingt konvergent.

9.11 Quotientenkriterium: Sei

$$q_n := \left| \frac{a_{n+1}}{a_n} \right| \quad \text{und, falls konvergent,} \quad q := \lim_{n\to\infty} q_n.$$

Die Reihe $\sum_{n=1}^{\infty} a_n$ ist absolut konvergent, wenn $q < 1$ und divergent, wenn $q > 1$. Im Fall $q = 1$ ist keine Entscheidung möglich.
Wenn die Folge q_n nicht konvergiert, dann gilt immer noch:

- Wenn es eine Zahl $q < 1$ gibt, sodass $q_n \leq q$ fast überall (d.h. für alle bis auf endlich viele Werte von n), dann ist die Reihe absolut konvergent.
- Wenn $q_n \geq 1$ fast überall, dann ist die Reihe divergent.

9.12 Wurzelkriterium: Sei

$$w_n := \sqrt[n]{|a_n|} \quad \text{und, falls konvergent,} \quad w := \lim_{n\to\infty} w_n.$$

Die Reihe $\sum_{n=1}^{\infty} a_n$ ist absolut konvergent, wenn $w < 1$ und divergent, wenn $w > 1$. Im Fall $w = 1$ ist keine Entscheidung möglich.
Wenn die Folge w_n nicht konvergiert, dann gilt immer noch:

- Wenn es eine Zahl $w < 1$ gibt, sodass $w_n \leq w$ fast überall, dann ist die Reihe absolut konvergent.

- Wenn $w_n \geq 1$ für unendlich viele n, dann ist die Reihe divergent.

Das Wurzelkriterium ist im folgenden Sinne stärker als das Quotientenkriterium: Wenn $\lim_{n\to\infty} q_n$ existiert, dann existiert auch $\lim_{n\to\infty} w_n$ und es gilt

$$q = \lim_{n\to\infty} q_n = \lim_{n\to\infty} w_n = w.$$

9.13 Beispiel: Sei $a_n = (-1)^n\, n^3\, 2^{-n}$. Das Quotientenkriterium liefert

$$q_n = \left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)^3\, 2^{-n-1}}{n^3\, 2^{-n}} = \frac{(1+1/n)^3}{2}, \quad q = \lim_{n\to\infty} q_n = \frac{1}{2}.$$

Genauso liefert das Wurzelkriterium

$$w_n = \sqrt[n]{|a_n|} = \frac{\sqrt[n]{n^3}}{2}, \quad w = \lim_{n\to\infty} w_n = \frac{1}{2}.$$

Die Reihe

$$\sum_{n=0}^{\infty} \frac{(-1)^n\, n^3}{2^n}$$

ist also absolut konvergent.

9.14 Beispiel: Sei

$$a_n = \begin{cases} 2^{-n} & \text{für } n \text{ gerade} \\ 2^{-n+2} & \text{für } n \text{ ungerade} \end{cases}$$

Es ist also

$$\sum_{n=0}^{\infty} a_n = 1 + 2 + \frac{1}{4} + \frac{1}{2} + \frac{1}{16} + \frac{1}{8} + \cdots .$$

Das Quotientenkriterium liefert die divergente Folge

$$q_n = \begin{cases} 2 & \text{für } n \text{ gerade} \\ 1/8 & \text{für } n \text{ ungerade,} \end{cases}$$

während das Wurzelkriterium mit

$$w_n = \begin{cases} 1/2 & \text{für } n \text{ gerade} \\ \sqrt[n]{4}/2 & \text{für } n \text{ ungerade} \end{cases} \quad \Rightarrow \quad w = \lim_{n\to\infty} w_n = 1/2$$

zeigt, dass die Reihe absolut konvergent ist.

9.15 Beispiel: Für $a_n = n^n/n!$ ist

$$q_n = \frac{(n+1)^{n+1}\, n!}{n^n\, (n+1)!} = \left(1 + \frac{1}{n}\right)^n, \quad w_n = n/\sqrt[n]{n!}.$$

Zum einen folgt aus $\lim_{n\to\infty} q_n = e$ die Divergenz der zugehörigen Reihe. Zum anderen liefert die Gleichheit der Grenzwerte $w = q$ die *Stirlingsche Formel* [→ 8.11]

$$\lim_{n\to\infty} \frac{n}{\sqrt[n]{n!}} = e.$$

9.16 Beispiel: Für $a_n = x^n/n!$ ist

$$q_n = \left| \frac{x^{n+1}\, n!}{(n+1)!\, x^n} \right| = \frac{|x|}{n+1}, \quad q = \lim_{n\to\infty} q_n = 0.$$

Die zugehörige Reihe ist also für alle $x \in \mathbb{R}$ konvergent und definiert die e-Funktion,

$$e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{R}.$$

10 Funktionengrenzwert und Stetigkeit

10.1 Reelle Funktion: Erinnerung [→ 1.6], [→ 1.16]: Eine *reelle Funktion* $f : D_f \to \mathbb{R}$ ordnet jedem Element $x \in D_f$ der Menge $D_f \subset \mathbb{R}$ eine reelle Zahl $y \in \mathbb{R}$ zu, und man schreibt

$$y = f(x), \quad x \in D.$$

- Die Menge D_f heißt *Definitionsmenge* von f.
- x heißt *Urbild von* y und y heißt *Bild von* x.
- Die *Bildmenge* von f ist die Menge aller Bilder,

$$B_f := \{y = f(x) : x \in D_f\}.$$

10.2 Häufungspunkt: $x_* \in \mathbb{R}$ heißt *Häufungspunkt* der Menge $D_f \subset \mathbb{R}$, wenn es eine Folge $(x_n)_n$ gibt mit

$$\lim_{n\to\infty} x_n = x_*, \quad x_n \in D_f \ , \quad x_n \neq x_*, \quad \text{f.ü..}$$

10.3 Beispiel:

- $x_* = 1$ ist ein Häufungspunkt der Menge $(0, 1)$.
- $x_* = 1$ ist kein Häufungspunkt der Menge $\mathbb{N}$.
- $x_* = \pi$ ist ein Häufungspunkt der Menge $\mathbb{Q}$ der rationalen Zahlen.

10.4 Grenzwert: Sei $f : D_f \to \mathbb{R}$ eine reelle Funktion. Man sagt, dass f an der Stelle x_* den *Grenzwert* f_* hat, wenn x_* ein Häufungspunkt von D_f ist und wenn für jede gegen x_* konvergente Folge $(x_n)_n$ gemäß 10.2 gilt

$$\lim_{n\to\infty} f(x_n) = f_*.$$

Man schreibt dann

$$\lim_{x\to x_*} f(x) = f_*.$$

Wenn die Folge $(f(x_n))_n$ stets bestimmt divergent ist, dann schreibt man

$$\lim_{x\to x_*} f(x) = \infty \quad \text{bzw.} \quad \lim_{x\to x_*} f(x) = -\infty.$$

Wenn $\lim_{n\to\infty} f(x_n) = f_*$ für jede bestimmt divergente Folge $(x_n)_n$ mit $\lim_{n\to\infty} x_n = \infty$, dann schreibt man

$$\lim_{x\to\infty} f(x) = f_*.$$

Analog sind die Ausdrücke

$$\lim_{x\to-\infty} f(x) = f_*, \quad \lim_{x\to\infty} f(x) = \pm\infty, \quad \lim_{x\to-\infty} f(x) = \pm\infty$$

erklärt.

10.5 Beispiel:

- $$\lim_{x\to 0} x\sin(1/x) = 0, \quad \lim_{x\to\infty} \arctan(x) = \pi/2, \quad \lim_{x\to 0} \ln|x| = -\infty.$$
- Die Funktion $f(x) = \sin(1/x), x \in \mathbb{R}_{\neq 0}$, hat an der Stelle $x_* = 0$ keinen Grenzwert.
- Die Funktion $f(x) = x, x \in \mathbb{N}$, hat an der Stelle $x_* = 1$ keinen Grenzwert, da $x_* = 1$ kein Häufungspunkt von $\mathbb{N}$ ist [→ 10.3].
- Die *Heaviside-Funktion* $H : \mathbb{R} \to \mathbb{R}$ ist definiert durch

  $$H(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 & \text{für } x \geq 0. \end{cases}$$

  Sie hat an der Stelle $x_* = 0$ keinen Grenzwert. Sei $f(x) := H(x) + H(-x)$, dann gilt

  $$\lim_{x\to 0} f(x) = 1,$$

  aber $f(0) = 2$.

10.6 Einseitiger Grenzwert: Man sagt, dass f an der Stelle x_* den *rechtsseitigen bzw. linksseitigen Grenzwert* f_* hat, wenn x_* ein Häufungspunkt von D_f ist und wenn für jede Folge $(x_n)_n$ gemäß 10.2 mit der zusätzlichen Eigenschaft

$$x_n > x_* \quad \text{bzw.} \quad x_n < x_*, \quad n \in \mathbb{N},$$

gilt

$$\lim_{n\to\infty} f(x_n) = f_*.$$

Man schreibt dann

$$\lim_{x\downarrow x_*} f(x) = f_* \quad \text{bzw.} \quad \lim_{x\uparrow x_*} f(x) = f_*.$$

10.7 Beispiel: Für die Heaviside-Funktion [→ 10.5] gilt

$$\lim_{x\uparrow 0} H(x) = 0, \quad \lim_{x\downarrow 0} H(x) = 1.$$

10.8 Stetigkeit: Eine Funktion $f : D_f \to \mathbb{R}$ heißt *stetig an der Stelle* $x_* \in D_f$, wenn

$$\lim_{x\to x_*} f(x) = f(x_*).$$

f heißt *stetig*, wenn f an allen Stellen $x \in D_f$ stetig ist.

10.9 Regeln:

- Alle elementaren Funktionen (Polynome, exp, sin, cos, tan und deren Umkehrfunktionen) sowie Betrag-, Potenz- und Wurzelfunktionen sind stetig auf ihrem gesamten Definitionsgebiet.
- Summe, Differenz, Produkt, Quotient und Verkettung stetiger Funktionen sind stetig auf ihrem gesamten Definitionsgebiet.

- Die Umkehrfunktion einer stetigen Funktion ist stetig, sofern das Definitionsgebiet ein Intervall ist.

- *Zwischenwertsatz:* Sei $f := [a, b] \to \mathbb{R}$ stetig, dann ist B_f ein Intervall. Das heißt insbesondere, dass die Funktion alle Werte zwischen $f(a)$ und $f(b)$ annimmt.

- *Satz von Weierstraß:* Sei $f : [a, b] \to \mathbb{R}$ stetig, dann gibt es Stellen $\underline{x}, \overline{x} \in [a, b]$ mit

$$f(\underline{x}) \leq f(x) \leq f(\overline{x}), \quad x \in [a, b].$$

  Eine derartige Aussage gilt in der Regel nicht, wenn das Definitionsgebiet kein abgeschlossenes Intervall ist.

10.10 Bemerkungen:

- f ist stetig an der Stelle x_* genau dann, wenn es für jedes $\varepsilon > 0$ es ein $\delta > 0$ gibt, sodass

$$|f(x) - f(x_*)| < \varepsilon \quad \text{falls} \quad |x - x_*| < \delta.$$

- Bei einer stetigen Funktion bewirken kleine Änderungen des Arguments kleine Änderungen des Funktionswerts.

- Das Schaubild einer stetigen Funktion besitzt keine Sprünge, sofern das Definitionsgebiet ein Intervall ist.

Nur die erste Aussage ist streng mathematischer Natur. Die beiden anderen sind unpräzise, aber gelegentlich hilfreich.

11 Differenziation

11.1 Differenzierbarkeit: Im Folgenden sei $f : I \to \mathbb{R}$ stets eine reelle Funktion, deren Definitionsgebiet $D_f = I$ ein offenes Intervall ist. Für zwei Stellen $x \neq x_0$ in I definieren wir den *Differenzenquotienten*

$$\frac{f(x) - f(x_0)}{x - x_0}.$$

Er gibt die Steigung der Sekante an, die durch die Punkte $(x, f(x))$ und $(x_0, f(x_0))$ gegeben ist. Die Funktion f heißt *differenzierbar an der Stelle* $x_0 \in I$, wenn der Grenzwert

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert.

- Der Grenzwert heißt *Ableitung* von f an der Stelle x_0 und man schreibt dafür

  $$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

  Andere gebräuchliche Bezeichnungen sind

  $$f'(x_0) = Df(x_0) = \frac{df}{dx}(x_0).$$

- Setzt man $h := x - x_0$, dann erhält man die äquivalente Darstellung

  $$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

- Die Funktion f heißt *differenzierbar*, wenn sie an allen Stellen $x_0 \in I$ differenzierbar ist. Die Funktion $f' : I \to \mathbb{R}$, die jedem Punkt $x_0 \in I$ den Wert der Ableitung zuweist, heißt *Ableitungsfunktion*.

- Ist f in x_0 differenzierbar, so ist f in x_0 stetig. Die Umkehrung gilt im Allgemeinen nicht, wie das Beispiel der Betragsfunktion zeigt.

- Die Ableitung kann verwendet werden, um *lokale Extremstellen*, also Maxima oder Minima, und Wendepunkte von f, zu bestimmen.

11.2 Beispiel:

- Für die lineare Funktion $f(x) = ax + b$ ist

  $$\frac{f(x) - f(x_0)}{x - x_0} = \frac{(ax + b) - (ax_0 + b)}{x - x_0} = a.$$

  Also ist f differenzierbar und es gilt $f'(x_0) = a$ für alle $x_0 \in \mathbb{R}$.

- Für die quadratische Funktion $f(x) = x^2$ ist

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

Der Grenzwert für x gegen x_0 ist

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} x + x_0 = 2x_0.$$

Also ist f differenzierbar und es gilt $f'(x_0) = 2x_0$ für alle $x_0 \in \mathbb{R}$.

- Für die Betragsfunktion $f(x) = |x|$ und $x_0 = 0$ ist

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{|x|}{x} = \begin{cases} +1 & \text{für } x > 0 \\ -1 & \text{für } x < 0. \end{cases}$$

Es existiert also kein Grenzwert für $x \to 0$ und die Betragsfunktion ist deshalb nicht differenzierbar an der Stelle $x_0 = 0$ (aber stetig).

11.3 Wichtige Ableitungen:

f	f'	
x^n	nx^{n-1}	$x \in \mathbb{R}$, $n \in \mathbb{N}$
x^n	nx^{n-1}	$x \neq 0$, $-n \in \mathbb{N}$
x^α	$\alpha x^{\alpha-1}$	$x > 0$, $\alpha \in \mathbb{R}$
e^x	e^x	$x \in \mathbb{R}$
$\ln x$	$1/x$	$x > 0$
$\sin x$	$\cos x$	$x \in \mathbb{R}$
$\cos x$	$-\sin x$	$x \in \mathbb{R}$
$\arctan x$	$1/(1+x^2)$	$x \in \mathbb{R}$
$\sinh x$	$\cosh x$	$x \in \mathbb{R}$
$\cosh x$	$\sinh x$	$x \in \mathbb{R}$
$\operatorname{artanh} x$	$1/(1-x^2)$	$x \in (-1, 1)$

11.4 Regeln: Seien f und g zwei differenzierbare Funktionen, dann gilt

- *Linearität:* Für $a, b \in \mathbb{R}$ ist die Funktion $af + bg$ differenzierbar,

$$(af + bg)' = af' + bg'.$$

- *Produktregel:* Die Funktion $f \cdot g$ ist differenzierbar,

$$(f \cdot g)' = fg' + f'g.$$

- *Quotientenregel:* Die Funktion f/g ist differenzierbar, sofern $g \neq 0$,

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

- *Kettenregel:* Die Funktion $f \circ g$ ist differenzierbar,

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

- *Umkehrfunktion:* Wenn $f' \neq 0$, dann ist f streng monoton [→ 11.8] und damit injektiv. Es existiert also die Umkehrfunktion [→ 1.11] $g := f^{-1}$, und es gilt

$$g'(y) = \frac{1}{f'(x)}, \quad \text{wobei} \quad f(x) = y, \quad g(y) = x.$$

11.5 Beispiel:

- Linearität:

$$(3\sin x - 2x^3)' = 3\cos x - 6x^2$$

- Produktregel:

$$(x^2 \ln x)' = 2x \ln x + x, \quad x > 0$$

- Quotientenregel:

$$(\tan x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = 1 + \tan^2 x, \quad x \in (-\pi/2, \pi/2)$$

- Kettenregel:

$$(\exp(x^2))' = \exp(x^2) \cdot (2x)$$
$$(\arctan(1/x))' = \frac{-1/x^2}{1 + (1/x)^2} = \frac{-1}{1 + x^2}, \quad x \neq 0$$

- Umkehrfunktion:

$$y = f(x) = \tan x, \quad f'(x) = 1 + \tan^2 x > 0$$
$$x = g(y) = \arctan y, \quad g'(y) = \frac{1}{1 + \tan^2 x} = \frac{1}{1 + y^2}$$

11.6 Tangente: Sei f an der Stelle x_0 differenzierbar. Die Funktion

$$t(x) := f(x_0) + f'(x_0)(x - x_0)$$

beschreibt eine Gerade, die man als *Tangente* von f an der Stelle x_0 bezeichnet. Sei $r(x) := f(x) - t(x)$ die Abweichung zwischen Funktion und Tangente, dann gilt

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x), \quad \lim_{x \to x_0} \frac{r(x)}{x - x_0} = 0,$$

vgl. Abbildung 20.
Ersetzt man die Ableitung $f'(x_0)$ durch eine beliebige andere reelle Zahl, so ist der Grenzwert des Quotienten $r(x)/(x - x_0)$ von Null verschieden. In diesem Sinne liefert die Ableitung $f'(x_0)$ die beste lineare Approximation der Funktion f in der Nähe des Punktes x_0.

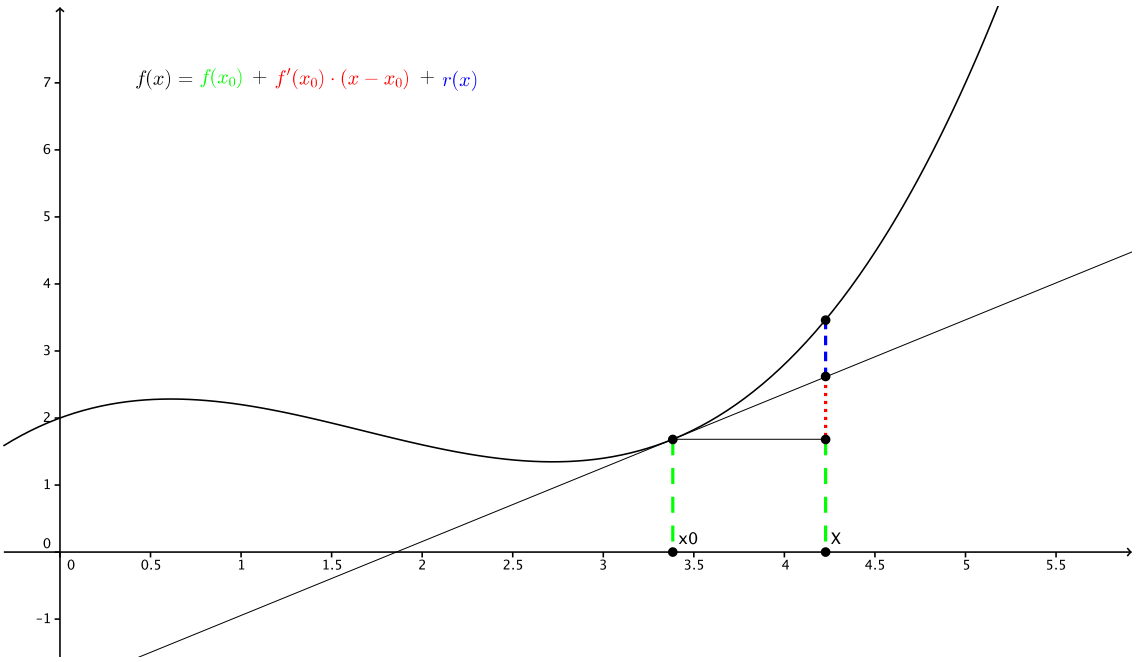


Abbildung 20: Die Tangente an f im Punkt x_0 und der Rest r

11.7 Mittelwertsatz: Sei f differenzierbar auf dem Intervall I, dann gibt eine Zahl $\vartheta \in (0,1)$, sodass

$$f(x_0 + h) = f(x_0) + hf'(x_0 + \vartheta h),$$

sofern x_0, $x_0 + h \in I$. Setzt man $a = x_0$, $b = x_0 + h$ und $\xi = x_0 + \vartheta h$ für $h > 0$, dann erhält man

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

Es gibt also an einer Zwischenstelle ξ eine Tangente, die parallel zur Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ ist.

11.8 Folgerungen: Sei $f : I \to \mathbb{R}$ eine differenzierbare Funktion.

- Wenn $f'(x) = 0$ für alle $x \in I$, dann gilt

  $$f(x_0 + h) = f(x_0) + h \cdot 0 = f(x_0)$$

  für alle h. Die Funktion f ist also konstant.

- Wenn $f'(x) > 0$ für alle $x \in I$, dann gilt

  $$f(x_0 + h) = f(x_0) + hf'(x_0 + \vartheta h) > f(x_0)$$

  für alle $h > 0$. Die Funktion f ist also streng monoton wachsend. Genauso ist f streng monoton fallend, wenn $f'(x) < 0$.

- Für die Aweichung r zwischen Funktion und Tangente [→ 11.6] gilt

  $$r(h) = h\big(f'(x_0 + \vartheta h) - f'(x_0)\big).$$

11.9 Beispiel: Sei $f(x) = \sin x$. Mit $f'(x) = \cos x$ erhält man im Punkt $x_0 = 0$ die Tangente

$$t(h) = h \quad \text{und damit} \quad r(h) = \sin h - h = h\big(\cos(\vartheta h) - 1\big).$$

Für $|h| \le 1/10$ gilt dann die Abschätzung

$$|\sin h - h| \le |h| \cdot (1 - \cos h) \le |h| \cdot (1 - \cos \tfrac{1}{10}) < \frac{|h|}{200}.$$

11.10 Regel von l'Hospital: Wenn f und g stetige Funktionen sind, dann gilt

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{f(x_0)}{g(x_0)}, \quad \text{falls} \quad g(x_0) \neq 0.$$

Im Fall $f(x_0) \neq 0$, $g(x_0) = 0$ existiert kein Grenzwert. Wenn aber f und g an der Stelle x_0 differenzierbar sind, dann erhält man für $f(x_0) = g(x_0) = 0$ gemäß 11.6 mit $x = x_0 + h$

$$f(x) = hf'(x_0) + r_1(h)$$
$$g(x) = hg'(x_0) + r_2(h)$$

und damit

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{h \to 0} \frac{f'(x_0) + r_1(h)/h}{g'(x_0) + r_2(h)/h} = \frac{f'(x_0)}{g'(x_0)}, \quad \text{falls} \quad g'(x_0) \neq 0.$$

Allgemeiner gilt: Wenn f und g differenzierbare Funktionen sind und

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = 0$$

oder

$$\lim_{x \to x_0} f(x) = \lim_{x \to x_0} g(x) = \infty,$$

dann ist

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{f'(x)}{g'(x)}.$$

Diese Regeln gelten auch vollkommen analog für Grenzwerte der Form $x \to \infty$.

11.11 Beispiel:

- $$\lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{\cos x}{1} = 1$$
- $$\lim_{x \to \infty} \frac{x^2}{e^x} = \lim_{x \to \infty} \frac{2x}{e^x} = \lim_{x \to \infty} \frac{2}{e^x} = 0$$
- $$\lim_{x \downarrow 0} x \ln x = \lim_{x \downarrow 0} \frac{\ln x}{1/x} = \lim_{x \downarrow 0} \frac{1/x}{-1/x^2} = \lim_{x \downarrow 0} (-x) = 0$$
- $$\lim_{x \to 1} \frac{\ln x - x + 1}{\cos(2 - 2x) - 1} = \lim_{x \to 1} \frac{1/x - 1}{2\sin(2 - 2x)} = \lim_{x \to 1} \frac{-1/x^2}{-4\cos(2 - 2x)} = \frac{1}{4}$$

11.12 Stetige Differenzierbarkeit: Eine Funktion $f : I \to \mathbb{R}$ heißt *stetig differenzierbar*, wenn sie differenzierbar ist und die Ableitungsfunktion f' zudem stetig ist.

11.13 Beispiel:

- Die Funktion $f(x) = \sqrt{|x|^3}, x \in \mathbb{R}$, ist stetig differenzierbar, denn die Ableitungsfunktion
$$f'(x) = \begin{cases} \frac{3}{2}\sqrt{x} & \text{für } x > 0 \\ 0 & \text{für } x = 0 \\ \frac{3}{2}\sqrt{-x} & \text{für } x < 0 \end{cases}$$
ist stetig.

- Die Funktion
$$g(x) = \begin{cases} x^2 \sin(1/x) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$$
ist differenzierbar mit
$$g'(x) = \begin{cases} 2x \sin(1/x) - \cos(1/x) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$$
Sie ist aber nicht stetig differenzierbar, da der Grenzwert $\lim_{x\to 0} g'(x)$ nicht existiert.

  Das hier zu beobachtende Verhalten ist typisch. Ableitungsfunktionen differenzierbarer Funktionen weisen niemals Sprungstellen auf. Unstetigkeiten sind vielmehr stets durch ein oszillierendes Verhalten der Ableitungsfunktionen bedingt, das zu einer Nichtexistienz von Grenzwerten führt.

11.14 Höhere Ableitungen: Eine Funktion $f : I \to \mathbb{R}$ heißt *k-mal differenzierbar*, wenn die Ableitungen in der Rekursion
$$\begin{aligned} f^{(0)} &:= f \\ f^{(n)} &:= \left(f^{(n-1)}\right)' \end{aligned}$$
für alle $n = 1, \ldots, k$ existieren. Die Funktion f heißt *unendlich oft differenzierbar*, wenn diese Ableitungen für alle $n \in \mathbb{N}$ existieren. Man nennt $f^{(k)}$ die k-te Ableitung von f und schreibt auch
$$f' = f^{(1)}, \quad f'' = f^{(2)}, \quad f''' = f^{(3)}$$
für Ableitungen niedriger Ordnung. Andere gebräuchliche Schreibweisen sind
$$f^{(k)} = D^k f = \frac{d^k f}{dx^k}.$$
Die Funktion f heißt *k-mal stetig differenzierbar*, wenn die Funktion $f^{(k)}$ stetig ist. Die Menge aller dieser Funktionen wird mit $C^k(I)$ bezeichnet. Insbesondere ist $C^0(I)$ die Menge der stetigen Funktionen auf dem Intervall I. Die Menge aller unendlich oft differenzierbaren Funktionen wird mit $C^\infty(I)$ bezeichnet.

11.15 Beispiel [→ 11.13]:

- Polynome und alle elementaren Funktionen $(\exp, \sin, \cos, \ln)$ sind auf ihrem Definitionsgebiet unendlich oft differenzierbar. Es gilt z.B.

$$f(x) = x^n \quad \Rightarrow \quad f^{(k)}(x) = \begin{cases} \frac{n!}{(n-k)!}\, x^{n-k} & \text{für } 0 \le k \le n \\ 0 & \text{für } k > n \end{cases}$$

$$f(x) = e^{2x} \quad \Rightarrow \quad f^{(k)}(x) = 2^k e^{2x}$$

$$f(x) = \sin x \quad \Rightarrow \quad f'(x) = \cos x, \quad f''(x) = -\sin x, \quad f'''(x) = -\cos x$$

- Für die *Stutzfunktion*

$$H_n(x) = \begin{cases} x^n & \text{für } x \ge 0 \\ 0 & \text{für } x < 0 \end{cases}$$

  vom Grad n gilt $H_n \in C^{n-1}(\mathbb{R})$. Insbesondere erhält man für $n = 0$ die Heaviside-Funktion [→ 10.5]. Diese ist unstetig und man schreibt $H_0 \in C^{-1}(\mathbb{R})$.

- Die Funktion f gemäß Beispiel 11.13 ist stetig differenzierbar, aber nicht zweimal differenzierbar, $f \in C^1(\mathbb{R})$.

- Die Funktion g gemäß Beispiel 11.13 ist stetig, aber nicht stetig differenzierbar, $g \in C^0(\mathbb{R})$.

12 Integration

12.1 Stammfunktion: Eine stetige Funktion $F : I \to \mathbb{R}$ heißt *Stammfunktion von* f auf dem Intervall I, wenn

$$F'(x) = f(x) \quad \text{für fast alle} \quad x \in I.$$

Dabei heißt „fast alle“, dass es endlich viele Stellen $x_1, \ldots, x_n \in I$ geben darf, an denen F nicht differenzierbar ist oder an denen F' nicht mit f übereinstimmt. Insbesondere muss die Funktion f an diesen Stellen noch nicht einmal definiert sein.

12.2 Beispiel:

- $$f(x) = 2x, \quad F(x) = x^2$$
- $$f(x) = 2x, \quad F(x) = (x-1)(x+1)$$
- $$f(x) = \begin{cases} 1 & \text{für } x > 0 \\ 7 & \text{für } x = 0 \\ -1 & \text{für } x < 0 \end{cases}, \quad F(x) = |x|$$
- $$f(x) = \frac{\sqrt{|x|}}{x}, \quad F(x) = 2\sqrt{|x|} \tag{12.1}$$

  Man beachte, dass die Funktion F auf ganz $\mathbb{R}$ definiert ist, während f an der Stelle $x = 0$ eine Definitionslücke hat.

12.3 Mehrdeutigkeit: Wenn F eine Stammfunktion von f ist, dann ist auch $\tilde{F} := F + c$ eine Stammfunktion, wobei $c \in \mathbb{R}$ eine beliebige Konstante ist. Umgekehrt gilt: Wenn F und $\tilde{F}$ Stammfunktionen von f sind, dann ist $\tilde{F}' - F' = 0$ und damit $\tilde{F} - F$ nach dem Mittelwertsatz konstant.

12.4 Unbestimmtes Integral: Die Menge aller Stammfunktionen einer Funktion f wird als das *unbestimmte Integral* von f bezeichnet,

$$\int f := \{F : F \text{ ist Stammfunktion von } f\}.$$

Wenn F eine beliebige Stammfunktion von f ist, dann gilt gemäß 12.3

$$\int f = \{F + c : c \in \mathbb{R}\}.$$

Meist wird das unbestimmte Integral nicht als Menge, sondern in der Form

$$\int f = F + c, \quad c \in \mathbb{R}$$

geschrieben. Gelegentlich wird in Formelsammlungen oder in Computeralgebra-Systemen (z.B. Maple) auf die Angabe der Integrationskonstanten c sogar ganz verzichtet. Wenn die unabhängige Veränderliche notiert werden soll, schreibt man

$$\int f(x)\,dx = F(x) + c, \quad c \in \mathbb{R}.$$

Man bezeichnet dann x auch als *Integrationsvariable* und dx als *Differenzial*.

12.5 Beispiel:

- $$\int e^{2x}\,dx = e^{2x}/2 + c, \quad c \in \mathbb{R}$$
- $$\int \cos = \sin + c, \quad c \in \mathbb{R}$$

12.6 Regeln: Seien F und G Stammfunktionen von f und g, und sei $c \in \mathbb{R}$.

- *Linearität:* Für reelle Zahlen α, β gilt
$$\int (\alpha f + \beta g) = \alpha F + \beta G + c$$
- *Partielle Integration:* (vgl. Produktregel)
$$\int fG = FG - \int Fg$$
- *Substitution:* (vgl. Kettenregel)
$$\int (f \circ G)g = F \circ G + c$$

12.7 Beispiel:

- Linearität:
$$\int (\sin x + 2/x) = -\cos x + 2\ln x + c, \quad x > 0$$
- Partielle Integration: Mit $f(x) = 1$ und $G(x) = \ln x$ gilt für $x > 0$
$$\int \ln x\,dx = \int 1 \cdot \ln x\,dx = x \ln x - \int x/x\,dx = x \ln x - x + c.$$
- Substitution: Mit $G(x) = \alpha x$ gilt
$$\int f(\alpha x)\alpha\,dx = F(\alpha x) + c$$

12.8 Wichtige Stammfunktionen:

f	F	
x^n	$\frac{1}{n+1}\,x^{n+1}$	$x \in \mathbb{R}$, $n \in \mathbb{N}_0$
x^α	$\frac{1}{\alpha+1}\,x^{\alpha+1}$	$x > 0$, $\alpha < -1$
x^α	$\frac{1}{\alpha+1}\,x^{\alpha+1}$	$x \geq 0$, $\alpha > -1$
$\frac{1}{x}$	$\ln\lvert x\rvert$	$x > 0$ oder $x < 0$
e^x	e^x	$x \in \mathbb{R}$
$\ln x$	$x \ln x - x$	$x > 0$
$\sin x$	$-\cos x$	$x \in \mathbb{R}$
$\cos x$	$\sin x$	$x \in \mathbb{R}$
$\frac{1}{1+x^2}$	$\arctan x$	$x \in \mathbb{R}$

12.9 Partialbruchzerlegung: Integrale der Form

$$\int \frac{p(x)}{q(x)}\,dx$$

mit Polynomen

$$p(x) = p_0 + p_1 x + \cdots + p_m x^m, \quad q(x) = q_0 + q_1 x + \cdots + q_n x^n, \quad q_n \neq 0,$$

lassen sich durch Partialbruchzerlegung bestimmen. Wir betrachten zunächst den Fall $m < n$.

1. Man bestimmt alle Nullstellen $x_1, \ldots, x_n$ von q und erhält damit die Faktorisierung

$$q(x) = q_n(x - x_1) \cdots (x - x_n).$$

2. Wenn alle Nullstellen verschieden sind, macht man den Ansatz

$$\frac{p(x)}{q(x)} = \frac{A_1}{x - x_1} + \cdots + \frac{A_n}{x - x_n}.$$

3. Man multipliziert den Ansatz mit $q(x)$ und erhält dann die unbekannten Koeffizienten $A_1, \ldots, A_n$ einfach durch Einsetzen der Nullstellen.

4. Für das Integral ergibt sich damit im Fall reeller Nullstellen

$$\int \frac{p(x)}{q(x)}\,dx = A_1 \ln\lvert x - x_1\rvert + \cdots + A_n \ln\lvert x - x_n\rvert + c.$$

Das Verfahren besitzt die folgenden Modifikationen:

- Im Fall $m \geq n$ liefert eine Polynomdivision

$$\frac{p(x)}{q(x)} = r(x) + \frac{\tilde{p}(x)}{q(x)},$$

wobei r eine Polynom vom Grad $m - n$ und der Grad $\tilde{m}$ von $\tilde{p}$ kleiner als n ist.

- Wenn z.B. $x_1 = \cdots = x_\ell$ eine mehrfache Nullstelle ist, dann sind die zugehörigen Summanden im Ansatz zu ersetzen durch

$$\frac{A_1}{x - x_1} + \frac{A_2}{(x - x_1)^2} + \cdots + \frac{A_\ell}{(x - x_1)^\ell}.$$

In diesem Fall ist die Koeffizientenbestimmung durch Einsetzen der Nullstellen nur für A_ℓ möglich. Die übrigen Koeffizienten können beispielsweise durch Ausmultiplizieren und Koeffizientenvergleich bestimmt werden.

- Wenn es komplexe Nullstellen gibt, dann ist das Verfahren im Prinzip ebenfalls anwendbar. Es ist dann lediglich bei der Bildung der Stammfunktion der Partialbrüche der *komplexe Logarithmus*

$$\ln_{\mathbb{C}}(x - x_k) := \ln|x - x_k| + i \arg(x - x_k)$$

zu verwenden. Da komplexe Nullstellen und damit auch die zugehörigen Partialbrüche immer in konjugiert komplexen Paaren auftreten, heben sich die Imaginärteile stets weg. Alternativ zur Verwendung komplexer Logarithmen kann man im Ansatz auch konjugiert komplexe Terme in der Form

$$\frac{A_1}{x - x_1} + \frac{A_2}{x - \overline{x}_1} \longrightarrow \frac{B_1 + B_2 x}{x^2 - 2x \operatorname{Re} x_1 + |x_1|^2}$$

zu einem reellen Summanden mit quadratischem Nenner zusammenfassen. Dessen Integral lässt sich dann mit Hilfe der reellen Logarithmus- und der Arcustangens-Funktion angeben.

12.10 Beispiel:

- Für

$$\frac{p(x)}{q(x)} = \frac{x^3 - x + 3}{x^2 - x - 2}$$

liefert Polynomdivision zunächst

$$\frac{p(x)}{q(x)} = x + 1 + \frac{2x + 5}{x^2 - x - 2}.$$

Wir betrachten nun nur den letzen Summanden. Die Nullstellen des Nenners sind $x_1 = 2$ und $x_2 = -1$, also $q(x) = x^2 - x - 2 = (x - 2)(x + 1)$. Der Ansatz

$$\frac{2x + 5}{(x - 2)(x + 1)} = \frac{A_1}{x - 2} + \frac{A_2}{x + 1}$$

liefert nach Multiplikation mit $q(x)$ die Gleichung

$$2x + 5 = (x + 1)A_1 + (x - 2)A_2.$$

Einsetzen der Nullstelle $x_1 = 2$ liefert $A_1 = 3$, und $x_2 = -1$ liefert $A_2 = -1$. Also ist

$$\frac{p(x)}{q(x)} = x + 1 + \frac{3}{x-2} - \frac{1}{x+1}$$

und

$$\int \frac{x^3 - x + 3}{x^2 - x - 2}\,dx = \frac{x^2}{2} + x + 3\ln|x-2| - \ln|x+1| + c, \quad c \in \mathbb{R}.$$

- Für

$$\frac{p(x)}{q(x)} = \frac{x^2 + 4}{x^3 - 2x^2}$$

erhält man die Nullstellen $x_1 = x_2 = 0$ und $x_3 = 2$, also $q(x) = x^3 - 2x^2 = x^2(x-2)$. Der Ansatz

$$\frac{p(x)}{q(x)} = \frac{A_1}{x} + \frac{A_2}{x^2} + \frac{A_3}{x-2}$$

liefert nach Multiplikation mit $q(x)$ die Gleichung

$$x^2 + 4 = x(x-2)A_1 + (x-2)A_2 + x^2 A_3.$$

Einsetzen der Nullstellen $x_1 = 0$ und $x_3 = 2$ liefert $A_2 = -2$ und $A_3 = 2$. Damit erhält der x^2-Term auf der rechten Seite den Koeffizienten $A_1 + 2$ und auf der linken Seite den Koeffizienten 1. Also ist $A_1 = -1$. Wir erhalten

$$\frac{p(x)}{q(x)} = -\frac{1}{x} - \frac{2}{x^2} + \frac{2}{x-2}$$

und

$$\int \frac{x^2 + 4}{x^3 - 2x^2}\,dx = -\ln|x| + \frac{2}{x} + 2\ln|x-2| + c, \quad c \in \mathbb{R}.$$

12.11 Bestimmtes Integral: Sei $F : I \to \mathbb{R}$ eine Stammfunktion von f. Wenn $[a, b] \subset I$, dann bezeichnet man

$$\int_a^b f(x)\,dx := F(x)\Big|_a^b := F(b) - F(a)$$

als das *bestimmte Integral* der Funktion f mit Integrationsgrenzen a, b. Sei $\tilde{F} = F + c$ eine andere Stammfunktion von f, dann erhält man ebenfalls

$$\int_a^b f(x)\,dx = \tilde{F}(b) - \tilde{F}(a) = \big(F(b) - c\big) - \big(F(a) - c\big) = F(b) - F(a).$$

Der Wert des bestimmten Integrals ist also nicht von der Wahl der Stammfunktion abhängig.

12.12 Regeln:

- *Linearität:* Für reelle Zahlen α, β gilt

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

- *Partielle Integration:*

$$\int_a^b f(x)G(x)\, dx = F(x)G(x)\big|_a^b - \int_a^b F(x)g(x)\, dx$$

- *Substitution:*

$$\int_a^b f(G(x))g(x)\, dx = F(G(x))\big|_a^b$$

  Setzt man $G(x) = u$, dann gilt weiter

$$F(G(x))\big|_{x=a}^b = F(u)\big|_{u=G(a)}^{G(b)} = \int_{u=G(a)}^{G(b)} f(u)\, du$$

  und damit

$$\int_{x=a}^b f(G(x))g(x)\, dx = \int_{u=G(a)}^{G(b)} f(u)\, du.$$

  Bei der Substititution $G(x) = u$ wird also mit Hilfe der Ableitung $G'(x) = g(x)$ formal $g(x)\, dx$ durch du ersetzt. Die Integrationsgrenzen transformieren sich durch Einsetzen in die Funktion G, also $x = a \;\Rightarrow\; u = G(a)$ und $x = b \;\Rightarrow\; u = G(b)$.

  Wenn H eine *invertierbare* Funktion ist, dann kann man $G := H^{-1}$ setzen und erhält die Regel

$$\int_{x=a}^b f(x)\, dx = \int_{u=H^{-1}(a)}^{H^{-1}(b)} f(H(u))h(u)\, du.$$

  Bei der Substititution $x = H(u)$ wird also mit Hilfe der Ableitung $H'(u) = h(u)$ formal dx durch $h(u)\, du$ ersetzt. Bei der Transformation der Integrationsgrenzen sind diejenigen u-Werte zu suchen, die den gegebenen x-Werten entsprechen, $x = a \;\Rightarrow\; u = H^{-1}(a)$ und $x = b \;\Rightarrow\; u = H^{-1}(b)$.

12.13 Beispiel:

- Linearität:

$$\int_0^\pi (\sin x + 2x)\, dx = \int_0^\pi \sin x\, dx + 2\int_0^\pi x\, dx = -\cos(x)\big|_0^\pi + x^2\big|_0^\pi = 2 + \pi^2$$

- Partielle Integration:

$$\int_0^1 xe^{2x}\, dx = \frac{x}{2}e^{2x}\Big|_0^1 - \frac{1}{2}\int_0^1 e^{2x}\, dx = \frac{e^2}{2} - \frac{1}{4}\,e^{2x}\Big|_0^1 = \frac{e^2}{2} - \frac{e^2}{4} + \frac{1}{4} = \frac{e^2}{4} + \frac{1}{4}$$

- Substitution $G(x) = u$: Zur Berechnung von $\int_0^2 x\cos(x^2)\,dx$ setzt man

$$G(x) = x^2 = u, \quad g(x) = 2x, \quad 2x\,dx = du.$$

Die Integrationsgrenzen transformieren sich gemäß

$$x = 0 \;\Rightarrow\; u = 0, \quad x = 2 \;\Rightarrow\; u = 4.$$

Damit erhält man

$$\int_{x=0}^{2} x\cos(x^2)\,dx = \frac{1}{2}\int_{u=0}^{4}\cos(u)\,du = \frac{1}{2}\,\sin(u)\Big|_{u=0}^{4} = \frac{1}{2}\,\sin(4).$$

- Substitution $x = H(u)$: Zur Berechnung von $\int_0^1 \sqrt{1-x^2}\,dx$ setzt man

$$x = H(u) = \sin(u), \quad h(u) = \cos(u), \quad dx = \cos(u)\,du.$$

Dabei verwendet man das Definitionsgebiet $u \in [0, \pi/2]$, sodass die Funktion H injektiv ist. Die Integrationsgrenzen transformieren sich gemäß

$$x = 0 \;\Rightarrow\; u = 0, \quad x = 1 \;\Rightarrow\; u = \pi/2.$$

Damit erhält man

$$\int_{x=0}^{1}\sqrt{1-x^2}\,dx = \int_{u=0}^{\pi/2}\sqrt{1-\sin^2(u)}\cos(u)\,du = \int_{u=0}^{\pi/2}\cos^2(u)\,du = \pi/4.$$

12.14 Variable obere Integrationsgrenze: Definiert man die Funktion $F_a : I \to \mathbb{R}$ gemäß

$$F_a(t) := \int_a^t f(x)\,dx, \quad t \in I,$$

und ist F eine Stammfunktion von f, dann gilt $F_a(t) = F(t) - F(a)$. Folglich ist auch F_a eine Stammfunktion von f, die durch $F_a(a) = 0$ charakterisiert ist. *Achtung:* Schreibweisen wie

$$\int_a^x f(x)\,dx,$$

bei denen die Integrationsvariable auch in den Grenzen auftaucht, sind nicht zulässig.

12.15 Additivität: Es gilt

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

und insbesondere

$$\int_a^b f(x)\,dx + \int_b^a f(x)\,dx = \int_a^a f(x)\,dx = 0.$$

12.16 Monotonie: Sei $a \le b$ und

$$f(x) \le g(x) \quad \text{für alle} \quad x \in [a,b],$$

dann gilt aufgrund des Mittelwertsatzes

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

Ein wichtiger Spezialfall hiervon ist die *Dreiecksungleichung*

$$\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx.$$

12.17 Flächeninhalt: Sei $f : I \to \mathbb{R}$ eine stetige Funktion. Wir betrachten die Punktmenge, die in x-Richtung durch die Abszissen a und b und in y-Richtung durch die x-Achse und den Graphen von f begrenzt ist. Wenn wir dieser Menge in sinnvoller Weise einen (orientierten) Flächeninhalt $S_a^b f$ zuweisen wollen, dann muss dieser offensichtlich folgende Eigenschaften haben:

- *Additivität:* $S_a^b f + S_b^c f = S_a^c f$
- *Monotonie:* Aus $f \le g$ und $a \le b$ folgt $S_a^b f \le S_a^b g$.
- *Konstante:* Wenn $f(x) = f_0$ eine konstante Funktion ist, dann gilt $S_a^b f = f_0(b-a)$.

Tatsächlich ist allein durch diese drei Vorgaben der Ausdruck $S_a^b f$ eindeutig bestimmt: Sei $F_a(t) := S_a^t f$, dann gilt aufgrund der Additivität

$$F_a(t+h) - F_a(t) = S_a^{t+h} f - S_a^t f = S_t^{t+h} f.$$

Sei nun $h > 0$ beliebig gewählt (für $h \le 0$ verlaufen die Überlegungen vollkommen analog). Dann definieren wir die konstanten Funktionen $f_{\min}^h$ und $f_{\max}^h$ durch das Minimum bzw. das Maximum der Funktion f auf dem Intervall $[t, t+h]$,

$$f_{\min}^h := \min\{f(x) : t \le x \le t+h\}, \quad f_{\max}^h := \max\{f(x) : t \le x \le t+h\}.$$

Es ist also

$$f_{\min}^h \le f(x) \le f_{\max}^h \quad \text{für alle} \quad x \in [t, t+h]$$

und damit folgt aufgrund der Monotonie

$$S_t^{t+h} f_{\min}^h \le F_a(t+h) - F_a(t) \le S_t^{t+h} f_{\max}^h.$$

Da es sich auf der linken und rechten Seite der Ungleichung um konstante Funktionen handelt, erhalten wir

$$h f_{\min}^h \le F_a(t+h) - F_a(t) \le h f_{\max}^h.$$

Wenn nun h gegen 0 geht, folgt aus der Stetigkeit von f

$$\lim_{h\to 0} f_{\min}^h = \lim_{h\to 0} f_{\max}^h = f(t)$$

und damit

$$F_a'(t) = \lim_{h\to 0} \frac{F_a(t+h) - F_a(t)}{h} = f(t).$$

F_a muss also eine Stammfunktion von f sein. Da außerdem $F_a(a) = S_a^a f = 0$ gilt, folgt

$$S_a^t f = F_a(t) = \int_a^t f(x)\,dx.$$

Diesen Zusammenhang nennt man auch den *Hauptsatz der Differenzial- und Integralrechnung*. Insbesondere ergibt sich für den Flächeninhalt zwischen den Abszissen a und b

$$S_a^b f = \int_a^b f(x)\,dx.$$

12.18 Andere Interpretationen des Integrals: In der vorangegangenen Herleitung des Flächeninhalts als bestimmtes Integral wurden nur die Additivität und die Monotonie sowie die Information über konstante Funktionen ausgenutzt. Diese drei Eigenschaften sind bei der Betrachtung des Flächeninhalts besonders offensichtlich, doch charakterisieren sie in vollkommen analoger Weise zahlreiche andere Phänomene:

- Sei $f(x)$ die Geschwindigkeit eines Teilchens zum Zeitpunkt x und $S_a^b f$ der zwischen den Zeitpunkten a und b zurückgelegte Weg, dann gilt:
  - *Additivität,* zurückgelegte Wegstrecken addieren sich.
  - *Monotonie,* wer schneller fährt, kommt weiter.
  - *Konstante Geschwindigkeit $f(x) = f_0$ liefert $S_a^b f = f_0(b-a)$,* das ist die Definition der Geschwindigkeit einer gleichförmigen Bewegung.
- Sei $f(x)$ die spezifische Dichte eines Drahtes mit Querschnitt 1 an der Stelle x und $S_a^b f$ die Masse des Drahtstücks zwischen den Punkten a und b, dann gilt:
  - *Additivität,* Teilmassen addieren sich.
  - *Monotonie,* was überall schwerer ist, ist auch insgesamt schwerer.
  - *Konstante Dichte $f(x) = f_0$ liefert $S_a^b f = f_0(b-a)$,* das ist die Definition der Dichte eines homogenen Materials.
- Sei $f(x)$ die Kraft, die auf ein Teilchen der Masse 1 an der Stelle x wirkt und $S_a^b f$ die Energie des Teilchens, nachdem es die Strecke von a nach b zurückgelegt hat, dann gilt:
  - *Additivität,* Teilenergien addieren sich.
  - *Monotonie,* größere Kraft ergibt höhere Energie.
  - *Konstante Kraft $f(x) = f_0$ liefert $S_a^b f = f_0(b-a)$,* das ist die Definition der Energie bei konstanter Krafteinwirkung.

In allen drei Fällen ergibt sich der Zusammenhang

$$S_a^b f = \int_a^b f(x)\,dx.$$

12.19 Existenz einer Stammfunktion: Wir wollen nun zeigen, dass jede stetige Funktion tatsächlich eine Stammfunktion besitzt. Sei also $f : [a, b] \to \mathbb{R}$ eine stetige Funktion. Eine Funktion u, die stückweise konstant und überall kleiner oder gleich f ist, heißt *untere Treppenfunktion* von f. Sind genauer die Abszissen $a = x_0 < x_1 < \cdots < x_n = b$ sowie die Funktionswerte $u_0, \ldots, u_n$ gegeben, dann hat die zugehörige Treppenfunktion die Form

$$u(x) = u_i \leq f(x) \quad \text{für} \quad x_i \leq x < x_{i+1}.$$

Wir definieren nun in naheliegender Weise die *Untersumme*

$$S_a^b u := \sum_{i=0}^{n-1} u_i (x_{i+1} - x_i).$$

Die Menge aller unteren Treppenfunktionen bezeichnen wir mit U_f. Sei M das Maximum von f auf dem Intervall $[a, b]$, dann gilt sicher $S_a^b u < M(b - a)$ für alle $u \in U_f$. Also ist die Menge aller Werte $S_a^b u$ mit $u \in U_f$ nach oben beschränkt und damit ein Intervall der Form

$$\{S_a^b u : u \in U_f\} = (-\infty, s) \quad \text{oder} \quad \{S_a^b u : u \in U_f\} = (-\infty, s].$$

Wenn f eine konstante Funktion $f(x) = f_0$ ist, dann erhalten wir den Wert $s = S_a^b f = f_0(b-a)$. Wir bezeichnen nun ganz allgemein für eine stetige Funktion f die Obergrenze der Intervalls des Untersummen mit

$$S_a^b f := s.$$

Man überzeugt sich leicht, dass $S_a^b f$ sowohl die Eigenschaft der Additivität als auch der Monotonie hat. Somit ist $F_a(t) := S_a^t f$ eine Stammfunktion von f. Vollkommen analoge Überlegungen lassen sich für obere Treppenfunktionen und die daraus resultierenden Obersummen anstellen. Insbesondere sieht man leicht, dass die Obergrenze der Untersummen mit der Untergrenze der Obersummen übereinstimmt, da beide Größen auf dieselbe Stammfunktion von f führen.

12.20 Uneigentliche Integrale: Man bezeichnet Integrale als *uneigentlich*, wenn entweder der Integrand oder das Integrationsgebiet unbeschränkt sind. Der Fall unbeschränkter Integranden erfordert nach dem hier gewählten Zugang im Prinzip keine gesonderte Behandlung. Beispielsweise ist gemäß (12.1)

$$\int_0^1 \frac{dx}{\sqrt{x}} = 2\sqrt{x}\Big|_0^1 = 2.$$

Dagegen existiert das Integral

$$\int_0^1 \frac{dx}{x}$$

nicht, da sich die Stammfunktion $\ln x$ nicht stetig in den Ursprung fortsetzen lässt. Für

$$\int_0^{1/2} \frac{dx}{x \ln^2 x}$$

erhält man zunächst die Stammfunktion $-1/\ln x$, die an der Stelle $x = 0$ nicht definiert ist. Allerdings existiert der Grenzwert $\lim_{x\to 0} 1/\ln x = 0$, sodass die Funktion

$$F(x) = \begin{cases} -1/\ln x & \text{für } x > 0 \\ 0 & \text{für } x = 0 \end{cases}$$

stetig und damit eine Stammfunktion von f ist. Damit erhält man

$$\int_0^{1/2} \frac{dx}{x \ln^2 x} = F(x)\Big|_0^{1/2} = -\frac{1}{\ln(1/2)} - 0 = \frac{1}{\ln 2}.$$

Die Grenzwertbildung kann auch direkt in die Integration einbezogen werden,

$$\int_0^{1/2} \frac{dx}{x \ln^2 x} = \lim_{a\to 0}\Big(\frac{-1}{\ln x}\Big|_a^{1/2}\Big) = -\frac{1}{\ln(1/2)} - \lim_{a\to 0}\frac{-1}{\ln a} = \frac{1}{\ln 2}.$$

Im Fall unbeschränkter Integrationsgebiete erklärt man die zugehörigen uneigentlichen Integrale mit Hilfe von Grenzwerten, sofern diese existieren:

- Sei $F : (a, \infty) \to \mathbb{R}$ eine Stammfunktion von f, dann ist

$$\int_a^\infty f(x)\,dx := F(x)\big|_a^\infty := \lim_{b\to\infty} F(b) - F(a).$$

- Sei $F : (-\infty, b) \to \mathbb{R}$ eine Stammfunktion von f, dann ist

$$\int_{-\infty}^b f(x)\,dx := F(x)\big|_{-\infty}^b := F(b) - \lim_{a\to-\infty} F(a).$$

- Sei $F : (-\infty, \infty) \to \mathbb{R}$ eine Stammfunktion von f, dann ist

$$\int_{-\infty}^\infty f(x)\,dx := F(x)\big|_{-\infty}^\infty := \lim_{b\to\infty} F(b) - \lim_{a\to-\infty} F(a).$$

12.21 Beispiel:

- Für $\alpha > 1$ ist

$$\int_1^\infty \frac{dx}{x^\alpha} = \frac{1}{(1-\alpha)x^{\alpha-1}}\Big|_1^\infty = \frac{1}{1-\alpha}\left(\lim_{b\to\infty}\frac{1}{b^{\alpha-1}} - 1\right) = \frac{1}{\alpha-1}.$$

  Für $\alpha \le 1$ existiert das uneigentliche Integral nicht.

- Für $\alpha < 1$ ist

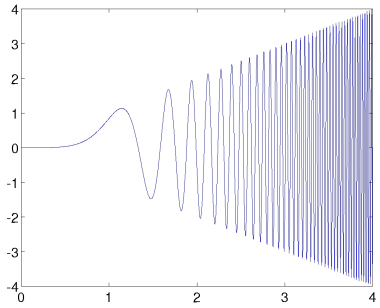
$$\int_0^1 \frac{dx}{x^\alpha} = \frac{1}{(1-\alpha)x^{\alpha-1}}\Big|_0^1 = \frac{1}{1-\alpha}\,(1-0) = \frac{1}{1-\alpha}.$$

  Für $\alpha \ge 1$ existiert das uneigentliche Integral nicht.

-

$$\begin{aligned}\int_{-\infty}^\infty \frac{dx}{1+x^2} &= \arctan x\Big|_{-\infty}^\infty = \lim_{b\to\infty}\arctan b - \lim_{a\to-\infty}\arctan a \\ &= \pi/2 - (-\pi/2) = \pi\end{aligned}$$

12.22 Existenzkriterien: Anders als bei der Existenz von Grenzwerten von Reihen können uneigentliche Integrale von Typ $\int_a^\infty f(x)\,dx$ selbst dann existieren, wenn der Integrand unbeschränkt anwächst. Es ist beispielsweise (ohne dass wir das hier nachrechnen könnten)



$$\int_0^\infty x\sin(x^4)\,dx = \frac{\sqrt{2\pi}}{8}\ .$$

Vollkommen analog zur Untersuchung von Reihen sind aber die folgenden Vergleichskriterien:

- *Konvergente Majorante:* Wenn $\int_a^\infty f(x)\,dx$ konvergiert und

  $$|g(x)| \le f(x) \quad \text{für alle} \quad x \ge x_0$$

  für ein $x_0 \in \mathbb{R}$, dann konvergiert auch $\int_{\tilde{a}}^\infty g(x)\,dx$.

- *Divergente Minorante:* Wenn $\int_a^\infty f(x)\,dx$ divergiert und

  $$0 \le f(x) \le g(x) \quad \text{für alle} \quad x \ge x_0$$

  für ein $x_0 \in \mathbb{R}$, dann divergiert auch $\int_{\tilde{a}}^\infty g(x)\,dx$.

Vollkommen analoge Kriterien gelten auch für uneigentliche Integrale mit unbeschränktem Integranden.

12.23 Beispiel:

$$\int_1^\infty \frac{\sin x}{x^2}\,dx$$

existiert, da

$$\left|\frac{\sin x}{x^2}\right| \le \frac{1}{x^2}$$

und $\int_1^\infty \frac{dx}{x^2}$ nach Beispiel 12.21 existiert.

12.24 Integration und Reihenbildung: Integration und Reihenbildung weisen zahlreiche Analogien auf. Sei $(a_n)_n$ eine Folge und f eine stetige Funktion, dann ist

$$F_0(t) := \int_0^t f(x)\,dx, \quad s_m := \sum_{n=0}^m a_n.$$

Der Differenziation der Stammfunktion entspricht eine Differenzenbildung der Partialsummen,

$$F_0'(t) = f(t), \quad s_m - s_{m-1} = a_m.$$

Das uneigentliche Integral und der Reihenwert entstehen durch Grenzwertbildung,

$$\int_0^\infty f(x)\,dx = \lim_{t\to\infty} F_0(t), \quad \sum_{n=0}^{\infty} a_n = \lim_{m\to\infty} s_m.$$

Definiert man die Funktion f als stückweise konstant mit Werten

$$f(x) = a_n \quad \text{für} \quad n \le x < n+1,$$

dann gilt

$$s_m = \sum_{n=0}^{m} a_n = \int_0^{m+1} f(x)\,dx = F_0(m+1), \quad m \in \mathbb{N}_0.$$

Hieraus lässt sich das sogenannte *Integralkriterium* für die Konvergenz von Reihen ableiten: Sei $(a_n)_n$ eine Folge und f eine monoton fallende Funktion mit $f(n) = a_n$. Dann konvergiert das uneigentliche Integral $\int_0^\infty f(x)\,dx$ genau dann, wenn die Reihe $\sum_{n=0}^{\infty} a_n$ konvergiert.

13 Taylor-Reihen

Ein Taschenrechner wird gebaut. Er beherrscht bisher nur die Grundrechenarten $+, -, \cdot$ und $\div$ mit Gleitkommazahlen, d.h. reellen Zahlen. Nun soll $\sin(x)$ für ein $x \in \mathbb{R}$ berechnet werden. Wie soll der Taschnrechner das machen?
Die Funktion $\sin(x)$ ist ein Beispiel einer *elementaren Funktion*. Hier können die Funktionswerte durch *Potenzreihen* beliebig genau berechnet werden. Im Fall von $\sin(x)$ ist beispielsweise

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm$$

Den Ausdruck rechts kann der Taschenrechner schrittweise berechnen: er beginnt mit $y_0 := x$, berechnet dann $y_1 := y_0 - \frac{x^3}{3!}$, dann $y_2 := y_1 + \frac{x^5}{5!}$ etc. Der Taschenrechner muss wissen, wann er mit seinen Berechnungen aufhören kann, weil der gesuchte Wert $\sin(x)$ mit genügender Genauigkeit (etwa für das Display) erreicht wurde.
In diesem Abschnitt geht es daher um Potenzreihen, das Aufstellen der *Taylor-Reihe* Tf zu einer gegebenen Funktionen f wie hier am Beispiel von $\sin(x)$, die Frage, ob die Taylor-Reihe $Tf(x)$ im Punkt x immer $f(x)$ approximiert und wie man die Genauigkeit mit dem *Restglied* messen kann.

13.1 Potenzreihen: Eine Reihe der Form

$$p(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

heißt *Potenzreihe* mit *Entwicklungspunkt* x_0 und Koeffizienten a_n. Der zugehörige *Konvergenzradius* ist die reelle Zahl R mit der Eigenschaft: Die Potenzreihe ist

- absolut konvergent für alle x mit $|x - x_0| < R$,
- divergent für $|x - x_0| > R$.

Sollte die Potenzreihe sogar für alle $x \in \mathbb{R}$ absolut konvergieren setzen wir formal $R = \infty$. Sollte sie für kein $x \neq x_0$ absolut konvergieren setzen wir $R = 0$. Aus dem Wurzel- bzw. Quotientenkriterium für Reihen folgt: Wenn einer der beiden Grenzwerte

$$\left(\lim_{n\to\infty} \sqrt[n]{|a_n|} \right)^{-1} \quad \text{bzw.} \quad \left(\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} \right)^{-1}$$

existiert, so ist er gleich dem Konvergenzradius. Wenn der Grenzwert in der Klammer 0 ist, ist $R = \infty$. Falls die Folge in der Klammer unbeschränkt ist, ist $R = 0$. Das Verhalten für $|x - x_0| = R$, also die Randpunkte $x = x_0 \pm R$, muss gesondert untersucht werden. Die Menge aller x, für die die Potenzreihe p konvergiert, heißt *Konvergenzbereich* und wird mit I_p bezeichnet.

13.2 Beispiel:

- Für $p(x) = e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$ erhält man $R = \infty$ und $I_p = \mathbb{R}$ (vgl. 9.15).
- Für $\sum_{n=0}^{\infty} x^n$ erhält man $R = 1$ und $I_p = (-1, 1)$ (vgl. 9.2).

- Für $p(x) = \sum_{n=1}^{\infty} \frac{2^n}{n}(x-2)^n$ kann man den Konvergenzradius nach beiden Formeln berechnen,

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

  Für $x = 2 + R = 5/2$ erhält man eine divergente und für $x = 2 - R = 3/2$ eine nach dem Leibniz-Kriterium konvergente Reihe. Die Potenzreihe konvergiert also für $x \in I_p = [3/2, 5/2)$.

- Für $p(x) = \sum_{n=0}^{\infty} n!\, x^n$ erhält man $R = 0$ und somit $I_p = \{0\}$.

- Für $p(x) = \sum_{k=1}^{\infty} \frac{2^k}{k} x^{3k}$, also

$$p(x) = 2x^3 + 2x^6 + \frac{8}{3}x^9 + 4x^{12} + \cdots,$$

  kann man den Konvergenzradius nach keiner der beiden Formeln berechnen, denn hier ist die Koeffizientenfolge nach Potenzen von x

$$0, 0, 2, 0, 0, 2, 0, 0, \frac{8}{3}, 0, 0, 4, 0, 0, \frac{32}{5}, 0, 0, \ldots$$

  und somit existiert keiner der beiden Grenzwerte. Man kann sich hier behelfen indem man $y = x^3$ subsituiert. Man erhält $\sum_{k=1}^{\infty} \frac{2^k}{k} y^k$ mit Koeffizientenfolge nach Potenzen in y

$$2, 2, \frac{8}{3}, 4, \frac{32}{5}, \ldots.$$

  Wie im dritten Beispiel ergeben die Formeln, dass die Reihe in y Konvergenzradius $\frac{1}{2}$ besitzt. Wegen $y = x^3$ ist der Konvergenzradius der Reihe in x damit $R = \sqrt[3]{\frac{1}{2}}$.

13.3 Eigenschaften:

- Jede Potenzreihe konvergiert für $x = x_0$ gegen den Wert $p(x_0) = a_0$. In der Summenschreibweise $p(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ wird also formal $0^0 = 1$ gesetzt.

- Für $|x - x_0| < R$ ist die Funktion $p(x)$ beliebig oft stetig differenzierbar und die Ableitungen können durch gliedweises Differenzieren bestimmt werden,

$$p'(x) = \sum_{n=1}^{\infty} n a_n \, (x - x_0)^{n-1}$$
$$p^{(k)}(x) = \sum_{n=k}^{\infty} \tfrac{n!}{(n-k)!} a_n \, (x - x_0)^{n-k}.$$

  Der Summationsbereich ist also immer so gewählt, dass keine negativen Exponenten auftreten. Insbesondere gilt für alle $k \in \mathbb{N}_0$

$$p^{(k)}(x_0) = k!\, a_k. \tag{13.1}$$

- Unbestimmte und bestimmte Integrale können durch gliedweise Integration bestimmt werden. Für $a, b, x \in I_p$ gilt also

$$\int p(x)\,dx = c + \sum_{n=0}^{\infty} \frac{a_n}{n+1}\,(x-x_0)^{n+1}\,,$$
$$\int_a^b p(x)\,dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1}\,(b^{n+1}-a^{n+1}).$$

- Die Konvergenzradien der integrierten oder differenzierten Potenzreihen stimmen mit dem der ursprünglichen Potenzreihe überein, d.h., $R_p = R_{p'} = R_{\int p}$. Die Konvergenzbereiche können verschieden sein.

- *Abelscher Grenzwertsatz:* Ist eine Potenzreihe in einem Randpunkt konvergent, so ist sie auch dort (einseitig) stetig.

- *Eindeutigkeitssatz:* Zwei Potenzreihen mit demselben Entwicklungspunkt und positiven Konvergenzradien haben genau dann überall dieselben Funktionswerte, wenn alle Koeffizienten übereinstimmen.

13.4 Beispiel:

- Für die Exponentialreihe gilt $R = \infty$ und daher

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

und

$$\int e^x dx = c + \sum_{n=0}^{\infty} \frac{1}{n+1}\frac{x^{n+1}}{n!} = c + \sum_{n=0}^{\infty} \frac{x^{n+1}}{(n+1)!} = c - 1 + e^x.$$

- Für $p(x) = \sum_{n=0}^{\infty} x^n$ gilt $R = 1$, also ist $R = 1$ auch für

$$p'(x) = \sum_{n=0}^{\infty} nx^{n-1}, \qquad \int p(x) = c + \sum_{n=0}^{\infty} \frac{1}{n+1}x^{n+1} = c + \sum_{n=1}^{\infty} \frac{x^n}{n}.$$

13.5 Taylor-Reihe einer differenzierbaren Funktion: Sei f eine im Punkt x_0 beliebig oft differenzierbare Funktion. Wenn $p(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine Potenzreihe mit positivem Konvergenzradius ist, die mit f auf I_p übereinstimmt, so muss gemäß (13.1) gelten

$$f^{(n)}(x_0) = p^{(n)}(x_0) = n!\,a_n \quad \Rightarrow \quad a_n := \frac{f^{(n)}(x_0)}{n!}.$$

Die Potenzreihe

$$Tf(x) := \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}\,(x-x_0)^n \tag{13.2}$$

heißt *Taylor-Reihe* der Funktion f im *Entwicklungspunkt* x_0. Wenn man nur endlich viele Summanden berücksichtigt, dann erhält man das *Taylor-Polynom*

$$T_m f(x) := \sum_{n=0}^{m} \frac{f^{(n)}(x_0)}{n!}\,(x-x_0)^n$$

vom *Grad* m von f im Entwicklungspunkt x_0. Beachte: wir bezeichnen $T_m f(x)$ als Taylor-Polynom vom Grad m, auch wenn im Fall $f^{(m)}(x_0) = 0$ der Koeffizient vor x^m Null ist.

13.6 Beispiel:

- $f(x) = e^x$, Entwicklungspunkt $x_0 = 0$,

$$Tf(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad T_3 f(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}.$$

- $f(x) = e^x$, Entwicklungspunkt $x_0 = 1$,

$$Tf(x) = \sum_{n=0}^{\infty} \frac{e(x-1)^n}{n!}, \quad T_1 f(x) = e + e(x-1).$$

- $f(x) = \frac{1}{1-x}$, Entwicklungspunkt $x_0 = 0$,

$$Tf(x) = \sum_{n=0}^{\infty} x^n, \quad T_4 f(x) = 1 + x + x^2 + x^3 + x^4.$$

- Allgemein gilt: $T_1 f(x)$ ist die Tangente an den Graphen von f im Punkt x_0.

13.7 Regeln:

$$\begin{aligned} T(f+g) &= Tf + Tg \\ T(f \cdot g) &= Tf \cdot Tg \\ T(f') &= (Tf)' \\ T(\textstyle\int f) &= \textstyle\int (Tf) \end{aligned}$$

13.8 Restglied: Für die Abweichung zwischen Funktionswert $f(x)$ und dem Wert des Taylor-Polynoms m–ten Grades $T_m f(x)$ gilt: Es gibt eine Zahl ξ zwischen x_0 und x, sodass

$$f(x) = T_m f(x) + R_{m+1} f(x), \quad R_{m+1} f(x) = \frac{f^{(m+1)}(\xi)}{(m+1)!} (x - x_0)^{m+1}.$$

$R_{m+1} f(x)$ heißt *Restglied* in Lagrange-Form. Für festes x gilt $f(x) = Tf(x)$ genau dann, wenn $(R_m f(x))_{m \in \mathbb{N}}$ eine Nullfolge ist.

13.9 Taylor-Reihe einer elementaren Funktion: Alle sogenannten elementaren Funktionen $f : D_f \to \mathbb{R}$ (dazu gehören Polynome, rationale Funktionen, $\exp x$,$\sin x$,$\cos x$, $\sinh x$, $\cosh x$, $\ln x$, ... sowie deren Verkettungen oder Umkehrungen) haben folgende Eigenschaft: Ist $x_0 \in D_f$ ein innerer Punkt, d.h. es gibt ein $\epsilon > 0$, so dass $(x_0 - \epsilon, x_0 + \epsilon) \subset D_f$, dann hat die Taylor-Reihe von f mit Entwicklungspunkt x_0 einen positiven Konvergenzradius und

$$f(x) = Tf(x) \quad \text{für } x \in I_{Tf}.$$

Mit anderen Worten approximiert $T_m f(x)$ den Funktionswert $f(x)$, sofern $x \approx x_0$. *Faustregel:* Die Approximation ist umso besser, je näher x bei x_0 liegt und je größer m ist.

13.10 Beispiel:

- Sei $x \in \mathbb{R}_{>0}$ fest. Für $f(x) = \sin(x)$ und $x_0 = 0$ ist

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

und

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

mit einem $\xi \in (0, x)$, das von x und m abhängt. Je nach Wahl von m ist $f^{(m+1)}(\xi)$ einer der vier Werte $\pm\sin(\xi)$ oder $\pm\cos(\xi)$, in jedem Fall aber ist $|f^{(m+1)}(\xi)| \leq 1$. Also

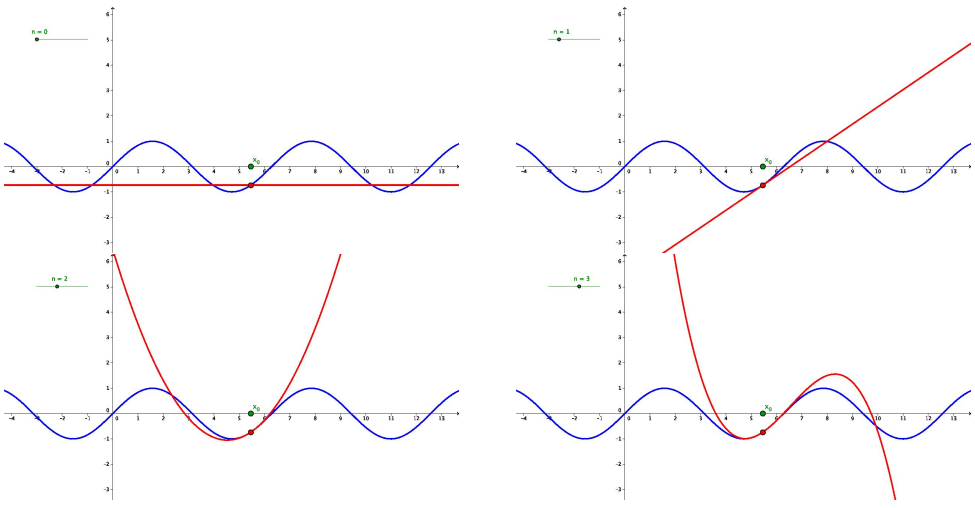
$$|R_{m+1} f(x)| \leq \frac{|x|^{m+1}}{(m+1)!}.$$

Da x fest ist, folgt $\lim_{m\to\infty} |R_{m+1} f(x)| = 0$. Also

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm,$$

d.h. der Wert, gegen die Reihe rechts für ein festes x konvergiert, ist $\sin(x)$. Wollen wir $\sin(x)$ auf 9 Nachkommastellen genau über die Taylor–Reihe berechnen, d.h. das Ziel ist $|\sin(x) - T_m \sin(x)| < 10^{-9}$, dann müssen wir m so groß wählen, dass $\frac{|x|^{m+1}}{(m+1)!} < 10^{-9}$. Je kleiner x ist, umso kleiner wird m sein.

- Bilder von $T_0 f$, $T_1 f$, $T_2 f$, $T_3 f$ für $f(x) = \sin(x)$ im Punkt x_0 (grün):



Auch das Bild zeigt: Je näher x an x_0 liegt und je größer m ist, desto besser approximiert $T_m f(x)$ den tatsächlichen Wert $\sin(x)$.

13.11 Bemerkung:

- Es gibt nicht-elementare Funktionen, bei denen $f(x)$ und $Tf(x)$ nur für $x = x_0$ übereinstimmt. Ein Beispiel ist

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

mit Entwicklungspunkt $x_0 = 0$. Man kann zeigen, dass f in $x_0 = 0$ beliebig oft differenzierbar ist mit $f^{(k)}(0) = 0$ für alle k. Also $Tf(x) = 0$ für alle x, während $f(x) = 0$ nur für $x = 0$. Hier wird $f(x)$ nicht durch $T_m f(x)$ approximiert.

- Mit der Substitution $h := x - x_0$ kann eine Taylor-Reihe (13.2) in die Form

$$Tf(x_0 + h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}\, h^n$$

gebracht werden. Analoges gilt für das Taylor-Polynom.

13.12 Beispiel: Nach 9.2 gilt

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n \quad \text{für } |x| < 1.$$

Aus dem Eindeutigkeitssatz folgt, dass die Potenzreihe auf der rechten Seite die Taylor-Reihe von $f(x) = \frac{1}{1-x}$ im Entwicklungspunkt $x_0 = 0$ sein muss. Dann ist

$$f^{(n)}(0) = n! \quad \text{für alle } n,$$

was direkt aus $f(x) = \frac{1}{1-x}$ nur mit Aufwand nachgerechnet werden kann. Reihe und Funktion haben dieselben Werte genau auf $I = (-1, 1)$. Für $x \in \mathbb{R}\backslash(-1, 1]$ kann $f(x)$ berechnet werden, die Reihe konvergiert aber nicht.
Multiplikation mit x liefert

$$\frac{x}{1-x} = \sum_{n=0}^{\infty} x^{n+1} \quad \text{für } |x| < 1.$$

Differenziation der Ausgangsgleichung auf beiden Seiten liefert

$$\frac{1}{(1-x)^2} = \sum_{n=1}^{\infty} n x^{n-1} \quad \text{für } |x| < 1.$$

Integration auf beiden Seiten liefert

$$-\ln(1-x) = \sum_{n=1}^{\infty} \frac{x^n}{n} \quad \text{für } |x| < 1.$$

Der Eindeutigkeitssatz zeigt wieder, dass ei allen drei Gleichungen jeweils auf der rechten Seite die Taylor-Reihe im Entwicklungspunkt $x_0 = 0$ der Funktion links steht. Bei der letzten Zeile hat man zuerst noch eine Konstante c, aber Einsetzen von $x = 0$ zeigt $c = 0$.
Die *Logarithmus-Reihe* $\sum_{n=1}^{\infty} \frac{x^n}{n}$ konvergiert für $x = -1$ nach dem Leibniz-Kriterium, nicht aber für $x = 1$. Hier ist also $I = [-1, 1)$. Da Funktion und Reihe auf I_0 übereinstimmen folgt

$$\ln 2 = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$$

wie in 9.6 behauptet.

13.13 Beispiel: Taylorreihen mit Entwicklungspunkt 0 und Konvergenzbereich I:

$$
\begin{aligned}
e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!}, && I = \mathbb{R} \\
\sin x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, && I = \mathbb{R} \\
\cos x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, && I = \mathbb{R} \\
\frac{1}{1-x} &= \sum_{n=0}^{\infty} x^n, && I = (-1,1) \\
\ln(1+x) &= \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}, && I = (-1,1] \\
\\
\frac{3x}{1-x} &= \sum_{n=0}^{\infty} 3x^{n+1}, && I = (-1,1) \\
\frac{1}{(1-x)^2} &= \sum_{n=1}^{\infty} n x^{n-1} && I = (-1,1) \\
\frac{1}{1+x^2} &= \sum_{n=0}^{\infty} (-1)^n x^{2n}, && I = (-1,1) \\
\arctan x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}, && I = [-1,1] \\
\frac{\sin x}{x} &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n+1)!}, && I = \mathbb{R} \\
\sinh x &= \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}, && I = \mathbb{R} \\
\cosh x &= \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}, && I = \mathbb{R} \\
(x-1)^4 &= 1 - 4x + 6x^2 - 4x^3 + x^4, && I = \mathbb{R}
\end{aligned}
$$

13.14 Landau-Symbole: Wir schreiben das *Landau Symbol* $O(m)$ um auszudrücken, dass nur Terme der Ordnung $\geq m$ folgen. Im Fall $x_0 = 0$ ist etwa

$$Tf(x) = \sum_{n=0}^{m-1} \frac{f^{(n)}(0)}{n!} x^n + O(x^m) = T_{m-1} f(x) + O(x^m),$$

d.h. $Tf(x) - T_{m-1} f(x)$ ist eine Potenzreihe beginnend mit x^m. Für beliebiges x_0 schreibt man entsprechend

$$Tf(x) = \sum_{n=0}^{m-1} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n + O\big((x - x_0)^m\big)$$

oder besser

$$Tf(x_0+h)=\sum_{n=0}^{m-1}\frac{f^{(n)}(x_0)}{n!}\,h^n+O(h^m).$$

13.15 Verkettete Funktionen: Ausgehend von der Formel

$$f(x)=T_{m-1}f(x)+O(x^m)$$

und den bekannten Taylor-Reihen elementarer Funktionen lassen sich Taylor-Polynome verketteter Funktionen ohne die Berechnung von Ableitungen bestimmen.

13.16 Beispiel:

- $f(x)=e^{-x^2}$, gesucht T_5f. Mit $u=-x^2$ erhält man

  $$f(x)=1+u+\frac{u^2}{2}+O(u^3)=1-x^2+\frac{x^4}{2}+O(x^6),$$

  also $T_5f(x)=1-x^2+x^4/2$.

- $f(x)=\tan x$, gesucht T_5f. Mit $\sin x=x-\frac{x^3}{6}+\frac{x^5}{120}+O(x^7)$ und $\cos x=1-\frac{x^2}{2}+\frac{x^4}{24}+O(x^6)$ erhält man zunächst

  $$\tan x=\frac{\sin x}{\cos x}=\frac{x-\frac{x^3}{6}+\frac{x^5}{120}+O(x^7)}{1-\frac{x^2}{2}+\frac{x^4}{24}+O(x^6)}.$$

  Die Substitution $u=x^2/2-x^4/24+O(x^6)$ liefert mit der geometrischen Reihe bzw. der Taylorreihe zu $\frac{1}{1-u}$

  $$\begin{aligned}\tan x &= (x-\tfrac{x^3}{6}+\tfrac{x^5}{120}+O(x^7))\tfrac{1}{1-u}\\ &= \left(x-\tfrac{x^3}{6}+\tfrac{x^5}{120}+O(x^7)\right)\cdot(1+u+u^2+O(u^3))\,.\end{aligned}$$

  Rücksubstitution ergibt

  $$\tan x=\left(x-\frac{x^3}{6}+\frac{x^5}{120}+O(x^7)\right)\cdot\left(1+\frac{x^2}{2}-\frac{x^4}{24}+\frac{x^4}{4}+O(x^6)\right)=x+\frac{x^3}{3}+\frac{2x^5}{15}+O(x^7).$$

  Also ist $T_5f(x)=x+\frac{x^3}{3}+\frac{2x^5}{15}$.

13.17 Bestimmung von Grenzwerten: Grenzwerte vom Typ

$$\lim_{x\to x_0}\frac{f(x)}{g(x)},\quad f(x_0)=g(x_0)=0,$$

lassen sich ohne die Regel von l'Hospital bestimmen, indem man die Taylor-Entwicklung von f und g bis zum ersten von Null verschiedenen Term explizit bestimmt.

13.18 Beispiel:

- $f(x) = \sin(x^2) = x^2 + O(x^3),\ g(x) = 1 - \cos x = x^2/2 + O(x^3)$, also

$$\lim_{x\to 0} \frac{\sin(x^2)}{1-\cos x} = \lim_{x\to 0} \frac{x^2 + O(x^3)}{x^2/2 + O(x^3)} = \lim_{x\to 0} \frac{1 + O(x)}{1/2 + O(x)} = 2.$$

- $f(x) = x^3(e^{2x} - 1) = 2x^4 + O(x^5),\ g(x) = \ln(1 + x^3) = x^3 + O(x^4)$, also

$$\lim_{x\to 0} \frac{x^3(e^{2x}-1)}{\ln(1+x^3)} = \lim_{x\to 0} \frac{2x^4 + O(x^5)}{x^3 + O(x^4)} = \lim_{x\to 0} \frac{2x + O(x^2)}{1 + O(x)} = 0.$$

- $f(x) = x^2 - \sin x^2 = x^6/6 + O(x^7),\ g(x) = (1 - \cos x)^3 = x^6/8 + O(x^7)$, also

$$\lim_{x\to 0} \frac{x^2 - \sin x^2}{(1-\cos x)^3} = \lim_{x\to 0} \frac{x^6/6 + O(x^7)}{x^6/8 + O(x^7)} = \lim_{x\to 0} \frac{1/6 + O(x)}{1/8 + O(x)} = \frac{4}{3}.$$

13.19 Minima/Maxima: Die Funktion $f : I = (a, b) \to \mathbb{R}$ hat in $x_0 \in I$ ein *lokales Minimum (bzw. Maximum)*, wenn es ein $\epsilon > 0$ gibt, s.d. $f(x) \geq f(x_0)$ (bzw. $f(x) \leq f(x_0)$) für alle $x \in I$ mit $|x - x_0| < \epsilon$.
Sei f nun differenzierbar, $f'(x_0) = \cdots = f^{(k-1)}(x_0) = 0$ und $f^{(k)}(x_0) \neq 0$, dann gilt

$$Tf(x_0 + h) = f(x_0) + \frac{f^{(k)}(x_0)}{k!}\, h^k + O(h^{k+1}).$$

Für x nahe bei x_0 ist h klein, daher sieht der Graph von f dort aus wie der Graph von $f(x_0) + \frac{f^{(k)}(x_0)}{k!}\, h^k$ bei $h = 0$. Daraus kann man folgendes ablesen:

- Wenn k gerade ist und $f^{(k)}(x_0) > 0$, dann hat f im Punkt x_0 ein lokales Minimum.
- Wenn k gerade ist und $f^{(k)}(x_0) < 0$, dann hat f im Punkt x_0 ein lokales Maximum.

13.20 Beispiel: Die Funktion

$$f(x) := \cos(x) + \cosh(x) = 2 + x^4/12 + O(x^5)$$

hat an der Stelle $x_0 = 0$ ein lokales Minimum.

14 Fourier-Reihen

14.1 Periodische Funktion: Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt *L-periodisch*, wenn

$$f(x + L) = f(x), \quad x \in \mathbb{R}$$

für eine Zahl $L > 0$. Eine L-periodische Funktion ist auch kL-periodisch für alle $k \in \mathbb{N}$. Die *Periode* einer Funktion ist definiert als die kleinste Zahl $L > 0$, für die die Funktion L-periodisch ist. Eine L-periodische Funktion ist durch ihre Werte auf einem beliebigen Intervall der Länge L vollständig bestimmt.

14.2 Trigonometrische Reihen: Eine Reihe der Form

$$\begin{aligned} t(x) &= \frac{a_0}{2} + a_1 \cos(x) + b_1 \sin(x) + a_2 \cos(2x) + b_2 \sin(2x) + a_3 \cos(3x) + b_3 \sin(3x) + \cdots \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx) \end{aligned} \tag{14.1}$$

heißt *trigonometrische Reihe* mit Koeffizienten a_n, b_n. Die Partialsumme bis m

$$t_m(x) = \frac{a_0}{2} + \sum_{n=1}^{m} a_n \cos(nx) + \sum_{n=1}^{m} b_n \sin(nx)$$

heißt *trigonometrisches Polynom* vom Grad m.

14.3 Eigenschaften:

- t konvergiert für alle $x \in \mathbb{R}$ absolut, wenn es Konstanten $\alpha, \beta > 1$ gibt, sodass

  $$\lim_{n\to\infty} n^{\alpha} a_n = \lim_{n\to\infty} n^{\beta} b_n = 0.$$

  Die Funktion t ist dann stetig.

- t ist k-mal stetig differenzierbar, wenn es Konstanten $\alpha, \beta > k + 1$ gibt, sodass

  $$\lim_{n\to\infty} n^{\alpha} a_n = \lim_{n\to\infty} n^{\beta} b_n = 0.$$

  Die Ableitungen sind wieder trigonometrische Reihen, deren Koeffizienten sich durch gliedweises Differenzieren bestimmen lassen.

- Das Integral $\int t_m(x)\,dx$ ist genau dann ein trigonometrischen Polynom, wenn $a_0 = 0$. Wenn bei der unendlichen Reihe t summandenweise integriert werden darf, dann ist $\int t(x)\,dx$ genau dann eine trigonometrische Reihe, wenn $a_0 = 0$.

- Die trigonometrischen Polynome sind 2π-periodisch, d.h., $t_m(x + 2\pi) = t_m(x)$ für $x \in \mathbb{R}$. Das Gleiche gilt für trigonometrische Reihen, d.h.: Wenn die Reihe $t(x)$ für einen Wert x kovergiert, dann konvergiert sie auch für $x+2\pi$ und $t(x) = t(x+2\pi)$.

- t bzw. $t_m(x)$ ist genau dann eine gerade Funktion, wenn alle b_n Null sind.

- t bzw. $t_m(x)$ ist genau dann eine ungerade Funktion, wenn alle a_n Null sind.

14.4 Orthogonalitätsrelationen: Für beliebige $k, l \in \mathbb{N}_0$ gilt

$$\int_0^{2\pi} \cos(kx)\sin(lx)\,dx = 0$$

$$\int_0^{2\pi} \cos(kx)\cos(lx)\,dx = \begin{cases} 2\pi & \text{falls } k = l = 0 \\ \pi & \text{falls } k = l > 0 \\ 0 & \text{sonst} \end{cases}$$

$$\int_0^{2\pi} \sin(kx)\sin(lx)\,dx = \begin{cases} \pi & \text{falls } k = l > 0 \\ 0 & \text{sonst} \end{cases}$$

Durch Anwendung der Orthogonalitätsrelationen sieht man, dass zwischen einem trigonometrischen Polynom und seinen Koeffizienten der folgende Zusammenhang besteht:

$$a_k = \frac{1}{\pi}\int_0^{2\pi} t_m(x)\cos(kx)\,dx, \quad b_k = \frac{1}{\pi}\int_0^{2\pi} t_m(x)\sin(kx)\,dx. \tag{14.2}$$

Insbesondere ist

$$a_0 = \frac{1}{\pi}\int_0^{2\pi} t_m(x)\,dx.$$

Das Gleiche gilt für $t(x)$, sobald klar ist, dass die Reihe summandenweise integriert werden kann.

14.5 Fourier-Reihe periodischer Funktionen: Sei $f : \mathbb{R} \to \mathbb{R}$ eine 2π-periodische Funktion, die integrierbar ist, also eine Stammfunktion besitzt. Wenn t eine trigonometrische Reihe gemäß (14.1) ist, die mit f übereinstimmt und wenn summandenweise integriert werden darf, dann muss wegen (14.2) für die Koeffizienten gelten

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(kx)\,dx, \quad b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(kx)\,dx. \tag{14.3}$$

Die trigonometrische Reihe

$$Ff(x) := \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n\cos(nx) + \sum_{n=1}^{\infty} b_n\sin(nx)$$

mit Koeffizienten gemäß (14.3) heißt *Fourier-Reihe* der Funktion f. Wenn man nur endlich viele Summanden berücksichtigt, so erhält man das trigonometrische *Fourier-Polynom* zu f

$$F_m f(x) := \frac{a_0}{2} + \sum_{n=1}^{m} a_n\cos(nx) + \sum_{n=1}^{m} b_n\sin(nx)$$

vom Grad m.

14.6 Bemerkungen:

- $a_0/2 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx$ ist der *Mittelwert* der Funktion f.
- Wenn f gerade ist, bzw. $f(x) = f(-x)$ für alle $x \in [0, 2\pi]$ mit höchstens endlich vielen Ausnahmen, dann ist Ff eine reine *Cosinus-Reihe*, d.h., $b_n = 0, n \in \mathbb{N}$.

- Wenn f ungerade ist, bzw. $f(x) = -f(-x)$ für alle $x \in [0, 2\pi]$ mit höchstens endlich vielen Ausnahmen, dann ist Ff eine reine *Sinus-Reihe*, d.h., $a_n = 0, n \in \mathbb{N}_0$.

- Das Integral $\int_0^{2\pi}$ in der Definition der Koeffizienten kann durch ein beliebiges anderes Integral über ein Intervall der Länge 2π ersetzt werden, z.B. durch $\int_{-\pi}^{\pi}$.

14.7 Fourier Reihe stückweise glatter Funktionen: Es sei $f : \mathbb{R} \to \mathbb{R}$ eine 2π-periodische Funktion, die *stückweise glatt* ist. D.h. es gebe endlich viele x–Werte

$$0 = x_0 < x_1 < x_2 < \cdots < x_n = 2\pi$$

so dass gilt:

- f ist auf jedem Teilintervall (x_{i-1}, x_i) stetig differenzierbar.

- In den Punkten $x_1, \ldots, x_n$ gibt es den linksseitigen Funktionsswert und die linksseitige Tangentensteigung, d.h. es existieren die Grenzwerte

$$f(x_i-) = \lim_{x \uparrow x_i} f(x) \quad \text{und} \quad \lim_{x \uparrow x_i} \frac{f(x) - f(x_i-)}{x - x_i}.$$

- In den Punkten $x_0, \ldots, x_{n-1}$ gibt es den rechtsseitigen Funktionsswert und die rechtsseitige Tangentensteigung, d.h. es existieren die Grenzwerte

$$f(x_i+) = \lim_{x \downarrow x_i} f(x) \quad \text{und} \quad \lim_{x \downarrow x_i} \frac{f(x) - f(x_i+)}{x - x_i}.$$

Dann kovergiert $Ff(x)$ für alle $x \in [0, 2\pi]$ und es gilt

$$Ff(x) = f(x), \quad x \neq x_i$$

und

$$Ff(x_i) = \frac{1}{2}(f(x_i-) + f(x_i+)),$$

d.h. $Ff(x_i)$ ist der Mittelwert aus linksseitigem und rechtsseitigem Grenzwert von f im Punkt x_i.

14.8 Beispiel: $f(x) = \max(0, x)$ für $x \in [-\pi, \pi)$.

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx)\, dx = \begin{cases} \pi/2 & \text{für } n = 0 \\ -2/(\pi n^2) & \text{für } n = 1, 3, 5, \ldots \\ 0 & \text{für } n = 2, 4, 6, \ldots \end{cases}$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx)\, dx = (-1)^{n+1}/n$$

Damit ergibt sich die Fourier-Reihe

$$Ff(x) = \frac{\pi}{4} - \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} \cos((2n-1)x) - \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \sin(nx)$$

und das Fourier-Polynom der Ordnung 4

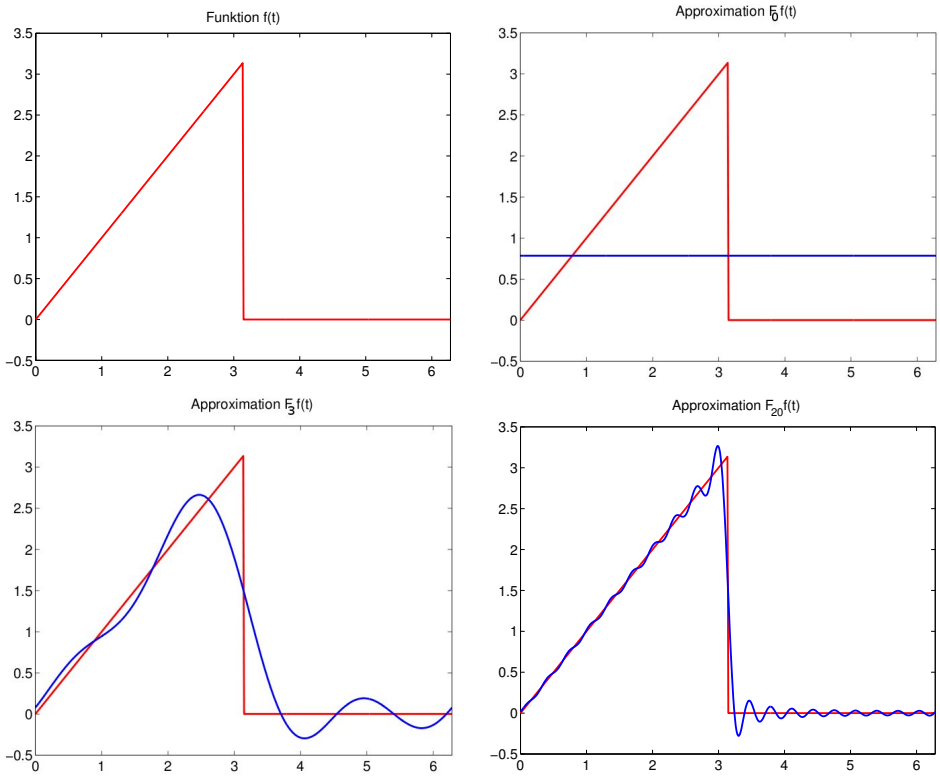
$$F_4 f(x) = \frac{\pi}{4} - \frac{2}{\pi}\cos(x) - \frac{2}{9\pi}\cos(3x) + \sin(x) - \frac{1}{2}\sin(2x) + \frac{1}{3}\sin(3x).$$

Setzt man speziell $x = 0$, so erhält man

$$Ff(0) = f(0) = 0 = \frac{\pi}{4} - \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2}$$

und damit

$$\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2} = 1 + \frac{1}{9} + \frac{1}{25} + \frac{1}{49}\cdots = \frac{\pi^2}{8}.$$



14.9 Beliebige Periode: Hat die gegebene Funktion f die Periode L, dann kann man diese durch die Transformation

$$g(x) := f(x/\omega), \quad \omega := \frac{2\pi}{L}$$

auf eine 2π-periodische Funktion g transformieren. Durch Rücktransformation der Fourier-Reihe von g erhält man für f die verallgemeinerte Fourier-Reihe

$$Ff(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\omega x) + \sum_{n=1}^{\infty} b_n \sin(n\omega x)$$

mit Koeffizienten

$$a_n = \frac{2}{L}\int_0^L f(x)\cos(n\omega x)\,dx$$

$$b_n = \frac{2}{L}\int_0^L f(x)\sin(n\omega x)\,dx.$$

Wieder kann statt von 0 bis L über jedes Intervall der Länge L integriert werden.

14.10 Fourier-Reihe in komplexer Form: Mit Hilfe der Eulerschen Formel

$$e^{inx} = \cos nx + i \sin nx$$

kann man die Fourier-Reihe (14.1) auch in der Form

$$Ff(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$$

schreiben. Dabei sind die Koeffizienten gegeben durch

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx}\, dx = \begin{cases} \frac{1}{2}\, a_0 & \text{für } n = 0 \\ \frac{1}{2}\,(a_n - i b_n) & \text{für } n > 0 \\ \frac{1}{2}\,(a_{-n} + i b_{-n}) & \text{für } n < 0. \end{cases}$$

Umgekehrt gilt

$$a_n = c_n + c_{-n}, \quad b_n = i c_n - i c_{-n}.$$

15 Funktionen mehrerer Veränderlicher

15.1 Erinnerung: Der Raum $\mathbb{R}^m$, Skalarprodukt, Norm und Folgen: Wie in Abschnitt 2 ist im folgenden

$$\mathbb{R}^m := \{\vec{x} = [x_1, x_2, \ldots, x_m]^T : x_1, \ldots, x_m \in \mathbb{R}\}$$

und

$$\begin{aligned} \vec{x} \pm \vec{y} &:= [x_1 \pm y_1,\ x_2 \pm y_2, \ldots, x_m \pm y_m]^T \\ c\vec{x} &:= [cx_1, cx_2, \ldots, cx_m]^T, \quad c \in \mathbb{R} \\ \langle \vec{x}, \vec{y} \rangle &:= \vec{x} \cdot \vec{y} := x_1y_1 + x_2y_2 + \cdots + x_my_m \\ \|\vec{x}\| &:= \sqrt{\langle \vec{x}, \vec{x} \rangle}. \end{aligned}$$

Dabei gilt

$$\begin{aligned} \|c\vec{x}\| &= |c| \cdot \|\vec{x}\|, \quad c \in \mathbb{R} \\ \|\vec{x} + \vec{y}\| &\le \|\vec{x}\| + \|\vec{y}\| \\ \|\vec{x} - \vec{y}\| &\ge |\, \|\vec{x}\| - \|\vec{y}\|\, | \\ \langle \vec{x}, \vec{y} \rangle &\le \|\vec{x}\| \cdot \|\vec{y}\| \end{aligned}$$

Eine Folge $(\vec{x}_k)_{k\in\mathbb{N}}$ von Vektoren mit Koordinaten $\vec{x}_k = [x_{1,k}, \ldots, x_{m,k}]^T$ in $\mathbb{R}^m$ heißt *konvergent*, wenn die Zahlenfolge jeder Koordinate konvergiert, also die Zahlenfolge $(x_{i,k})_{k\in\mathbb{N}}$ für $i = 1, \ldots, m$. Der *Grenzwert* $\vec{x}_0$ ist dann gegeben als Vektor der Grenzwerte der einzelnen Koordinaten,

$$\vec{x}_0 := \begin{bmatrix} \lim\limits_{k\to\infty} x_{1,k} \\ \vdots \\ \lim\limits_{k\to\infty} x_{m,k} \end{bmatrix},$$

und man schreibt

$$\lim_{k\to\infty} \vec{x}_k = \vec{x}_0.$$

Ein Folge, die nicht konvergent ist, heißt *divergent*.

- Die Folge $\vec{x}_k$ konvergiert genau dann gegen $\vec{x}_0$, wenn $\|\vec{x}_k - \vec{x}_0\|$ eine Nullfolge ist.
- Seien $\vec{x}_k$ und $\vec{y}_k$ konvergente Folgen mit Grenzwert $\vec{x}_0$ bzw. $\vec{y}_0$. Dann gilt

$$\begin{aligned} \lim_{k\to\infty} (\vec{x}_k + \vec{y}_k) &= \vec{x}_0 + \vec{y}_0, & \lim_{k\to\infty} c\vec{x}_k &= c\vec{x}_0 \\ \lim_{k\to\infty} \langle \vec{x}_k, \vec{y}_k \rangle &= \langle \vec{x}_0, \vec{y}_0 \rangle, & \lim_{k\to\infty} \|\vec{x}_k\| &= \|\vec{x}_0\|. \end{aligned}$$

15.2 Beispiel:

- Die Folge $(\vec{x}_k)_{k\in\mathbb{N}}$ mit $\vec{x}_k = [1/k, \sqrt[k]{k}]^T$ ist konvergent zum Grenzwert

$$\vec{x}_0 = \lim_{k\to\infty} \vec{x}_k = \begin{bmatrix} \lim\limits_{k\to\infty} 1/k \\ \lim\limits_{k\to\infty} \sqrt[k]{k} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

- Die Folge $(\vec{x}_k)_{k\in\mathbb{N}}$ mit $\vec{x}_k = [1/k, 3, k]^T$ ist divergent, da die dritte Koordinate divergiert.

15.3 Topologische Begriffe: Es sei $M \subset \mathbb{R}^m$.

- Die Menge
$$U_\varepsilon(\vec{x}_0) := \{\vec{x} \in \mathbb{R}^m : \|\vec{x} - \vec{x}_0\| < \varepsilon\}$$
heißt *ε-Kugel um $\vec{x}_0$*. Hier ist $\varepsilon > 0$ ein fester Radius.
- Ein Punkt $\vec{x}_0$ heißt *innerer Punkt* von M, wenn es eine ε-Kugel $U_\varepsilon(\vec{x}_0)$ gibt, die vollständig in M enthalten ist.
- Die Menge M heißt *offen*, wenn alle ihre Punkte innere Punkte sind.
- Ein Punkt $\vec{x}_0$ heißt *Häufungspunkt* von M, wenn es eine Folge $(\vec{x}_k)_{k\in\mathbb{N}}$ von Punkten in M gibt, die gegen $\vec{x}_0$ konvergiert, wobei $\vec{x}_k \neq \vec{x}_0$ für alle k.
- Die Menge M heißt *abgeschlossen*, wenn sie alle ihre Häufungspunkte enthält.
- Die Menge M heißt *beschränkt*, wenn es eine Konstante R gibt, sodass für alle $\vec{x}$ in M gilt $\|\vec{x}\| < R$.
- Die Menge M heißt *kompakt*, wenn sie beschränkt und abgeschlossen ist.

15.4 Beispiel:

- $M = U_\varepsilon(\vec{x}_0)$ ist offen und beschränkt.
- $M = K_\varepsilon(\vec{x}_0) := \{\vec{x} \in \mathbb{R}^m : \|\vec{x} - \vec{x}_0\| \leq \varepsilon\}$ ist kompakt.
- $M = \mathbb{R}^m$ ist offen und abgeschlossen, aber nicht beschränkt.
- $M = K_\varepsilon(\vec{x}_0)\backslash\{\vec{x}_0\}$ ist weder offen noch abgeschlossen, aber beschränkt.

15.5 Bemerkungen:

- Eine Menge $M \subset \mathbb{R}^m$ ist genau dann abgeschlossen, wenn ihr *Komplement*
$$M^C := \{\vec{x} \in \mathbb{R}^m : \vec{x} \notin M\}$$
offen ist.
- Seien M_1 und M_2 offen. Dann sind auch $M_1 \cup M_2$ und $M_1 \cap M_2$ offen.
- Seien M_1 und M_2 abgeschlossen. Dann sind auch $M_1 \cup M_2$ und $M_1 \cap M_2$ abgeschlossen.

15.6 Funktionen mehrerer Veränderlicher: Sei $M \subset \mathbb{R}^m$. Eine Vorschrift $f : M \to \mathbb{R}$, die jedem Element $\vec{x} \in M$ eine reelle Zahl $y = f(\vec{x}) = f(x_1, \ldots, x_m)$ zuordnet, heißt *skalare (oder reelle) Funktion von m Veränderlichen* mit Definitionsgebiet M. Wir verwenden synonym die Schreibweisen
$$f(\vec{x}) = f(x_1, \ldots, x_m).$$

15.7 Spezielle Funktionen:

- Sei $c \in \mathbb{R}$, dann ist $f(\vec{x}) = c$ eine *konstante Funktion.*
- Sei $\vec{a} \in \mathbb{R}^m$, dann ist $f(\vec{x}) = \langle \vec{a}, \vec{x} \rangle = \vec{a}^T \vec{x}$ eine *lineare Funktion.*
- Sei A eine $(m \times m)$-Matrix, dann ist $f(\vec{x}) = \langle \vec{x}, A\vec{x} \rangle = \vec{x}^T A \vec{x}$ eine *quadratische Funktion.*
- $f(\vec{x}) = \|\vec{x} - P\|$ ist die *Abstands-Funktion* zum Punkt P.

Dabei ist jeweils das Definitionsgebiet $M = \mathbb{R}^m$ möglich.

15.8 Stetigkeit: Sei $f : M \to \mathbb{R}$ eine Funktion und $\vec{x}_0 \in M$ ein Punkt im Definitionsgebiet. Die Funktion f heißt *stetig im Punkt* $\vec{x}_0$, wenn für jede Folge $\vec{x}_k$ in M, die gegen $\vec{x}_0$ konvergiert, gilt

$$\lim_{k\to\infty} f(\vec{x}_k) = f(\lim_{k\to\infty} \vec{x}_k) = f(\vec{x}_0).$$

Anderenfalls heißt f *unstetig im Punkt* $\vec{x}_0$. Die Funktion f heißt *stetig*, wenn sie in allen Punkten von M stetig ist.

15.9 Regeln: Wie im Fall $m = 1$ (vgl. 10.9) hat man

- Alle Verknüpfungen der elementaren Funktionen mit mehreren Variablen sind auf dem Definitionsgebiet stetig.
- Wenn $f : M \to \mathbb{R}$ und $g : M \to \mathbb{R}$ stetig sind, dann sind auch $f \pm g$, fg sowie $\max(f, g)$ und $\min(f, g)$ stetig. Der Quotient f/g ist stetig in $\vec{x}_0$, falls $g(\vec{x}_0) \neq 0$.
- Wenn M kompakt ist und $f : M \to \mathbb{R}$ stetig, dann nimmt f auf M sein Maximum und sein Minimum an, d.h., es gibt Punkte $\vec{x}_{\min}$ und $\vec{x}_{\max}$ in M mit

$$f(\vec{x}_{\min}) \leq f(\vec{x}) \leq f(\vec{x}_{\max}) \quad \text{für alle } \vec{x} \in M.$$

15.10 Beispiel:

- Konstante, lineare und quadratische Funktionen sowie Abstands-Funktionen sind stetig auf $\mathbb{R}^m$.
- $f(x, y) = x^2 \sin(x - y^3) + xy$ ist stetig auf $\mathbb{R}^2$.
- Die Funktion

$$f_1(x, y) = \begin{cases} \frac{x^2 y}{x^2+y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

  ist stetig auf $\mathbb{R}^2$.
- Die Funktion

$$f_2(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{für } (x, y) \neq (0, 0) \\ 0 & \text{für } (x, y) = (0, 0) \end{cases}$$

  ist stetig auf $\mathbb{R}^2 \backslash \{\vec{0}\}$ und unstetig in $\vec{0}$.

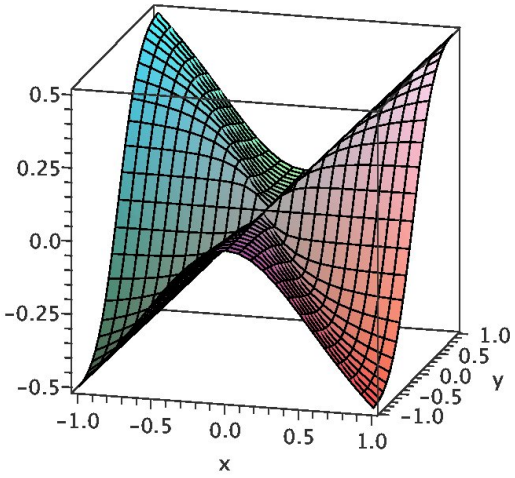


Abbildung 21: Die Funktion f_1 ist stetig im Ursprung. Egal, wie man sich mit Punkten $[x_n, y_n]^T$ dem Ursprung nähert, die Funktionswerte $f_1(x_n, y_n)$ konvergieren gegen $0 = f_1(0, 0)$.

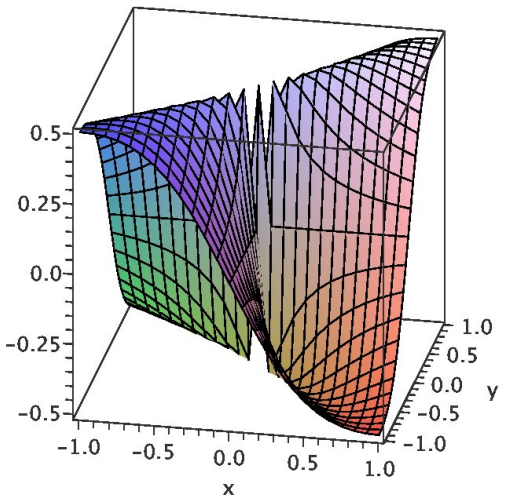


Abbildung 22: Die Funktion f_2 ist unstetig im Ursprung. Egal welchen Wert $c \in [-0,5, 0,5]$ man wählt, man findet eine Folge $([x_n, y_n]^T)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty}[x_n, y_n]^T = \vec{0}$ und $\lim_{n\to\infty} f_2(x_n, y_n) = c$.

15.11 Vektorfelder: Seien $f_1, \ldots, f_n$ skalare Funktionen mit einem gemeinsamen Definitionsgebiet $M \subset \mathbb{R}^m$. Die daraus gebildete Funktion $F : M \to \mathbb{R}^n$ mit

$$F(\vec{x}) = \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_n(\vec{x}) \end{bmatrix} = \begin{bmatrix} f_1(x_1, \ldots, x_m) \\ \vdots \\ f_n(x_1, \ldots, x_m) \end{bmatrix}$$

heißt *vektorwertige Funktion* von m Veränderlichen oder auch kurz *Vektorfeld*. Die Funktionen $f_1, \ldots, f_n$ heißen *Koordinatenfunktionen*. F ist genau dann stetig, wenn alle Koordinatenfunktionen stetig sind.

15.12 Spezielle Funktionen:

- Sei $\vec{c} \in \mathbb{R}^n$, dann ist $F(\vec{x}) = \vec{c}$ eine *konstante Funktion*.
- Sei A eine $(n \times m)$-Matrix, dann ist $F(\vec{x}) = A\vec{x}$ eine *lineare Funktion*.
- $\mathrm{Id}(\vec{x}) := \vec{x}$ ist die *identische Abbildung* (oder kurz die *Identität*). Dies ist die lineare Abbildung mit $A = E$. Wenn das Definitionsgebiet auf die Menge $M \subset \mathbb{R}^m$ eingeschränkt ist, dann schreibt man auch Id_M.

15.13 Verkettung: Seien $G : \mathbb{R}^k \to \mathbb{R}^m$ und $F : \mathbb{R}^m \to \mathbb{R}^n$ zwei Vektorfelder. Dann ist die *verkettete Funktion* $V := F \circ G$ definiert durch

$$V : \mathbb{R}^k \to \mathbb{R}^n, \quad V(\vec{x}) := F(G(\vec{x})).$$

Wenn F und G stetig sind, dann ist auch $V = F \circ G$ stetig.

15.14 Umkehrfunktion: Sei $F : M \to \mathbb{R}^n$ ein Vektorfeld mit Bildbereich

$$B := \{F(\vec{x}) : \vec{x} \in M\} \subset \mathbb{R}^n.$$

Wenn es eine Funktion $G : B \to M$ gibt mit

$$G \circ F = \mathrm{Id}_M, \quad F \circ G = \mathrm{Id}_B,$$

dann heißt G *Umkehrfunktion* von F und man schreibt $G = F^{-1}$.

15.15 Beispiel: Sei $M = \mathbb{R} \times (-\pi/2, \pi/2)$ und

$$F(p, q) = \begin{bmatrix} e^p \cos q \\ e^p \sin q \end{bmatrix}, \quad G(x, y, z) = \begin{bmatrix} x + y \\ xz \end{bmatrix},$$

dann ist

$$F^{-1}(r, s) = \begin{bmatrix} \frac{1}{2} \ln(r^2 + s^2) \\ \arctan(s/r) \end{bmatrix}, \quad V(x, y, z) = \begin{bmatrix} e^{x+y} \cos xz \\ e^{x+y} \sin xz \end{bmatrix}.$$

15.16 Graph und Bild: Es geht nun um bildliche Veranschaulichungen von Funktionen. Dazu sei $F : M \to \mathbb{R}^n$ gegeben, $M \subset \mathbb{R}^m$. Die f_i seien aus elementaren Funktionen zusammen gesetzt.
Im Fall $f : M \to \mathbb{R}$ mit $M \subset \mathbb{R}$ zeichnen wir seit der Schulzeit den *Graphen* im $\mathbb{R}^2$. Wir markieren zu jedem x–Wert in M seinen $y = f(x)$–Wert, zeichnen also die Punktmenge

$$G_f = \{(x, f(x)) \mid x \in M\} \subset \mathbb{R} \times \mathbb{R} = \mathbb{R}^2.$$

In direkter Verallgemeinerung ist *der Graph von F* die Punktmenge

$$G_F = \{(\vec{x}, F(\vec{x})) \mid \vec{x} \in M\} \subset \mathbb{R}^m \times \mathbb{R}^n = \mathbb{R}^{m+n}.$$

Da wir nur Bilder bis Dimension 3 zeichnen können, kann G_F nur für $m+n \leq 3$ bildlich dargestellt werden, also in folgenden Fällen:

- $F : M \to \mathbb{R}, M \subset \mathbb{R}$ (kennen wir aus der Schule).
- $F : M \to \mathbb{R}, M \subset \mathbb{R}^2$. Wir erhalten G_F wieder, indem wir über jedem $[x, y]^T \in M$ den Punkt $[x, y, F(x, y)]^T$ markieren. Speziell im Fall

  $$F : M = [a, b] \times [c, d] \to \mathbb{R}$$

  ist G_F ein "gekrümmtes Flächenstück" über dem Rechteck $[a, b] \times [c, d]$ in der (x, y)–Ebene (Bilder 24). Man kann G_F wie in den Bildern durch ein gekrümmtes Gitternetz darstellen. Einen Eindruck vermitteln auch die *Höhenlinien* in der (x, y)–Ebene: Für festes $c \in \mathbb{R}$ sind das die Punkte

  $$\{\vec{x} \in M \mid F(\vec{x}) = c\}.$$

- $F : M \to \mathbb{R}^2, M \subset \mathbb{R}$. Wir erhalten G_F, indem wir über jedem $x \in M$ den Punkt $[x, f_1(x), f_2(x)]^T$ markieren. Speziell im Fall

  $$F : [a, b] \to \mathbb{R}^2$$

  ist G_F eine "sich um $[a, b]$ windende Kurve". Man zeichnet hier allerdings zumeist B_F statt G_F.

Statt den Graphen zu betrachten kann man nämlich auch die Bildmenge von F, also

$$B_F = \{F(\vec{x}) \mid \vec{x} \in M\} \subset \mathbb{R}^n$$

betrachten. Geometrisch erhalten wir B_F aus G_F durch Projektion auf die Funktionswerte. In den ersten beiden Punkten oben ist $B_F \subset \mathbb{R}$ die Menge der y– bzw. z–Werte von F. Im dritten Punkt ist B_F eine *Kurve* in $\mathbb{R}^2$ (Linkes Bild 23).
Manchmal kann B_F bildlich veranschaulicht werden, G_F aber nicht. Z.B. kann im Fall $F : M \to \mathbb{R}^3$, $M \subset \mathbb{R}$ oder $M \subset \mathbb{R}^2$, der Graph G_F nicht mehr gezeichnet werden, wohl aber B_F. Die interessanten Fälle bekommen Namen:

- $F : [a, b] \to \mathbb{R}^n$ ist eine *Kurve* (auch Kurvenstück oder Weg). Man bezeichnet manchmal auch das Bild B_F als Kurve, parametrisierte Kurve oder Spur der Kurve (Bilder 23).
- $F : [a, b] \times [c, d] \to \mathbb{R}^3$ ist eine *Fläche*. Auch das Bild B_F heißt Fläche oder Flächenstück (Bild 25).

Wir nennen F *(globale) Parametrisierung* von B_F. Wir werden die Begriffe Kurve, Fläche später allgemeiner definieren, benötigen dazu aber Begriffe wie Stetigkeit, Differenzierbarkeit.

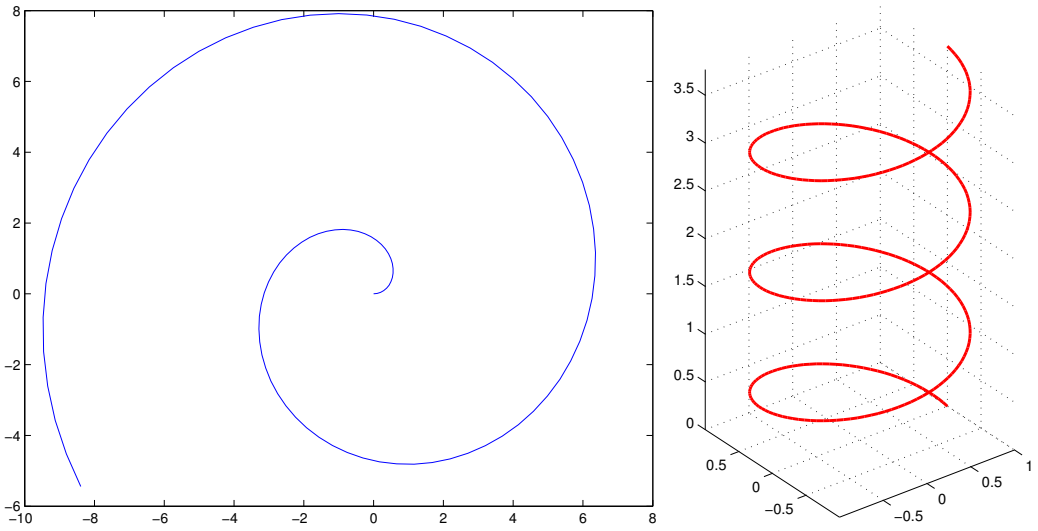


Abbildung 23: Zwei Bilder B_F. Links für $F : [0, 10] \to \mathbb{R}^2$, $F(t) = [t \cos t, t \sin t]^T$, rechts für $F : [0, 6\pi] \to \mathbb{R}^3$, $F(t) = [\cos t, \sin t, t/5]^T$

Graphen kann man parametrisieren: Der Graph der Funktion $F : [a, b] \times [c, d] \to \mathbb{R}$ wird parametrisiert durch $H = [a, b] \times [c, d] \to \mathbb{R}^3, [x, y]^T \mapsto [x, y, f(x, y)]^T$, denn es gilt $G_F = B_H$. Es ist andererseits

$$F : [0, 2\pi] \to \mathbb{R}^2, \quad F(t) = (\cos(t), \sin(t))^t$$

eine Parametrisierung des Kreises, der Kreis selbst ist aber nicht Graph einer Funktion $f : M \to \mathbb{R}$, $M \subset \mathbb{R}$.

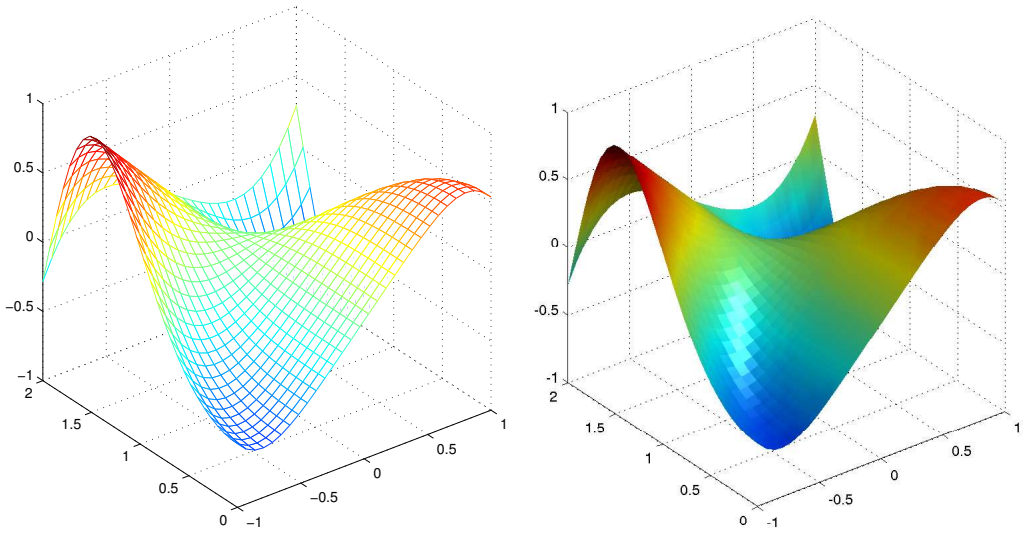


Abbildung 24: Graph G_F für $F : [-1, 1] \times [0, 2] \to \mathbb{R}, F(x, y) = x \cos(x^2 + y^2)$.

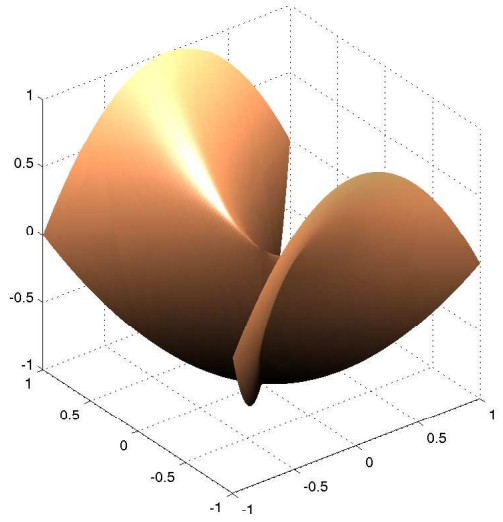


Abbildung 25: Bild B_F für $F : [-1, 1] \times [-1, 1] \to \mathbb{R}^3, F(u, v) = [uv, u, u^2 - v^2]^T$.

16 Differenziation in mehreren Veränderlichen

Die Einführung der Differentialrechnung in mehreren Variablen birgt ein **Problem:**
Für $M \subset \mathbb{R}^m$ sei $f\colon M \to \mathbb{R}$ eine Funktion, also $f(\vec{x}) = f(x_1, x_2, \ldots, x_m)$. Diese Funktion f soll abgeleitet werden.
Folgendes funktioniert nicht:

$$f'(\vec{x}_0) = \lim_{\vec{x}\to\vec{x}_0} \frac{f(\vec{x})-f(\vec{x}_0)}{\vec{x}-\vec{x}_0},$$

denn im Nenner steht der Vektor $\vec{x} - \vec{x}_0$, und eine Division durch Vektoren gibt es nicht (und wird es auch nie geben).
Der anschauliche Hintergrund ist: das Steigungsverhalten von f im Punkt $\vec{x}_0$ kann nicht durch eine einzelne Zahl beschrieben werden. Im $\mathbb{R}^m$ gibt es (unendlich) viele Richtungen, und in jeder einzelnen Richtung hat die Funktion f ein gewisses Verhalten. All diese (womöglich unterschiedlichen) Verhaltensweisen müssen sich im zu definierenden Ableitungsbegriff widerspiegeln.
Ab jetzt entwickeln wir zwei Konzepte zur Lösung dieses Problems:

1. totale Ableitung

- Diese entspricht der bekannten eindimensionalen Ableitung.
- Es ist ein starkes Konzept, welches eine schöne Theorie ermöglicht.
- Leider ist der Begriff der totalen Ableitung zum konkreten Rechnen nicht direkt geeignet.

2. partielle Ableitung

- Dies ist ein schwaches Konzept, das nur eine hässliche Theorie gestattet.
- Aber die Berechnung verläuft recht einfach wie in Mathematik I.

Und entscheidend ist natürlich

3. der Zusammenhang zwischen beiden Ableitungsbegriffen.

Wir beginnen mit der partiellen Ableitung und betrachten die totale Ableitung danach.

16.1 Richtungsableitung: Sei $M \subset \mathbb{R}^m$ eine offene Menge, $f\colon M \to \mathbb{R}$ eine Funktion, $\vec{x}_0 \in M$ ein Punkt, und sei $\vec{v} \in \mathbb{R}^m \setminus \{\vec{0}\}$ eine Richtung. Wenn der Grenzwert

$$(\partial_{\vec{v}} f)(\vec{x}_0) := \lim_{h\to 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0)}{h}$$

existiert, dann heißt f im Punkt $\vec{x}_0$ *in Richtung* $\vec{v}$ *differenzierbar*. Die Zahl $(\partial_{\vec{v}} f)(\vec{x}_0)$ heißt *Richtungsableitung von* f *im Punkt* $\vec{x}_0$ *in Richtung* $\vec{v}$.
Anschaulich bedeutet dies: in der Menge M befinden wir uns am Punkt $\vec{x}_0$. Wir laufen von dort in Richtung $\vec{v}$, und der Anstieg von f, den wir bei diesem Weg spüren, ist die Richtungsableitung von f im Punkt $\vec{x}_0$ in Richtung $\vec{v}$, siehe Abbildung 26.

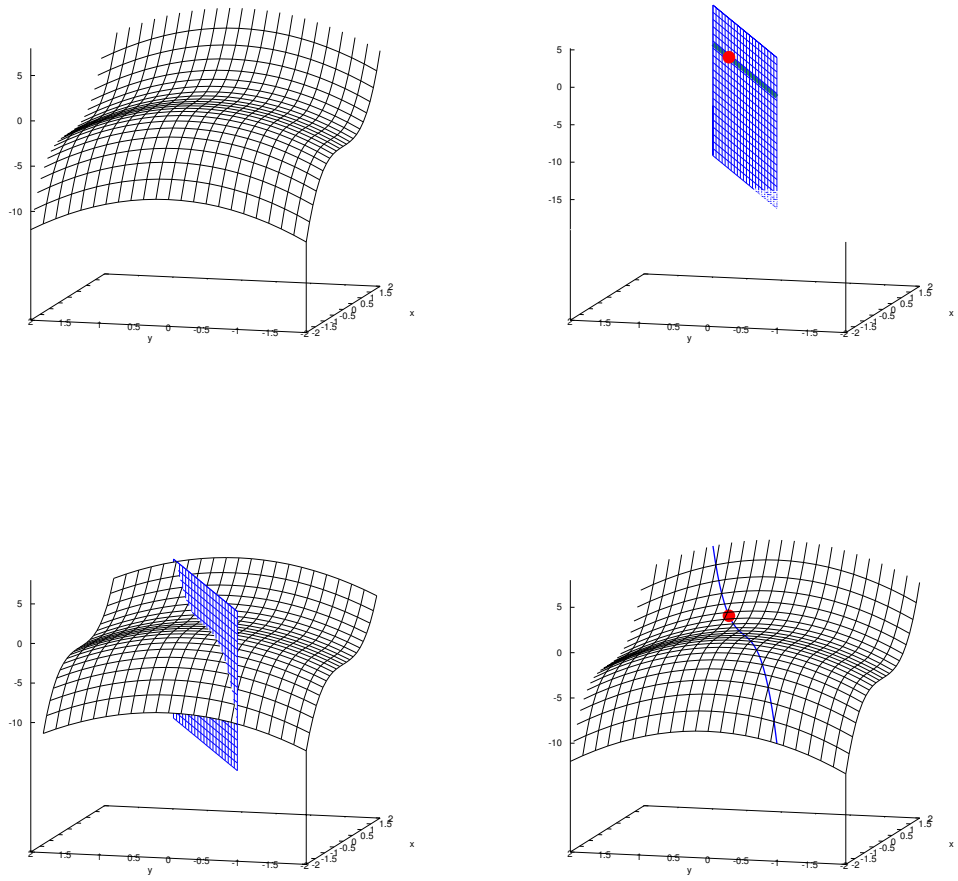


Abbildung 26: Links oben: die durch die Funktion $z = f(x, y) = x^3 - y^2$ beschriebene Fläche. Wir wollen im Punkt $[x_0, y_0]^T = [1, 0.5]^T$ die Richtungsableitung in Richtung $\vec{v} = \frac{1}{\sqrt{5}}[2, 1]^T$ bestimmen. Rechts oben: die Gerade durch den roten Punkt $[x_0, y_0]^T$ in Richtung $\vec{v}$ ist eine grüne Gerade in der xy-Ebene. Der Vektor $\vec{v}$ zeigt entlang der grünen Geraden „in das Papier hinein“. Wenn wir diese Gerade entlang der z-Achse parallel verschieben, entsteht eine blaue Ebene. Links unten: wir schneiden diese blaue Ebene mit der Fläche zur Funktion $z = f(x, y) = x^3 - y^2$. Rechts unten: als Schnittfigur entsteht eine blaue Kurve. Von dieser Kurve interessieren wir uns für das Steigungsverhalten am roten Punkt $[1, 0.5, 0.75]^T$. Beachte, dass der Vektor $\vec{v}$ in der xy-Ebene nach „links hinten“ zeigt, also wird die Richtungsableitung positiv sein. Weil wir $\vec{v}$ auf Länge Eins normiert haben, ist die Richtungsableitung von f gleich der schultypischen Ableitung, wenn wir die blaue Ebene als Papier-Ebene wählen. Bei anderen Längen von $\vec{v}$ käme noch ein entsprechender Korrekturfaktor hinzu.

16.2 Partielle Ableitung:

- Wenn $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_m$ die Standardbasis des $\mathbb{R}^m$ bezeichnet, dann nennen wir
$$\partial_j f(\vec{x}_0) := (\partial_{\vec{e}_j} f)(\vec{x}_0), \quad j = 1, 2, \ldots, m,$$
die *partielle Ableitung von* f in $\vec{x}_0$ nach der j-ten Koordinate. Wenn alle m partiellen Ableitungen im Punkt $\vec{x}_0$ existieren, dann nennt man f *partiell differenzierbar* in $\vec{x}_0$. Synonyme Schreibweisen sind
$$\partial_j f(\vec{x}_0) = \partial_{x_j} f(\vec{x}_0) = \frac{\partial f}{\partial x_j}(\vec{x}_0) = f_{x_j}(\vec{x}_0).$$

- Wenn $f\colon M \to \mathbb{R}$ in jedem Punkt $\vec{x}_0 \in M$ partiell differenzierbar ist, und wenn die m Funktionen $\vec{x}_0 \mapsto \partial_j f(\vec{x}_0)$ (wobei $j = 1, \ldots, m$) stetig in M sind, so heißt die Funktion f *auf* M *stetig partiell differenzierbar*.

16.3 Bestimmung der partiellen Ableitungen: Wir wählen die Richtung $\vec{v} = \vec{e}_1$. Dann ist
$$\begin{aligned}\partial_1 f(x_1, x_2, \ldots, x_m) &= \lim_{h\to 0} \frac{f(\vec{x} + h\vec{e}_1) - f(\vec{x})}{h} \\ &= \lim_{h\to 0} \frac{f(x_1 + h, x_2, \ldots, x_m) - f(x_1, x_2, \ldots, x_m)}{h}.\end{aligned}$$
Das bedeutet: wir betrachten die restlichen Variablen x_2, …, x_m als festgefrorene Parameter, und lediglich x_1 ist beweglich. Wir leiten diese Funktion dann „wie gewohnt“ nach der Variablen x_1 ab.
Die Ableitungen nach den anderen Variablen x_2, x_3, …, x_m werden entsprechend gebildet.

16.4 Beispiel: Wir betrachten die Funktion $f\colon \mathbb{R}^3 \to \mathbb{R}$ mit $f(x, y, z) = xe^{xz+y^2}$. Dann rechnen wir wie folgt:
$$\begin{aligned}\frac{\partial f}{\partial x}(x, y, z) &= \partial_1 f(x, y, z) = e^{xz+y^2} + xe^{xz+y^2} \cdot z, \\ \frac{\partial f}{\partial y}(x, y, z) &= \partial_2 f(x, y, z) = xe^{xz+y^2} \cdot 2y, \\ \frac{\partial f}{\partial z}(x, y, z) &= \partial_3 f(x, y, z) = xe^{xz+y^2} \cdot x = x^2 e^{xz+y^2}.\end{aligned}$$

16.5 Gradient: Sei die Funktion f im Punkt $\vec{x}_0$ partiell differenzierbar. Dann heißt
$$\nabla f(\vec{x}_0) := [\partial_1 f(\vec{x}_0), \partial_2 f(\vec{x}_0), \ldots, \partial_m f(\vec{x}_0)]$$
Gradient von f in $\vec{x}_0$.
Man beachte, dass der Gradient ein Zeilenvektor und kein Spaltenvektor ist.

Der Gradient besitzt folgende anschauliche Bedeutung: wenn wir innerhalb des $\mathbb{R}^m$ an dem Punkt $\vec{x}_0$ den Vektor $\nabla f(\vec{x}_0)$ anheften, dann liegt dieser Vektor wie folgt:

- der Vektor $\nabla f(\vec{x}_0)$ zeigt die Richtung an, in der der Graph von f am steilsten ansteigt,
- die Länge von $\nabla f(\vec{x}_0)$, also $\|\nabla f(\vec{x}_0)\|$, ist die Steigung dieses steilsten Anstiegs,
- der Vektor $\nabla f(\vec{x}_0)$ steht senkrecht auf der Höhenlinie durch den Punkt $\vec{x}_0$.

16.6 Beispiel: Wir wählen im $\mathbb{R}^2$ die offene Einheitskreisscheibe M mit

$$M := \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \colon x^2 + y^2 < 1 \right\},$$

und dann betrachten wir die Funktion $f \colon M \to \mathbb{R}$, gegeben durch

$$f(x,y) := \sqrt{1 - x^2 - y^2}.$$

Anschaulich beschreibt der Graph dieser Funktion eine obere Halbkugel. Dann haben wir den Gradienten

$$\begin{aligned} \nabla f(x,y) &= \left[\frac{\partial f}{\partial x}(x,y), \quad \frac{\partial f}{\partial y}(x,y) \right] = \left[\frac{-2x}{2\sqrt{1-x^2-y^2}}, \quad \frac{-2y}{2\sqrt{1-x^2-y^2}} \right] \\ &= \frac{-1}{\sqrt{1-x^2-y^2}} [x, \quad y]. \end{aligned}$$

Siehe Abbildung 27.

16.7 Partielle Ableitungen von Vektorfeldern: Sei $M \subset \mathbb{R}^m$ und $\vec{x}_0 \in M$. Ein Vektorfeld $F = [f_1, f_2, \ldots, f_n]^T \colon M \to \mathbb{R}^n$ heißt *partiell differenzierbar* in $\vec{x}_0$, falls die Komponenten f_1, f_2, …, f_n dies sind. Wir schreiben dann

$$J_F(\vec{x}_0) := \begin{bmatrix} \nabla f_1(\vec{x}_0) \\ \nabla f_2(\vec{x}_0) \\ \vdots \\ \nabla f_n(\vec{x}_0) \end{bmatrix} = \begin{bmatrix} \partial_1 f_1(\vec{x}_0) & \partial_2 f_1(\vec{x}_0) & \ldots & \partial_m f_1(\vec{x}_0) \\ \partial_1 f_2(\vec{x}_0) & \partial_2 f_2(\vec{x}_0) & \ldots & \partial_m f_2(\vec{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_n(\vec{x}_0) & \partial_2 f_n(\vec{x}_0) & \ldots & \partial_m f_n(\vec{x}_0) \end{bmatrix} \in \mathbb{R}^{n \times m}$$

und nennen diese Matrix *Jacobi-Matrix von F in $\vec{x}_0$*.
Wir führen auch partielle Ableitungen des Vektorfelds F ein:

$$\partial_j F(\vec{x}_0) = \begin{bmatrix} \partial_j f_1(\vec{x}_0) \\ \partial_j f_2(\vec{x}_0) \\ \vdots \\ \partial_j f_n(\vec{x}_0) \end{bmatrix}, \quad j = 1, \ldots, m.$$

Schließlich sind auch die Richtungsableitungen des Vektorfeldes F durch jene der Komponentenfunktionen gegeben, d.h. es gilt für $\vec{v} \in \mathbb{R}^m \setminus \{\vec{0}\}$

$$\partial_{\vec{v}} F(\vec{x}_0) = \begin{bmatrix} \partial_{\vec{v}} f_1(\vec{x}_0) \\ \partial_{\vec{v}} f_2(\vec{x}_0) \\ \vdots \\ \partial_{\vec{v}} f_n(\vec{x}_0) \end{bmatrix}.$$

16.8 Beispiel: Wir betrachten die Funktion F, die von Polarkoordinaten zu den üblichen kartesischen Koordinaten umrechnet, vgl. 3.6. Dabei ist $F \colon (0,\infty) \times (-\pi, \pi) \to \mathbb{R}^2$ mit

$$F(r,\varphi) = \begin{bmatrix} r\cos\varphi \\ r\sin\varphi \end{bmatrix} = \begin{bmatrix} f_1(r,\varphi) \\ f_2(r,\varphi) \end{bmatrix}.$$

Die Jacobi-Matrix berechnen wir dann so:

$$J_F(r,\varphi) = \begin{bmatrix} \partial_r f_1(r,\varphi) & \partial_\varphi f_1(r,\varphi) \\ \partial_r f_2(r,\varphi) & \partial_\varphi f_2(r,\varphi) \end{bmatrix} = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix}.$$

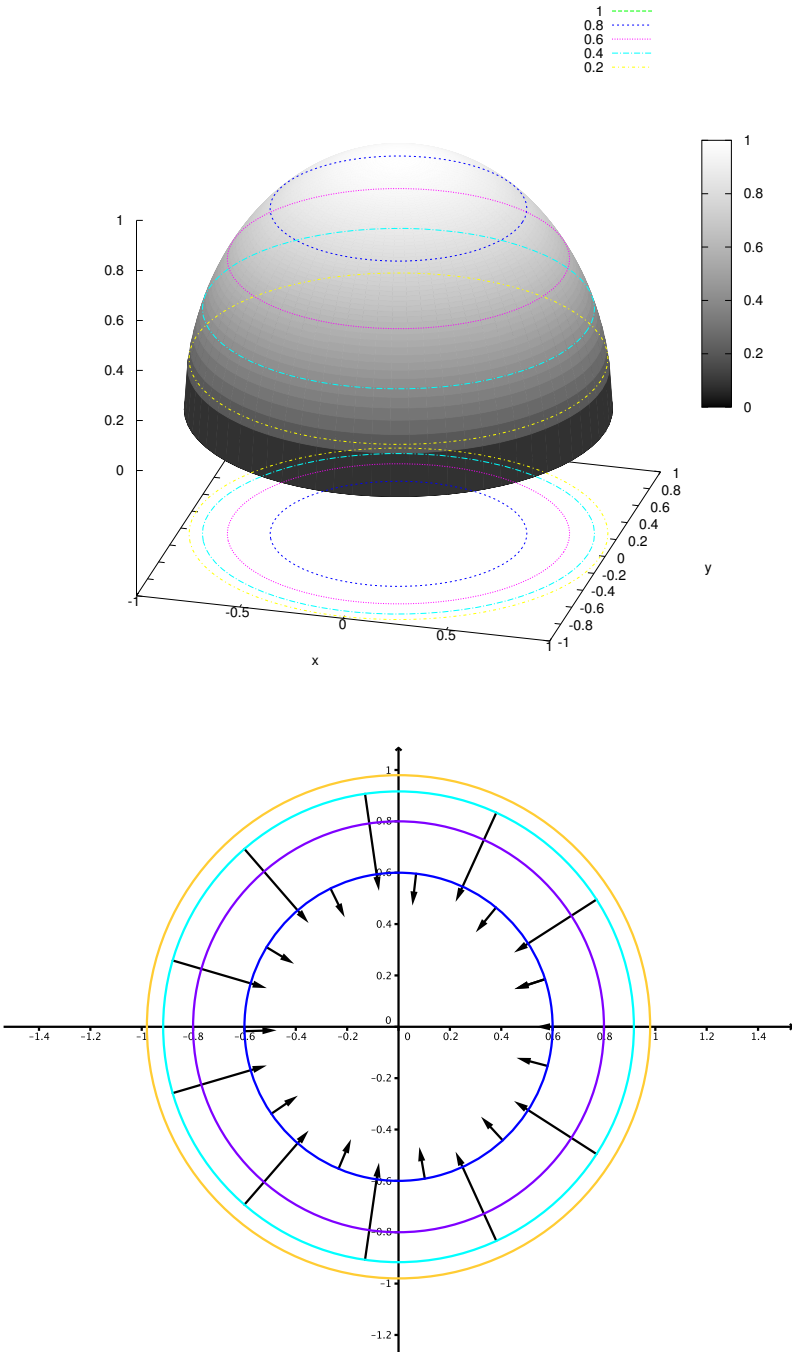


Abbildung 27: Oben: die skalare Funktion $f(x,y) = \sqrt{1-x^2-y^2}$. Die Höhenlinien sind zweifach eingetragen worden: einmal auf der Halbkugelfläche, und ein weiteres mal im Definitionsbereich $M \subset \mathbb{R}^2$. Unten sieht man die Höhenlinien in M und einige Gradientenvektoren. Die Gradientenvektoren sind um den Faktor 6 verkürzt eingezeichnet, damit die Graphik nicht überquillt. In der Nähe des Äquators ist die Halbkugelfläche steil, und deshalb sind dort die Gradientenvektoren lang. In der Nähe des Nordpols ist die Halbkugelfläche vergleichsweise „unsteil“, und deshalb sind dort die Gradientenvektoren kurz.

16.9 Kettenregel: Seien $F\colon \mathbb{R}^m \to \mathbb{R}^n$ und $G\colon \mathbb{R}^n \to \mathbb{R}^p$ stetig partiell differenzierbar. Dann ist auch die zusammengesetzte Funktion

$$H := G \circ F\colon \mathbb{R}^m \to \mathbb{R}^p$$

partiell differenzierbar, und es ist

$$J_H(\vec{x}_0) = J_G(F(\vec{x}_0)) \cdot J_F(\vec{x}_0).$$

Man beobachte, dass die Matrix-Matrix-Multiplikation auf der rechten Seite tatsächlich durchführbar ist, weil die Abmessungen der beiden Matrizen wirklich kompatibel sind!

16.10 Beispiel: Wir betrachten eine „äußere“ Funktion $G\colon \mathbb{R}^2 \to \mathbb{R}^2$, gegeben durch

$$G(p,q) = \begin{bmatrix} e^p \cos(q) \\ e^p \sin(q) \end{bmatrix},$$

und eine „innere“ Funktion $F\colon \mathbb{R}^3 \to \mathbb{R}^2$, gegeben durch

$$F(x,y,z) = \begin{bmatrix} x+y \\ xz \end{bmatrix}.$$

Dann ergibt sich die zusammengesetzte („verkettete“) Funktion

$$(G \circ F)(x,y,z) = G(F(x,y,z)) = G(x+y, xz) = \begin{bmatrix} e^{x+y}\cos(xz) \\ e^{x+y}\sin(xz) \end{bmatrix}.$$

Die Jacobi-Matrizen der drei Funktionen $G\circ F$, G und F stehen zueinander in Beziehung wie folgt:

$$J_{G\circ F}(x,y,z) = J_G(F(x,y,z)) \cdot J_F(x,y,z).$$

Wir rechnen die linke Seite aus, indem wir die beiden Faktoren auf der rechten Seite ermitteln: Zunächst ist

$$J_G(p,q) = \begin{bmatrix} e^p\cos(q) & -e^p\sin(q) \\ e^p\sin(q) & e^p\cos(q) \end{bmatrix}.$$

Für (p,q) setzen wir jetzt $F(x,y,z)$ ein:

$$\begin{aligned} J_G(F(x,y,z)) = J_G(x+y,xz) &= \begin{bmatrix} e^{x+y}\cos(xz) & -e^{x+y}\sin(xz) \\ e^{x+y}\sin(xz) & e^{x+y}\cos(xz) \end{bmatrix} \\ &= e^{x+y}\begin{bmatrix} \cos(xz) & -\sin(xz) \\ \sin(xz) & \cos(xz) \end{bmatrix}. \end{aligned}$$

Damit erhalten wir also, wie angekündigt:

$$\begin{aligned} J_{G\circ F}(x,y,z) &= e^{x+y}\begin{bmatrix} \cos(xz) & -\sin(xz) \\ \sin(xz) & \cos(xz) \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 0 \\ z & 0 & x \end{bmatrix} \\ &= e^{x+y}\begin{bmatrix} \cos(xz) - z\sin(xz) & \cos(xz) & -x\sin(xz) \\ \sin(xz) + z\cos(xz) & \sin(xz) & x\cos(xz) \end{bmatrix}. \end{aligned}$$

16.11 Partielle Ableitungen höherer Ordnung: Sei $M \subset \mathbb{R}^m$, und sei $f\colon M \to \mathbb{R}$ stetig partiell differenzierbar auf M. Sind die partiellen Ableitungen von f in $\vec{x}_0 \in M$ erneut partiell differenzierbare Funktionen, so setzen wir

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{x}_0) := \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(\vec{x}_0).$$

Die folgenden alternativen Schreibweisen sind gebräuchlich:

$$\partial_i \partial_j f(\vec{x}_0), \qquad \partial_{x_i}\partial_{x_j} f(\vec{x}_0), \qquad f_{x_i x_j}(\vec{x}_0).$$

Analog führt man höhere Ableitungen (und deren Schreibweisenvarianten) ein, zum Beispiel

$$\partial_1\partial_3\partial_1 f, \quad \frac{\partial^3 f}{\partial x_1 \partial x_1^2}, \quad f_{xzy}, \quad f_{xxxxxy}, \quad \ldots.$$

Die Anzahl der Differenziationen bezeichnet man dabei als *Ordnung* der partiellen Ableitung. Man sagt eine Funktion ist *k-mal differenzierbar*, falls alle partiellen Ableitungen der Ordnung k existieren und sie ist *k-mal stetig differenzierbar*, wenn alle partiellen Ableitungen der Ordnung k existieren und selbst wieder stetige Funktionen sind.

Die partiellen Ableitungen zweiter Ordnung ordnen wir jetzt etwas systematischer an:

16.12 Hesse-Matrix: Wenn die Funktion $f\colon \mathbb{R}^m \to \mathbb{R}$ zweimal im Punkt $\vec{x}_0$ partiell differenzierbar ist, dann nennen wir die Matrix

$$H_f(\vec{x}_0) := \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\vec{x}_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{x}_0) & \ldots & \frac{\partial^2 f}{\partial x_1 \partial x_m}(\vec{x}_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\vec{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(\vec{x}_0) & \ldots & \frac{\partial^2 f}{\partial x_2 \partial x_m}(\vec{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1}(\vec{x}_0) & \frac{\partial^2 f}{\partial x_m \partial x_2}(\vec{x}_0) & \ldots & \frac{\partial^2 f}{\partial x_m \partial x_m}(\vec{x}_0) \end{bmatrix}$$

Hesse-Matrix von f in $\vec{x}_0$.

16.13 Beispiel: Wir betrachten die Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$, gegeben durch die Formel $f(x,y) = x^3 y + xe^y$. Die partiellen Ableitungen erster Ordnung bilden den Gradienten von f:

$$\nabla f(x,y) = [f_x(x,y), \quad f_y(x,y)] = \left[3x^2 y + e^y, \quad x^3 + xe^y\right].$$

Die vier partiellen Ableitungen zweiter Ordnung stehen in der Hesse-Matrix:

$$H_f(x,y) = \begin{bmatrix} f_{xx}(x,y) & f_{yx}(x,y) \\ f_{xy}(x,y) & f_{yy}(x,y) \end{bmatrix} = \begin{bmatrix} 6xy & 3x^2 + e^y \\ 3x^2 + e^y & xe^y \end{bmatrix}.$$

Und dann gibt es noch die Ableitungen dritter Ordnung (insgesamt 8 Stück):

$$\begin{aligned} &f_{xxx}(x,y) = 6y, \\ &f_{xxy}(x,y) = 6x, && f_{xyx}(x,y) = 6x, && f_{yxx}(x,y) = 6x, \\ &f_{yyx}(x,y) = e^y, && f_{yxy}(x,y) = e^y, && f_{xyy}(x,y) = e^y, \\ &f_{yyy}(x,y) = xe^y. \end{aligned}$$

Wir beobachten, dass die Hesse-Matrix symmetrisch ist. Und wir beobachten auch, dass einige der Ableitungen dritter Ordnung gleich sind. Das ist kein Zufall, wie sich aus der nächsten Nummer ergibt.

16.14 Vertauschen von partiellen Ableitungen: Es gilt der *Satz von Schwarz*:

Sei $M \subset \mathbb{R}^m$ offen, und die Funktion $f\colon M \to \mathbb{R}$ sei k mal stetig partiell differenzierbar, d.h. alle partiellen Ableitungen bis zur Ordnung k existieren und sind allesamt stetig. Dann sind alle partiellen Ableitungen bis zur Ordnung k vertauschbar (also $\partial_1\partial_2 f = \partial_2\partial_1 f$ usw.).

In der Einleitung wurde bereits erwähnt, dass die partiellen Ableitungen zwar recht einfach ausgerechnet werden können, aber nur eine hässliche Theorie zulassen. Jetzt kommen wir zu diesen unangenehmen Eigenschaften der partiellen Ableitungen, wobei wir die schreckliche Funktion aus Beispiel 15.10 noch einmal aufgreifen. Siehe Bild 22 für eine bildliche Darstellung dieser Funktion, die im Nullpunkt eben nicht stetig ist.

16.15 Beispiel:

- Wir betrachten die Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$, die gegeben ist durch

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & : \text{ wenn } (x,y) \neq (0,0), \\ 0 & : \text{ wenn } (x,y) = (0,0). \end{cases}$$

  Für jeden Punkt $(x,y) \neq (0,0)$ gilt

$$\frac{\partial f}{\partial x}(x,y) = \frac{y(x^2+y^2) - 2x^2y}{(x^2+y^2)^2} = \frac{y^3 - x^2y}{(x^2+y^2)^2}, \qquad \frac{\partial f}{\partial y}(x,y) = (\text{analog}) = \frac{x^3 - 2xy^2}{(x^2+y^2)^2}.$$

  Im Ursprung gilt (gemäß Definition der partiellen Ableitungen)

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{0}{h^2} - 0}{h} = 0, \qquad \frac{\partial f}{\partial y}(0,0) = (\text{analog}) = 0.$$

  Das bedeutet: diese Funktion f ist im Ursprung partiell differenzierbar (denn dort existieren die partiellen Ableitungen tatsächlich), aber im Nullpunkt unstetig. Das ist seltsam, denn eigentlich würde man erwarten, dass Differenzierbarkeit eine stärkere Eigenschaft ist als Stetigkeit.

- Noch wilder ist folgende Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$:

$$f(x,y) = \begin{cases} 0 & : \text{ wenn } 0 < x^3 < y < x^2, \\ 1 & : \text{ sonst.} \end{cases}$$

  Bei dieser Funktion existieren im Ursprung sämtliche Richtungsableitungen, und alle haben den Wert Null. Trotzdem ist die Funktion f nicht stetig im Ursprung. Siehe Bild 28.

Die Angelegenheit wird viel schöner, wenn wir zum Konzept der **Totalen Differenzierbarkeit** übergehen.

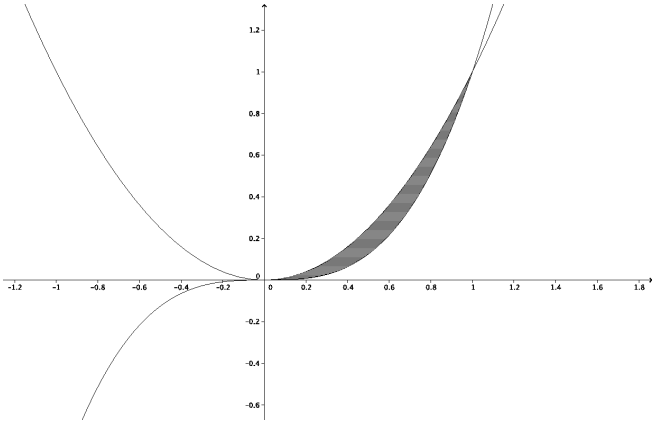


Abbildung 28: Die Funktion aus Beispiel 16.15 hat den Wert 0 im schraffierten Bereich, und überall sonst (insbesondere im Nullpunkt) hat sie den Wert 1.

16.16 Erinnerung: In der Vorlesung Mathematik I haben wir gelernt, vgl. 11.6: eine Funktion $f\colon \mathbb{R} \to \mathbb{R}$ ist genau dann differenzierbar im Punkt x_0, wenn man die Werte $f(x)$ für x aus einer Umgebung von x_0 wie folgt in drei Anteile zerlegen kann:

$$f(x) = f(x_0) + f'(x_0) \cdot (x - x_0) + r(x),$$

siehe auch Bild 20, wobei der Restsummand $r(x)$ folgende Bedingung erfüllt:

$$\lim_{x \to x_0} \frac{r(x)}{x - x_0} = 0.$$

Dies bedeutet, dass $r(x)$ für $x \to x_0$ schneller als $x - x_0$ nach Null strebt. Wir können die angegebene Bedingung auch umformulieren zu

$$\lim_{x \to x_0} \frac{r(x)}{|x - x_0|} = 0.$$

Im Nenner steht jetzt der Abstand der Zahlen x und x_0 auf der Zahlengeraden. Der Abstand ist eine reelle Zahl, und durch Zahlen ($\neq 0$) darf man dividieren. Daraus gewinnen wir eine Idee, wie wir einen Differenzierbarkeitsbegriff im $\mathbb{R}^m$ definieren könnten, siehe 16.17.

16.17 Totale Differenzierbarkeit: Sei $M \subset \mathbb{R}^m$ eine offene Menge, $\vec{x}_0 \in M$. Eine Funktion $F\colon M \to \mathbb{R}^n$ heißt *total differenzierbar* in $\vec{x}_0$, wenn es eine Matrix $A \in \mathbb{R}^{n \times m}$ und eine Funktion $R\colon M \to \mathbb{R}^n$ gibt mit der Eigenschaft

$$F(\vec{x}) = F(\vec{x}_0) + A \cdot (\vec{x} - \vec{x}_0) + R(\vec{x}), \quad \vec{x} \in M,$$

wobei der Summand $R(\vec{x})$ folgende Bedingung erfüllt:

$$\lim_{\vec{x} \to \vec{x}_0} \frac{R(\vec{x})}{\|\vec{x} - \vec{x}_0\|} = \vec{0}.$$

Die Matrix A heißt *totale Ableitung von F in $\vec{x}_0$*. Wir schreiben $DF(\vec{x}_0) := A$.

16.18 Differenzierbarkeit und Stetigkeit: Wir machen uns nun klar, dass der Begriff der totalen Differenzierbarkeit unsere Erwartungen erfüllt:
Wenn eine Funktion F im Punkt $\vec{x}_0$ total differenzierbar ist, dann ist F in diesem Punkt stetig.
Um das einzusehen, zerlegen wir $F(\vec{x})$ wie in 16.17 in die drei Anteile. Wir wollen zeigen, dass die Funktion F im Punkt $\vec{x}_0$ stetig ist, also dass $\lim_{\vec{x}\to\vec{x}_0} F(\vec{x}) = F(\vec{x}_0)$ gilt. Dazu rechnen wir den Limes einfach aus:

$$
\begin{aligned}
&\lim_{\vec{x}\to\vec{x}_0} F(\vec{x}) && \text{zerlege } F(\vec{x}) \text{ in die drei Teile}\\
&= \lim_{\vec{x}\to\vec{x}_0} \Big(F(\vec{x}_0) + A\cdot(\vec{x}-\vec{x}_0) + R(\vec{x})\Big) && \text{ziehe den Limes auf jeden Summanden}\\
&= \lim_{\vec{x}\to\vec{x}_0} F(\vec{x}_0) + \lim_{\vec{x}\to\vec{x}_0} A\cdot(\vec{x}-\vec{x}_0) + \lim_{\vec{x}\to\vec{x}_0} R(\vec{x}) && \text{vereinfache ersten Summanden}\\
&= F(\vec{x}_0) + \lim_{\vec{x}\to\vec{x}_0} A\cdot(\vec{x}-\vec{x}_0) + \lim_{\vec{x}\to\vec{x}_0} R(\vec{x}) && \text{lineare Abbildungen sind stetig (s. 15.10)}\\
&= F(\vec{x}_0) + A\cdot\vec{0} + \lim_{\vec{x}\to\vec{x}_0} R(\vec{x}) && \text{erweitere clever}\\
&= F(\vec{x}_0) + \lim_{\vec{x}\to\vec{x}_0} \left(\frac{R(\vec{x})}{\|\vec{x}-\vec{x}_0\|}\cdot\|\vec{x}-\vec{x}_0\|\right) && \text{ziehe Limes auf jeden Faktor}\\
&= F(\vec{x}_0) + \left(\lim_{\vec{x}\to\vec{x}_0} \frac{R(\vec{x})}{\|\vec{x}-\vec{x}_0\|}\right)\cdot \lim_{\vec{x}\to\vec{x}_0} \|\vec{x}-\vec{x}_0\| && \text{nutze Wissen über } R(\vec{x})\\
&= F(\vec{x}_0) + \vec{0}\cdot 0 = F(\vec{x}_0).
\end{aligned}
$$

Das wollten wir haben.

16.19 Zusammenhang zwischen partiellen und totalen Ableitungen: Sei $M \subset \mathbb{R}^m$ offen, $\vec{x}_0 \in M$ und $F\colon M \to \mathbb{R}^n$.

- Wenn F total differenzierbar in $\vec{x}_0$ ist, dann existieren alle Richtungsableitungen $\partial_{\vec{v}}F(\vec{x}_0)$ im Punkt $\vec{x}_0$, für alle $\vec{v} \in \mathbb{R}^m \setminus \{0\}$, und es gilt

  $$\partial_{\vec{v}}F(\vec{x}_0) = DF(\vec{x}_0)\cdot\vec{v},$$

  wobei rechts ein Produkt von Matrix und Vektor steht.

  Zusätzlich gilt: die Ableitung $DF(\vec{x}_0)$ ist gleich der Jacobi-Matrix $J_F(\vec{x}_0)$ (bzw. gleich dem Gradienten von F, wenn $n = 1$).

  Um das einzusehen, verwenden wir wieder, dass wir die in $\vec{x}_0$ total differenzierbare Funktion F in die drei Teile zerlegen können. Damit können wir dann $\partial_{\vec{v}}F(\vec{x}_0)$ auszurechnen. Sei dazu $\vec{v} \neq \vec{0}$ eine beliebige Richtung. Dann haben wir:

  $$
  \begin{aligned}
  (\partial_{\vec{v}}F)(\vec{x}_0) &= \lim_{h\to 0} \frac{1}{h}\Big(F(\vec{x}_0+h\vec{v}) - F(\vec{x}_0)\Big)\\
  &= \lim_{h\to 0} \frac{1}{h}\Big(\cancel{F(\vec{x}_0)} + DF(\vec{x}_0)\cdot h\vec{v} + R(\vec{x}_0+h\vec{v}) - \cancel{F(\vec{x}_0)}\Big)\\
  &= \lim_{h\to 0} \frac{DF(\vec{x}_0)\cdot \cancel{h}\vec{v}}{\cancel{h}} + \lim_{h\to 0} \frac{R(\vec{x}_0+h\vec{v})}{h}\\
  &= DF(\vec{x}_0)\cdot\vec{v} + \lim_{h\to 0} \left(\frac{R(\vec{x}_0+h\vec{v})}{\|(\vec{x}_0+h\vec{v})-\vec{x}_0\|}\cdot\|\vec{v}\|\right)
  \end{aligned}
  $$

$$= DF(\vec{x}_0) \cdot \vec{v} + \left(\lim_{h \to 0} \frac{R(\vec{x}_0 + h\vec{v})}{\|(\vec{x}_0 + h\vec{v}) - \vec{x}_0\|} \right) \cdot \|\vec{v}\|$$
$$= DF(\vec{x}_0) \cdot \vec{v} + \vec{0} \cdot \|\vec{v}\| = DF(\vec{x}_0) \cdot \vec{v}.$$

Das wollten wir zeigen.

Insbesondere gilt für $j = 1, \ldots, m$, dass

$$\partial_j F(\vec{x}_0) = \partial_{\vec{e}_j} F(\vec{x}_0) = DF(\vec{x}_0) \cdot \vec{e}_j.$$

Gemäß Definition 16.7 steht links die j-te Spalte der Jacobi-Matrix $J_F(\vec{x}_0)$, und da $\vec{e}_j$ der j-te Standardbasisvektor ist, steht rechts die j-te Spalte der Matrix $DF(\vec{x}_0)$. Diese Argumentation klappt für jedes $j = 1, \ldots, m$. Also ist $DF(\vec{x}_0) = J_F(\vec{x}_0)$, wenn F in $\vec{x}_0$ total differenzierbar ist.

- Die Funktion f aus Beispiel 16.15 ist partiell differenzierbar im Nullpunkt, aber sie kann dort nicht total differenzierbar sein, denn dann wäre sie ja stetig (wegen 16.18).

  Das zeigt uns: Wenn F partiell differenzierbar ist, dann braucht F nicht total differenzierbar sein. Jedoch gilt: wenn die partiellen Ableitungen stetig sind (also wenn die Funktion F stetig partiell differenzierbar ist), dann ist die Funktion F total differenzierbar.

Insgesamt bekommen wir folgendes **Verfahren zur Bestimmung der totalen Ableitung** von F:

- wir bestimmen die partiellen Ableitungen $\frac{\partial F_i}{\partial x_k}$ mit den Rechenmethoden aus der Mathematik I (bzw. der Schule), indem wir die anderen Variablen festhalten und lediglich x_k laufen lassen,
- all diese partiellen Ableitungen bilden die Einträge der Jacobi-Matrix $J_F(\vec{x}_0)$,
- wir prüfen nach, ob diese partiellen Ableitungen $\frac{\partial F_i}{\partial x_k}$ stetig sind,
- wenn ja: dann ist die totale Ableitung gleich der Jacobi-Matrix $J_F(\vec{x}_0)$, bzw. dem Gradienten, wenn die Funktion F skalarwertig ist.

Die Glattheitseigenschaften von Funktionen stehen wie folgt in Beziehung zueinander:

stetig partiell differenzierbar	$\Longrightarrow$	total differenzierbar	$\Longrightarrow$	stetig
$\Downarrow$		$\Downarrow$		
partiell differenzierbar	$\Longleftarrow$	alle Richtungsableitungen existieren		

Alle Umkehrungen sind falsch!

16.20 Beispiel: Eine totale Ableitung nach einer einzelnen Variablen (z.B. nach einer Zeitvariablen t) wird oft als $\frac{\mathrm{d}}{\mathrm{d}t}$ geschrieben. Wir müssen sorgfältig unterscheiden:

$\frac{\partial}{\partial t}$ gibt uns die partielle Ableitung nach der **Koordinate** t,

$\frac{\mathrm{d}}{\mathrm{d}t}$ gibt uns die totale Ableitung nach der **Variablen** t.

Als Beispiel betrachten wir die Funktion f mit der Gleichung $f(x,y) = x^3y + xe^y$, vgl. 16.13, und suchen $\frac{\mathrm{d}}{\mathrm{d}t} f(t^2, t^3)$.
Dann ist

$$f(t^2, t^3) = (t^2)^3 t^3 + t^2 e^{t^3} = t^9 + t^2 e^{t^3},$$

also bekommen wir das Ergebnis

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t^2, t^3) = 9t^8 + 2te^{t^3} + t^2 e^{t^3} \cdot 3t^2 = 9t^8 + e^{t^3}(2t + 3t^4).$$

Alternativ bekommt man das selbe Ergebnis mit der Kettenregel, vgl. 16.9:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} f(t^2, t^3) &= \frac{\partial f}{\partial x}(t^2, t^3) \cdot 2t + \frac{\partial f}{\partial y}(t^2, t^3) \cdot 3t^2 = \left(3(t^2)^2 t^3 + e^{t^3}\right) \cdot 2t + \left((t^2)^3 + t^2 e^{t^3}\right) \cdot 3t^2 \\ &= 6t^8 + 2te^{t^3} + 3t^8 + 3t^4 e^{t^3} = 9t^8 + e^{t^3}(3t^4 + 2t).\end{aligned}$$

17 Taylor-Reihen in mehreren Veränderlichen

Unser **Ziel** ist der Satz von Taylor für Funktionen von mehreren Variablen.
Wir **erinnern** an Abschnitt 13: dazu betrachten wir eine Funktion $f\colon \mathbb{R} \to \mathbb{R}$ und eine Entwicklungsstelle $x_0 \in \mathbb{R}$. Dann haben wir die Zerlegung

$$f(x) = T_k f(x) + R_k f(x),$$

wobei $T_k f$ das *Taylorpolynom* vom Grade k ist, d.h.

$$T_k f(x) = \sum_{j=0}^{k} \frac{f^{(j)}(x_0)}{j!}(x - x_0)^j,$$

und für das *Restglied* $R_k f(x)$ gibt es die Formel

$$R_k(x) = \frac{f^{(k+1)}(\xi)}{(k+1)!}(x - x_0)^{k+1},$$

wobei von der Zahl ξ lediglich bekannt ist, dass sie irgendwo zwischen x und x_0 liegt.
Jetzt gehen wir mit dem x vom $\mathbb{R}^1$ in den $\mathbb{R}^m$ und betrachten zunächst die Taylorentwicklung mit quadratischem Restglied.

17.1 Taylorpolynom ersten Grades und Restglied: Sei $f\colon \mathbb{R}^m \to \mathbb{R}$ zweimal stetig partiell differenzierbar und $\vec{x}_0 \in \mathbb{R}^m$. Dann gibt es für jedes $\vec{x} \in \mathbb{R}^m$ ein $\vec{\xi}$ auf der Verbindungsstrecke von $\vec{x}$ nach $\vec{x}_0$ mit

$$f(\vec{x}) = \underbrace{f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)}_{\text{Taylorpolynom ersten Grades}} + \underbrace{\frac{1}{2}(\vec{x} - \vec{x}_0)^T \cdot H_f(\vec{\xi}) \cdot (\vec{x} - \vec{x}_0)}_{\text{Restglied}}.$$

Beachte, dass $\nabla f(\vec{x}_0)$ ein Zeilenvektor ist und $(\vec{x} - \vec{x}_0)$ ein Spaltenvektor, sodass im Produkt $\nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$ die Abmessungen kompatibel sind. Entsprechendes gilt für das Drei-Faktor-Produkt im Restglied.

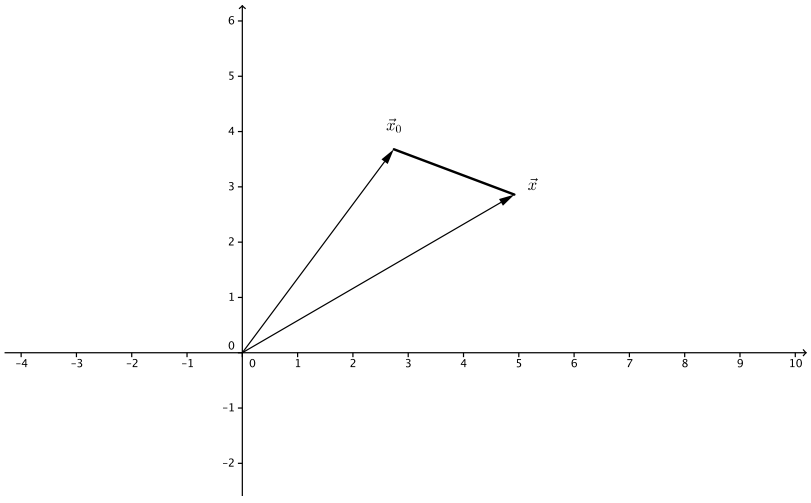


Abbildung 29: Die Verbindungsstrecke zwischen $\vec{x}$ und $\vec{x}_0$.

Die *Verbindungsstrecke* S zwischen $\vec{x}$ und $\vec{x}_0$ kann man wie folgt beschreiben:

$$S = \{\vec{x}_0 + \lambda(\vec{x} - \vec{x}_0)\colon 0 \le \lambda \le 1\} = \{\lambda\vec{x} + (1 - \lambda)\vec{x}_0\colon 0 \le \lambda \le 1\}.$$

Irgendwo auf der Strecke S liegt der Punkt $\vec{\xi}$ (mehr wissen wir über $\vec{\xi}$ nicht).

17.2 Taylorpolynom zweiten Grades: Man kann auch quadratische Taylorpolynome betrachten, die dann aussehen wie folgt:

$$f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + \frac{1}{2}(\vec{x} - \vec{x}_0)^T \cdot H_f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0),$$

und das dazugehörige Restglied wäre dann „kubisch in Bezug auf die Differenz $\vec{x} - \vec{x}_0$“. Allerdings ist die Formel für das Restglied dann sehr kompliziert, weshalb wir darauf verzichten, es hinzuschreiben.
Wir wollen exemplarisch für den Fall $m = 2$ (also für zwei Variablen x und y) das Taylorpolynom zweiten Grades ermitteln. Wir haben also

$$\vec{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \vec{x}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix},$$

und das Taylorpolynom ergibt sich dann zu

$$\begin{aligned}
&f(\vec{x}_0) + \begin{bmatrix} \frac{\partial f}{\partial x}(\vec{x}_0) & \frac{\partial f}{\partial y}(\vec{x}_0) \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} x - x_0 & y - y_0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(\vec{x}_0) & \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) \\ \frac{\partial^2 f}{\partial y \partial x}(\vec{x}_0) & \frac{\partial^2 f}{\partial y^2}(\vec{x}_0) \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\
&= f(\vec{x}_0) + \frac{\partial f}{\partial x}(\vec{x}_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(\vec{x}_0) \cdot (y - y_0) \\
&\quad + \frac{1}{2} \begin{bmatrix} x - x_0 & y - y_0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(\vec{x}_0) \cdot (x - x_0) + \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) \cdot (y - y_0) \\ \frac{\partial^2 f}{\partial y \partial x}(\vec{x}_0) \cdot (x - x_0) + \frac{\partial^2 f}{\partial y^2}(\vec{x}_0) \cdot (y - y_0) \end{bmatrix} \\
&= f(\vec{x}_0) + \frac{\partial f}{\partial x}(\vec{x}_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(\vec{x}_0) \cdot (y - y_0) \\
&\quad + \frac{1}{2} \left(\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) \cdot (x - x_0)^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) \cdot (x - x_0) \cdot (y - y_0) + \frac{\partial^2 f}{\partial y^2}(\vec{x}_0) \cdot (y - y_0)^2 \right) \\
&= \frac{f(\vec{x}_0)}{0!0!} \cdot (x - x_0)^0 \cdot (y - y_0)^0 + \frac{\frac{\partial f}{\partial x}(\vec{x}_0)}{1!0!} \cdot (x - x_0)^1 \cdot (y - y_0)^0 + \frac{\frac{\partial f}{\partial y}(\vec{x}_0)}{0!1!} \cdot (x - x_0)^0 \cdot (y - y_0)^1 \\
&\quad + \frac{\frac{\partial^2 f}{\partial x^2}(\vec{x}_0)}{2!0!} \cdot (x - x_0)^2 \cdot (y - y_0)^0 + \frac{\frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0)}{1!1!} \cdot (x - x_0)^1 \cdot (y - y_0)^1 \\
&\qquad\qquad + \frac{\frac{\partial^2 f}{\partial y^2}(\vec{x}_0)}{0!2!} \cdot (x - x_0)^0 \cdot (y - y_0)^2.
\end{aligned}$$

Hierbei haben wir verwendet: $0! = 1$, $1! = 1$, $2! = 1 \cdot 2 = 2$ usw. sowie $s^0 = 1$ für alle $s \in \mathbb{R}$.
Nach diesem Bildungsprinzip werden dann auch Taylorpolynome noch höheren Grades zusammengebaut.

17.3 Beispiel: Wir wollen für die Funktion f mit der Gleichung $f(x, y) = (1 + 3y^2)e^x$ im Punkt $[x_0, y_0]^T = [0, 0]^T$ das Taylorpolynom zweiten Grades aufstellen. Die Ableitungen und deren Werte sind:

$$\begin{aligned}
f(x, y) &= (1 + 3y^2)e^x, & f(0, 0) &= 1, \\
\frac{\partial f}{\partial x}(x, y) &= (1 + 3y^2)e^x, & \frac{\partial f}{\partial x}(0, 0) &= 1, \\
\frac{\partial f}{\partial y}(x, y) &= 6ye^x, & \frac{\partial f}{\partial x}(0, 0) &= 0,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2}(x,y) &= (1+3y^2)e^x, & \frac{\partial^2 f}{\partial x^2}(0,0) &= 1,\\
\frac{\partial^2 f}{\partial x \partial y}(x,y) &= 6ye^x, & \frac{\partial^2 f}{\partial x \partial y}(0,0) &= \frac{\partial^2 f}{\partial y \partial x}(0,0) = 0,\\
\frac{\partial^2 f}{\partial y^2}(x,y) &= 6e^x, & \frac{\partial^2 f}{\partial y^2}(0,0) &= 6.
\end{aligned}$$

Damit erhalten wir folgendes Taylorpolynom zweiten Grades:

$$\begin{aligned}
&f(0,0) + \frac{\partial f}{\partial x}(0,0)\cdot(x-0) + \frac{\partial f}{\partial y}(0,0)\cdot(y-0)\\
&\qquad + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(0,0)\cdot(x-0)^2 + \frac{\partial^2 f}{\partial x \partial y}(0,0)\cdot(x-0)\cdot(y-0) + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(0,0)\cdot(y-0)^2\\
&= 1 + 1\cdot x + 0\cdot y + \frac{1}{2}\cdot 1\cdot x^2 + 0\cdot x\cdot y + \frac{1}{2}\cdot 6\cdot y^2\\
&= 1 + x + \frac{1}{2}x^2 + 3y^2.
\end{aligned}$$

Wir hätten auch anders rechnen können, indem wir uns an die Reihe für e^x aus dem ersten Semester erinnern:

$$\begin{aligned}
f(x,y) = (1+3y^2)e^x &= (1+3y^2)\cdot\left(1 + x + \frac{x^2}{2} + O(x^3)\right) \qquad \Big| \text{ ausmultiplizieren}\\
&= \underbrace{1 + x + \frac{x^2}{2} + 3y^2}_{\text{alle Terme bis Grad 2}} + \underbrace{\ldots\ldots\ldots\ldots}_{\text{alle Terme ab Grad 3}}.
\end{aligned}$$

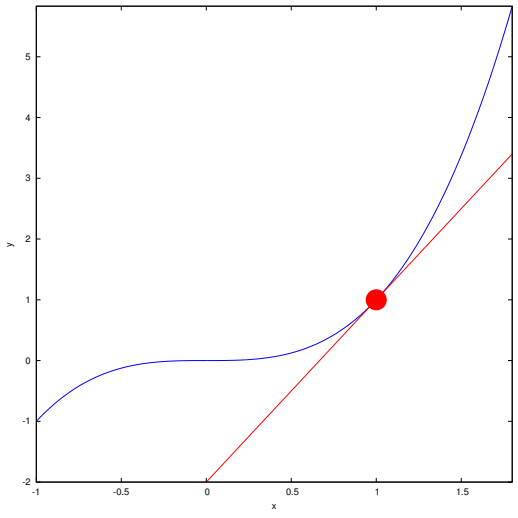


Abbildung 30: Die Tangente an den Graph der Funktion $y = f(x)$, genommen im Punkt $[x_0, f(x_0)]^T$, hört auf die Gleichung $y = f(x_0) + f'(x_0)\cdot(x - x_0)$. Wir bekommen diese Tangentengleichung, indem wir die Taylorentwicklung hinschreiben und alles ab dem quadratischen Term wegwerfen. Die Tangente ist diejenige Gerade, die den Graphen der Funktion f in der Nähe von x_0 bestmöglich annähert. Eine Funktion f ist genau dann im Punkt x_0 differenzierbar, wenn es dort eine eindeutig bestimmte Tangente gibt.

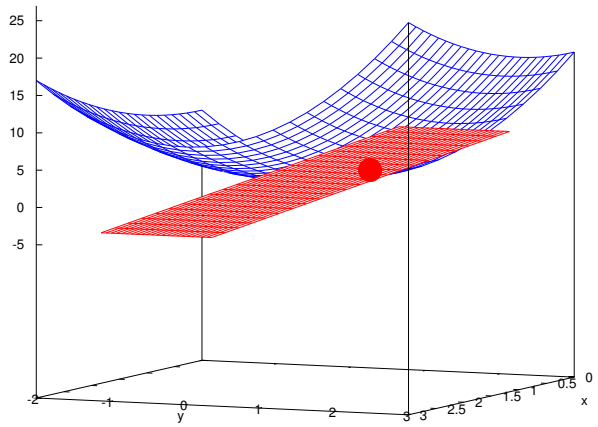


Abbildung 31: Die Tangentialebene an den Graph der Funktion $y = f(\vec{x})$, genommen im Punkt $[\vec{x}_0^T, f(\vec{x}_0)]^T$, hört auf die Gleichung $y = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$. Wir bekommen diese Ebenengleichung, indem wir die Taylorentwicklung hinschreiben und alles ab den quadratischen Termen wegwerfen. Die Tangentialebene ist diejenige Ebene, die den Graphen der Funktion f in der Nähe von $\vec{x}_0$ bestmöglich annähert. Eine Funktion f ist genau dann im Punkt $\vec{x}_0$ *total* differenzierbar, wenn es dort eine eindeutig bestimmte Tangentialebene gibt.

Die sprachlichen Parallelen in den Bildunterschriften sind didaktische Absicht!

18 Extrema

18.1 Extremwerte: Sei $M \subset \mathbb{R}^m$ eine (nicht unbedingt offene) Menge und $f\colon M \to \mathbb{R}$. Ein Punkt $\vec{x}_0 \in M$ heißt

- **globales Maximum** von f, wenn $f(\vec{x}) \leq f(\vec{x}_0)$ für jedes $\vec{x} \in M$ gilt,
- **globales Minimum** von f, wenn $f(\vec{x}) \geq f(\vec{x}_0)$ für jedes $\vec{x} \in M$ gilt,
- **lokales Maximum** von f, wenn es einen Radius $\delta > 0$ gibt mit $f(\vec{x}) \leq f(\vec{x}_0)$ für jedes $\vec{x} \in U_\delta(\vec{x}_0) \cap M$,
- **lokales Minimum** von f, wenn es einen Radius $\delta > 0$ gibt mit $f(\vec{x}) \geq f(\vec{x}_0)$ für jedes $\vec{x} \in U_\delta(\vec{x}_0) \cap M$,
- **globales Extremum** von f, wenn $\vec{x}_0$ ein globales Maximum oder ein globales Minimum von f ist,
- **lokales Extremum** von f, wenn $\vec{x}_0$ ein lokales Maximum oder ein lokales Minimum von f ist.

Wir **erinnern** an den Schulstoff: für Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ gilt:

$$\begin{array}{llll} f'(x_0) = 0 & \text{und} & f''(x_0) > 0 & \Longrightarrow \quad x_0 \text{ ist lokales Minimum,} \\ f'(x_0) = 0 & \text{und} & f''(x_0) < 0 & \Longrightarrow \quad x_0 \text{ ist lokales Maximum.} \end{array}$$

18.2 Notwendige Bedingung für Extrema: Sei $M \subset \mathbb{R}^m$, $\vec{x}_0$ ein innerer Punkt von M, und sei $f\colon M \to \mathbb{R}$ total differenzierbar in $\vec{x}_0$ mit lokalem Extremum in $\vec{x}_0$. Dann gilt $\nabla f(\vec{x}_0) = \vec{0}^T$.
Wir führen eine Sprechweise ein: wenn eine Funktion f in $\vec{x}_0$ total differenzierbar ist und wenn dort $\nabla f(\vec{x}_0) = \vec{0}^T$ gilt, dann nennen wir $\vec{x}_0$ einen *kritischen Punkt*.
Damit haben wir: jedes innere lokale Extremum einer total differenzierbaren Funktion ist ein kritischer Punkt.
Warnung: Wie im Eindimensionalen gilt: nicht jeder kritische Punkt ist eine Extremstelle! Man denke z.B. an die Funktion f mit der Gleichung $f(x, y) = x^3 - y^2$ (siehe Bild 26, links oben) oder an die „Affensattelfunktion“ $f(x, y) = x^3 - 3xy^2$ (siehe Bild 32). Kritische Punkte, die keine Extrema sind, nennt man *Sattelpunkte*.

18.3 Vorüberlegung für hinreichende Bedingungen: Sei $M \subset \mathbb{R}^m$ eine offene Menge, $f\colon M \to \mathbb{R}$ sei zweimal stetig partiell differenzierbar, und für $\vec{x}_0 \in M$ gelte $\nabla f(\vec{x}_0) = \vec{0}^T$. Der Satz von Taylor, vgl. 17.1, sagt uns dann: für jedes $\vec{x}$ in der Nähe von $\vec{x}_0$ gibt es einen Punkt $\vec{\xi}$ auf der Verbindungsstrecke von $\vec{x}$ und $\vec{x}_0$, sodass

$$f(\vec{x}) = f(\vec{x}_0) + \underbrace{\nabla f(\vec{x}_0)}_{=\vec{0}^T} \cdot (\vec{x} - \vec{x}_0) + \frac{1}{2}(\vec{x} - \vec{x}_0)^T \cdot H_f(\vec{\xi}) \cdot (\vec{x} - \vec{x}_0).$$

Also haben wir dann

$$f(\vec{x}) - f(\vec{x}_0) = \frac{1}{2}(\vec{x} - \vec{x}_0)^T \cdot H_f(\vec{\xi}) \cdot (\vec{x} - \vec{x}_0). \qquad (\spadesuit)$$

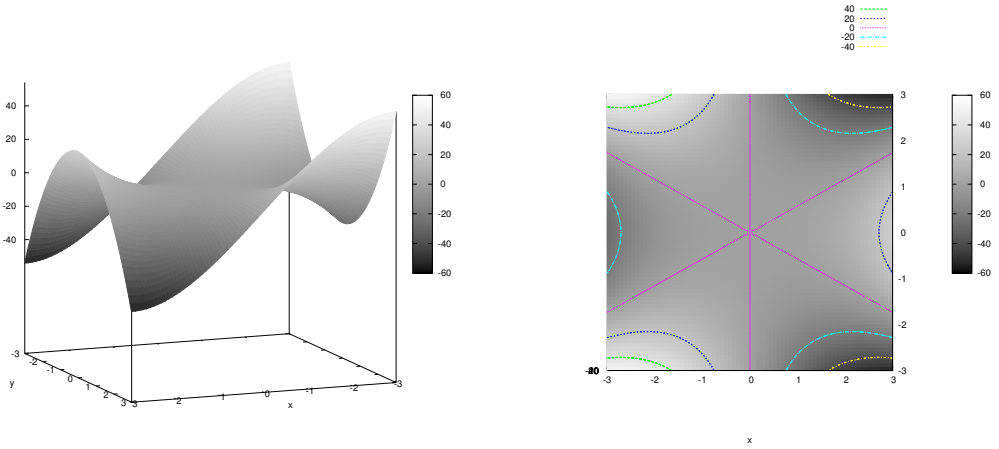


Abbildung 32: Links: die Affensattelfunktion $f(x,y) = x^3 - 3xy^2$. Wir sehen drei Richtungen, wo die Funktion „runtergeht“, und drei Richtungen, wo die Funktion „raufgeht“. Rechts der dazugehörige Höhenlinienplot.

Nach dem Satz von Schwarz ist die Hesse-Matrix $H_f(\vec{\xi}) \in \mathbb{R}^{n\times n}$ eine symmetrische Matrix. Nach Voraussetzung der zweimaligen stetigen Differenzierbarkeit ist $H_f(\vec{\xi})$ aber auch eine stetige Matrix-Funktion in der Vektorvariablen $\vec{\xi}$. Das bedeutet also $H_f(\vec{\xi}) \approx H_f(\vec{x}_0)$ für $\vec{\xi}$ in der Nähe von $\vec{x}_0$.
Wenn die rechte Seite von (♠) immer ≥ 0 ist für alle $\vec{\xi}$ in der Nähe von $\vec{x}_0$, dann ist auch die linke Seite immer positiv und damit $\vec{x}_0$ ein lokales Minimum von f. Und wenn die rechte Seite immer ≥ 0 ist für alle $\vec{\xi}$ in der Nähe von $\vec{x}_0$, dann ist genauso $\vec{x}_0$ ein lokales Maximum von f.
Damit sind die folgenden Begriffe motiviert:

18.4 Definite Matrizen: Sei $A \in \mathbb{R}^{m\times m}$ eine symmetrische Matrix. Dann heißt

- A **positiv definit**, wenn $\vec{x}^T \cdot A \cdot \vec{x} > 0$ für alle $\vec{x} \neq \vec{0}$,
- A **negativ definit**, wenn $\vec{x}^T \cdot A \cdot \vec{x} < 0$ für alle $\vec{x} \neq \vec{0}$,
- A **indefinit**, wenn es zwei Vektoren $\vec{x}$ und $\vec{y}$ aus $\mathbb{R}^m$ gibt mit $\vec{x}^T \cdot A \cdot \vec{x} > 0$ und $\vec{y}^T \cdot A \cdot \vec{y} < 0$.

18.5 Hinreichende Bedingung für Extrema: Sei $M \subset \mathbb{R}^m$ offen, $f\colon M \to \mathbb{R}$ sei zweimal stetig partiell differenzierbar, und $\vec{x}_0 \in M$ sei ein kritischer Punkt von f. Dann gilt:

Wenn $H_f(\vec{x}_0)$ positiv definit ist, dann ist $\vec{x}_0$ ein lokales Minimum von f.

Wenn $H_f(\vec{x}_0)$ negativ definit ist, dann ist $\vec{x}_0$ ein lokales Maximum von f.

Wenn $H_f(\vec{x}_0)$ indefinit ist, dann ist $\vec{x}_0$ ein Sattelpunkt von f.

Wie schon im Falle von $n = 1$ bekannt, gibt es auch kritische Punkte, die von diesem Kriterium nicht erfasst werden („unentscheidbare Fälle“). Das kann zum Beispiel

aussehen wie folgt:

$$\begin{aligned}&\text{Es ist } \vec{x}^T \cdot H_f(\vec{x}_0) \cdot \vec{x} \geq 0 \quad \text{für alle } \vec{x},\\ &\text{aber es gibt auch ein } \vec{y} \neq 0 \quad \text{mit} \quad \vec{y}^T \cdot H_f(\vec{x}_0) \cdot \vec{y} = 0.\end{aligned}$$

In diesem Fall nennt man die Hesse-Matrix $H_f(\vec{x}_0)$ *positiv semidefinit.*
Eine analoge Begriffsbildung gibt es mit ≤ 0: *negativ semidefinit.*
Das Affensattelbeispiel wird von unserem hinreichenden Kriterium übrigens auch nicht erfasst, weil dann für $\vec{x}_0 = [0,0]^T$ die Hesse-Matrix $H_f(\vec{x}_0)$ gleich der 2×2-Nullmatrix ist.

18.6 Praktischer Nachweis der Definitheit: Sei $A \in \mathbb{R}^{m\times m}$ symmetrisch. Dann gilt:

- A ist positiv definit $\iff$ alle Eigenwerte von A sind strikt positiv
  $\iff$ Die Unterdeterminanten
  $$\left|a_{11}\right|, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \ \ldots$$
  sind alle positiv.

  Diese Unterdeterminanten nennt man *Minoren.*

  (Bemerkung: Die Vorzeichenreihenfolge der Minoren kann man sich merken, indem man an die Einheitsmatrix E_n für A denkt.)

- A ist negativ definit $\iff$ alle Eigenwerte von A sind strikt negativ
  $\iff$ Die Minoren sind abwechselnd negativ, positiv, negativ, positiv usw.

  (Bemerkung: Die Vorzeichenreihenfolge der Minoren kann man sich merken, indem man an die Matrix $-E_n$ für A denkt, denn $-E_n$ ist negativ definit und hat die Minoren -1, $+1$, -1, $\ldots$)

- A ist indefinit $\iff$ es gibt einen strikt positiven und einen strikt negativen Eigenwert von A
  $\impliedby$ die Minoren sind alle $\neq 0$ und passen in keines der beiden obigen Schemata.

  (Achtung: in der letzten Zeile gilt wirklich nur $\impliedby$, nicht $\iff$.)

18.7 Beispiel:

- Die Matrix $A = \left[\begin{smallmatrix} 3 & -2 \\ -2 & 4 \end{smallmatrix}\right]$ ist symmetrisch. Die Minoren sind
  $$\det\begin{bmatrix}3\end{bmatrix} = 3 > 0, \qquad \det\begin{bmatrix} 3 & -2 \\ -2 & 4 \end{bmatrix} = 12 - 4 = 8 > 0,$$
  also ist die Matrix A positiv definit.

- Wir betrachten die Matrix
  $$B = \begin{bmatrix} -1 & -1 & 0 \\ -1 & -3 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

Diese ist symmetrisch, und die Minoren sind

$$\det\big[-1\big] = -1 < 0, \quad \det\begin{bmatrix}-1 & -1\\ -1 & -3\end{bmatrix} = 3 - 1 = 2 > 0,$$
$$\det\begin{bmatrix}-1 & -1 & 0\\ -1 & -3 & 1\\ 0 & 1 & -1\end{bmatrix} = \det\begin{bmatrix}-1 & -1 & 0\\ -1 & -2 & 1\\ 0 & 0 & -1\end{bmatrix} = -\det\begin{bmatrix}-1 & -1\\ -1 & -2\end{bmatrix} = -1 < 0.$$

Also ist B negativ definit.

- Wir betrachten die symmetrische Matrix $C = \left[\begin{smallmatrix}1 & 3\\ 3 & 2\end{smallmatrix}\right]$. Die Minoren sind

$$\det\big[1\big] = 1 > 0, \qquad \det\begin{bmatrix}1 & 3\\ 3 & 2\end{bmatrix} = 2 - 9 = -7 < 0.$$

Also ist C indefinit.

18.8 Beispiel: Für die Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$ mit $f(x,y) = (x^2 + 2y^2)e^{-x^2-y^2}$ wollen wir alle Extrema finden. Zuerst suchen wir die kritischen Punkte; also alle diejenigen $[x,y]^T$ mit $\nabla f(x,y) = [0,0]$. Dazu bestimmen wir den Gradienten von f:

$$\begin{aligned}\nabla f(x,y) &= \Big[2xe^{-x^2-y^2} + (x^2+2y^2)e^{-x^2-y^2}\cdot(-2x), \quad 4ye^{-x^2-y^2} + (x^2+2y^2)e^{-x^2-y^2}\cdot(-2y)\Big]\\ &= \Big[2xe^{-x^2-y^2}(1-x^2-2y^2), 2ye^{-x^2-y^2}(2-x^2-2y^2)\Big] \overset{!}{=} [0,0].\end{aligned}$$

Weil $2e^{-x^2-y^2}$ niemals Null werden kann, müssen kritische Punkte $[x,y]^T$ folgende beiden Gleichungen erfüllen:

$$x(1-x^2-2y^2) = 0, \tag{I}$$
$$y(2-x^2-2y^2) = 0. \tag{II}$$

Sei $x = 0$: dann ist (I) erfüllt, und aus (II) folgt $y(2-2y^2) = 0$, also $0 = y(1-y^2) = y(1-y)(1+y)$, also $y = 0$ oder $y = 1$ oder $y = -1$.

Sei $x \neq 0$: dann liefert (I) uns $(1-x^2-2y^2) = 0$, und aus (II) bekommen wir

$$0 = y\cdot(2-x^2-2y^2) = y\cdot(1+(1-x^2-2y^2)) = y\cdot(1+0) = y,$$

also $y = 0$. Damit gehen wir wieder zu (I) zurück und bekommen $(1-x^2) = 0$, also $x = 1$ oder $x = -1$.

Insgesamt finden wir die folgenden kritischen Stellen:

$$\begin{bmatrix}0\\0\end{bmatrix}, \quad \begin{bmatrix}0\\1\end{bmatrix}, \quad \begin{bmatrix}0\\-1\end{bmatrix}, \quad \begin{bmatrix}1\\0\end{bmatrix}, \quad \begin{bmatrix}-1\\0\end{bmatrix}.$$

Jetzt ermitteln wir den ersten Eintrag der Hesse-Matrix:

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2}(x,y) &= 2e^{-x^2-y^2}(1-x^2-2y^2) + 2xe^{-x^2-y^2}\cdot(-2x)\cdot(1-x^2-2y^2) + 2xe^{-x^2-y^2}\cdot(-2x)\\ &= 2e^{-x^2-y^2}\cdot\big(1-x^2-2y^2-2x^2(1-x^2-2y^2)-2x^2\big)\\ &= 2e^{-x^2-y^2}\left(1-x^2-2y^2-2x^2+2x^4+4x^2y^2-2x^2\right)\\ &= 2e^{-x^2-y^2}\left(1-5x^2-2y^2+2x^4+4x^2y^2\right).\end{aligned}$$

Auf ähnlichem Wege findet man

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = -4xye^{-x^2-y^2}\left(3-x^2-2y^2\right),$$
$$\frac{\partial^2 f}{\partial y^2}(x,y) = 2e^{-x^2-y^2}\left(2-x^2-10y^2+2x^2y^2+4y^4\right),$$

und somit ist

$$H_f(x,y) = 2e^{-x^2-y^2}\begin{bmatrix} 1-5x^2-2y^2+2x^4+4x^2y^2 & -2xy(3-x^2-2y^2) \\ -2xy(3-x^2-2y^2) & 2-x^2-10y^2+2x^2y^2+4y^4 \end{bmatrix}.$$

Wir setzen darin die kritischen Stellen ein:

$$H_f(0,0) = 2\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{mit Eigenwerten 2 und 4, also hat } f \text{ lokales Minimum in } \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

$$H_f(0,1) = \frac{2}{e}\begin{bmatrix} 1-2 & 0 \\ 0 & 2-10+4 \end{bmatrix} = \frac{2}{e}\begin{bmatrix} -1 & 0 \\ 0 & -4 \end{bmatrix}$$

$$\text{mit Eigenwerten } -\frac{2}{e} \text{ und } -\frac{8}{e}, \text{ also hat } f \text{ lokales Maximum in } \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$H_f(0,-1) = H_f(0,1), \quad \text{also ebenso lokales Maximum in } \begin{bmatrix} 0 \\ -1 \end{bmatrix},$$

$$H_f(1,0) = \frac{2}{e}\begin{bmatrix} 1-5+2 & 0 \\ 0 & 2-1 \end{bmatrix} = \frac{2}{e}\begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\text{mit Eigenwerten } -\frac{4}{e} \text{ und } \frac{2}{e}, \text{ also einem Sattelpunkt in } \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$H_f(-1,0) = H_f(1,0), \quad \text{also ebenso ein Sattelpunkt in } \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

18.9 Beispiel: (Extremwerte unter Nebenbedingungen)

- Wir suchen Extrema der Funktion $f(x,y) = x^3 - y^2$ auf der Menge

$$N = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \colon x^2 + y^2 = 1 \right\}.$$

Die Gleichung $x^2 + y^2 = 1$ nennen wir *Nebenbedingung*. Siehe Bild 34.

Zunächst beobachten wir: die Funktion f ist stetig und die Menge N ist kompakt. Also, vgl. 15.9, existieren tatsächlich Minima und Maxima der Funktion f auf N; wir müssen diese Extrema „nur noch“ finden. Bei dieser Aufgabe kommen wir mit **Einsetzen** zum Ziel: Für alle $[x,y]^T \in N$ ist $y^2 = 1 - x^2$, also ist dort

$$f(x,y) = x^3 - y^2 = x^3 - 1 + x^2 =: \varphi(x),$$

und jetzt suchen wir kritische Punkte von φ:

$$0 = \varphi'(x) = 3x^2 + 2x \qquad \Longleftrightarrow \qquad x = 0 \text{ oder } 3x + 2 = 0$$
$$\Longleftrightarrow \qquad x = 0 \text{ oder } x = -\frac{2}{3}.$$

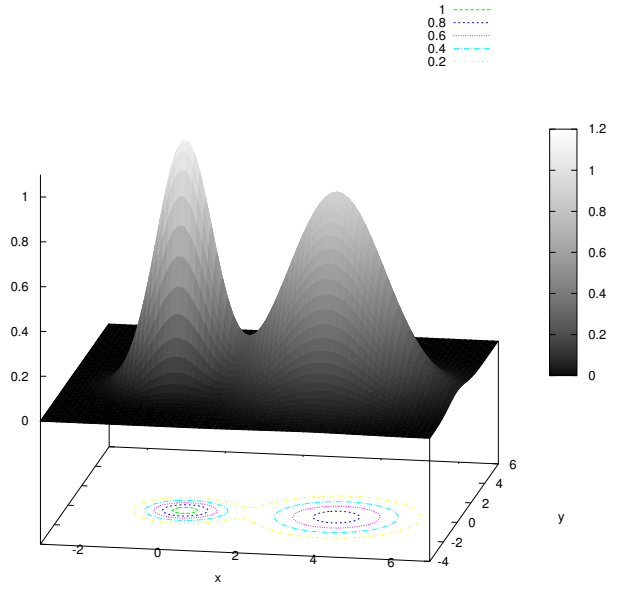


Abbildung 33: Achtung: diese Funktion $f\colon \mathbb{R}^2 \to \mathbb{R}$, gegeben durch die Funktionsgleichung $f(x,y) = 1.1e^{-x^2-y^2} + 0.9e^{-(x-3.9)^2/2-y^2/3}$ hat zwei Maxima, aber dazwischen kein Minimum (denn sie hat überhaupt kein Minimum auf dem $\mathbb{R}^2$). Funktionen im $\mathbb{R}^2$ können sich also anders verhalten als Funktionen im $\mathbb{R}^1$.

Wir setzen diese Werte in φ ein: es ist $\varphi(0) = -1$ und

$$\varphi\left(-\frac{2}{3}\right) = \left(-\frac{2}{3}\right)^3 - 1 + \left(-\frac{2}{3}\right)^2 = -\frac{8}{27} - 1 + \frac{4}{9} = \frac{-8-27+12}{27} = -\frac{23}{27}.$$

Die Maxima und Minima müssen wegen unserer Beobachtung existieren, also liegt das Maximum von f vor für $x = -\frac{2}{3}$; und das dazugehörige y ergibt sich dann aus der Gleichung $x^2 + y^2 = 1$, also $\frac{4}{9} + y^2 = 1$, d.h. $y = \pm\frac{\sqrt{5}}{3}$. Es hat also f Maxima in den Punkten

$$\begin{bmatrix} -\frac{2}{3} \\ \frac{\sqrt{5}}{3} \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} -\frac{2}{3} \\ -\frac{\sqrt{5}}{3} \end{bmatrix}.$$

Die Minima gibt es dann für $x = 0$, und das dazugehörige y ist gegeben über die Gleichung $x^2 + y^2 = 1$, also $0^2 + y^2 = 1$, d.h. $y = \pm 1$. Es hat also f Minima in den Punkten

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

- Jetzt suchen wir Extrema der Funktion $f(x,y) = x^3 - y^3$ auf derselben Menge N wie oben. Wir können die Nebenbedingung $x^2 + y^2 = 1$ in die Gestalt bringen $g(x,y) = 0$, mit $g(x,y) := x^2 + y^2 - 1$. Aber leider ist der Rechenweg von oben jetzt nicht mehr so schön durchführbar, wie man schnell bemerkt. Wir wollen uns eine andere Herangehensweise überlegen.

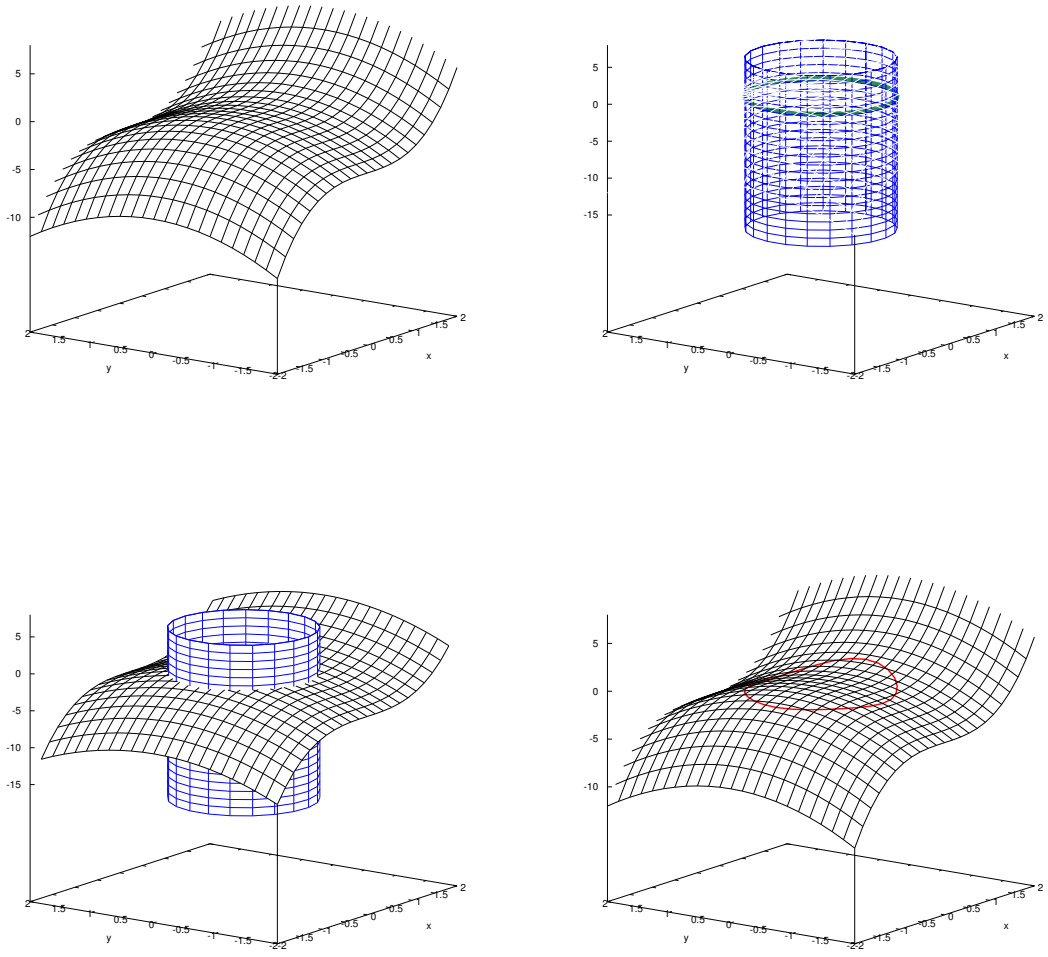


Abbildung 34: Links oben: die durch die Funktion $z = f(x, y) = x^3 - y^2$ beschriebene Fläche. Rechts oben: die Nebenbedingung $x^2 + y^2 = 1$ beschreibt in der xy-Ebene einen Kreis (grün). Wenn wir diesen entlang der z-Achse parallel verschieben, entsteht ein blauer Zylinder. Links unten: wir schneiden den blauen Zylinder mit der Fläche zur Funktion $z = f(x, y) = x^3 - y^2$. Rechts unten: als Schnittfigur entsteht eine rote Kurve. Von dieser Kurve suchen wir die lokal höchsten und tiefsten Punkte. Siehe Beispiel 18.9.

18.10 Problemstellung: Optimierung unter Nebenbedingungen: Wir wollen folgendes allgemeine **Problem** lösen: Seien $f\colon \mathbb{R}^m \to \mathbb{R}$ und $g\colon \mathbb{R}^m \to \mathbb{R}$ gegebene Funktionen. Wir suchen Extrema von f auf der Menge $N := \{\vec{x} \in \mathbb{R}^m \colon g(\vec{x}) = 0\}$. Die Gleichung $g(\vec{x}) = 0$ nennen wir *Nebenbedingung*.
Der Abbildung 35 und ihrem Begleittext entnehmen wir eine heuristische Überlegung, die uns hilft, die gesuchten Extrema zu finden. Wir machen uns plausibel, dass die gesuchten Extrema von f auf N nur an denjenigen Stellen überhaupt liegen können, an denen

- die Höhenlinien von f und g einander berühren, ohne einander zu schneiden,
- also die Tangenten an die Höhenlinien zueinander parallel sind.

Wir **erinnern** uns: es steht $\nabla f(\vec{x})$ senkrecht auf der Höhenlinie zu f durch den Punkt $\vec{x}$, und entsprechendes gilt für die Funktion g.
Das bedeutet: Extrema von f auf N kann es nur an den Stellen geben, an denen $\nabla f(\vec{x})$ und $\nabla g(\vec{x})$ parallel zueinander sind. Es muss also ein $\lambda \in \mathbb{R}$ geben mit $\nabla f(\vec{x}) = \lambda \nabla g(\vec{x})$.

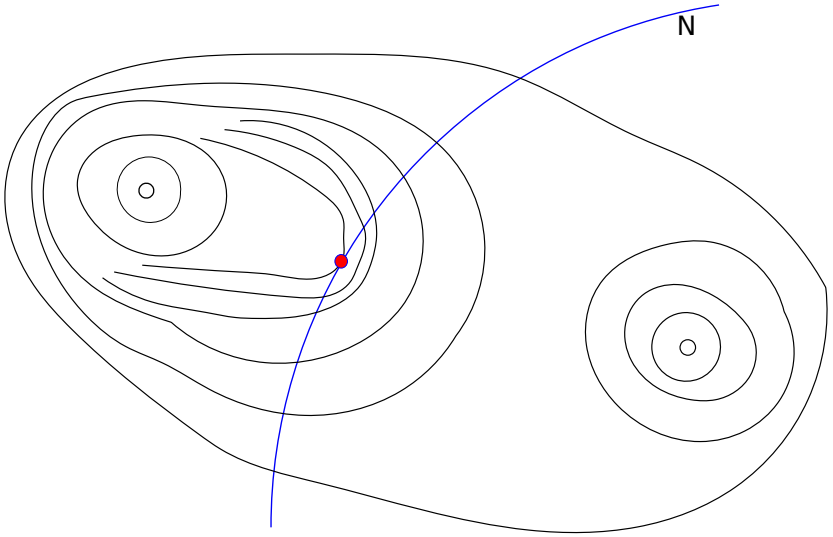


Abbildung 35: Eine handgemalte Skizze zur Suche nach Maxima der Funktion $f(\vec{x})$ für $\vec{x} \in \mathbb{R}^2$ unter der Nebenbedingung $g(\vec{x}) = 0$. Die Menge aller derjenigen $\vec{x} \in \mathbb{R}^2$ mit $g(\vec{x}) = 0$ ist gerade die Menge N (blaue Kurve), und dies ist die Höhenlinie der Funktion g zur Höhe Null. Die Höhenlinien der Funktion f sind schwarz skizziert, und in diesem Beispiel soll f genau zwei lokale Maxima besitzen. Der höchste Punkt von f auf der blauen Linie ist rot markiert. Diesen roten Punkt suchen wir. Am roten Punkt gilt: wenn wir die blaue Kurve nach „rechts oben“ entlanggehen, dann treffen wir dort nur niedere Niveaulinien von f. Und wenn wir vom roten Punkt aus die blaue Kurve nach „links unten“ entlanggehen, dann treffen wir ebenfalls nur niedere Niveaulinien von f. Also berührt die durch den roten Punkt laufende f-Niveaulinie die blaue Kurve N tangential.

18.11 Methode von Lagrange: Seien $M \subset \mathbb{R}^m$ offen, $f\colon M \to \mathbb{R}$ und $g\colon M \to \mathbb{R}$ beide stetig partiell differenzierbar, und sei $\nabla g(\vec{x}) \neq \vec{0}^T$ für alle $\vec{x} \in M$.
Dann gilt: falls f in $\vec{x}_0 \in M$ ein Extremum unter der Nebenbedingung $g = 0$ hat, so existiert ein $\lambda \in \mathbb{R}$ mit $\nabla f(\vec{x}_0) = \lambda \nabla g(\vec{x}_0)$.

18.12 Beispiel [→ 18.9]: Wir suchen Extrema von $f(x,y) = x^3 - y^3$ unter der Nebenbedingung $g(x,y) = x^2 + y^2 - 1 = 0$.
Die Lagrange-Bedingung lautet: es gibt ein $\lambda \in \mathbb{R}$ mit

$$\nabla f(x,y) = \lambda \nabla g(x,y) \quad \Longleftrightarrow \quad \begin{bmatrix} 3x^2 & -3y^2 \end{bmatrix} = \lambda \begin{bmatrix} 2x & 2y \end{bmatrix},$$

und das führt uns zum Gleichungssystem

$$\begin{aligned} 3x^2 &= 2\lambda x, \\ -3y^2 &= 2\lambda y, \\ x^2 + y^2 - 1 &= 0. \end{aligned}$$

(Beachte: dies sind zwar drei Gleichungen für die drei Unbekannten x, y und λ, aber damit ist über die Lösbarkeit dieses Gleichungssystems noch gar nichts gesagt.)
Dieses Gleichungssystem kann gelöst werden, und wir finden die Lösungen

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Dies sind alle Stellen, an denen es überhaupt Extrema geben könnte.
Allerdings ist f stetig auf der Menge

$$N = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \colon g(x,y) = 0 \right\} = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \colon x^2 + y^2 = 1 \right\},$$

und N ist die Einheitskreislinie im $\mathbb{R}^2$. Offenkundig ist N beschränkt und abgeschlossen, also kompakt. Nach 15.9 besitzt f auf N tatsächlich ein Maximum und ein Minimum. Wir brauchen also bloß noch die Werte von f an den extremwertverdächtigen Stellen ausrechnen und vergleichen:

$$\begin{aligned} f(1,0) = 1, \qquad f(-1,0) = -1, \qquad f(0,1) = -1, \qquad f(0,-1) = 1, \\ f\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}^3} + \frac{1}{\sqrt{2}^3} = \frac{2}{\sqrt{2}^3} = \frac{1}{\sqrt{2}}, \\ f\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = -\frac{1}{\sqrt{2}^3} - \frac{1}{\sqrt{2}^3} = -\frac{2}{\sqrt{2}^3} = -\frac{1}{\sqrt{2}}, \end{aligned}$$

und deshalb besitzt f auf der Menge N

$$\begin{aligned} &\text{Maxima in } \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -1 \end{bmatrix} \quad \text{mit Funktionswert } 1, \\ &\text{Minima in } \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{mit Funktionswert } -1. \end{aligned}$$

19 Parameterintegrale

19.1 Parameterintegral: Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ und $a, b \in \mathbb{R}$. Weiter sei für jedes $x \in \mathbb{R}$ die Abbildung $y \mapsto f(x,y)$ stückweise stetig. Dann definieren wir

$$g(x) := \int_{y=a}^{b} f(x,y)\,\mathrm{d}y$$

und nennen dies *Parameterintegral*. Die Schreibweise $\int_a^b f(x,y)\,\mathrm{d}y$ ist auch gebräuchlich.

19.2 Beispiel:

1. Wir setzen
$$g(x) := \int_{y=1}^{2} ye^{xy}\,\mathrm{d}y, \quad x \in \mathbb{R},$$
und fragen uns nach $\int_{x=c}^{d} g(x)\,\mathrm{d}x$ sowie $g'(x)$.

2. Sei die Funktion $u_0\colon \mathbb{R} \to \mathbb{R}$ stetig. Wir definieren die *Gaußglockenfunktion*
$$G(t,x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x|^2}{4t}}$$
und betrachten das uneigentliche Parameterintegral
$$u(t,x) := \int_{y=-\infty}^{\infty} G(t,x-y)u_0(y)\,\mathrm{d}y = \frac{1}{\sqrt{4\pi t}} \int_{y=-\infty}^{\infty} e^{-\frac{|x-y|^2}{4t}} u_0(y)\,\mathrm{d}y. \qquad (\spadesuit)$$
Hier haben wir sogar zwei Parameter, nämlich t und x.

Die **physikalische Bedeutung** ist: sei $u_0(x)$ die Temperatur zur Zeit $t = 0$ am Ort x, dann ist $u(t,x)$ die Temperatur zur Zeit t am Ort x bei einem ortsunabhängigen Wärmeleitungskoeffizienten. Das bedeutet, dass die Funktion $u(t,x)$ folgende Wärmeleitungsgleichung löst:
$$\frac{\partial u}{\partial t}(t,x) = \frac{\partial^2 u}{\partial x^2}(t,x), \qquad u(0,x) = u_0(x). \qquad (\heartsuit)$$
Man nennt die Funktion G auch *Fundamentallösung* der Wärmeleitungsgleichung. Um nachzuprüfen, dass unser u tatsächlich die Wärmeleitungsdifferentialgleichung ($\heartsuit$) löst, muss das Parameterintegral in ($\spadesuit$) einmal nach t bzw. zweimal nach x differenziert werden.

Wir haben also die Frage noch zu beantworten, wie man Parameterintegrale nach dem Parameter differenziert.

19.3 Integration von Parameterintegralen: Seien a, b, c, $d \in \mathbb{R}$. Wenn sowohl für jedes $x \in [c,d]$ die Funktionen $y \mapsto f(x,y)$ als auch für jedes $y \in [a,b]$ die Funktionen $x \mapsto f(x,y)$ stückweise stetig sind, so gilt

$$\int_{x=c}^{d} g(x)\,\mathrm{d}x = \int_{x=c}^{d} \int_{y=a}^{b} f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_{y=a}^{b} \int_{x=c}^{d} f(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

was bedeutet, dass die beiden Integrale getauscht werden dürfen.

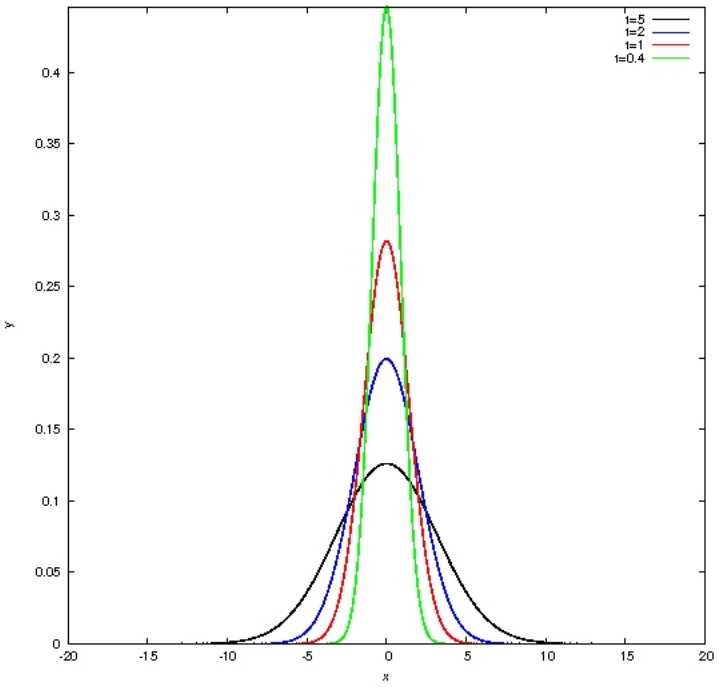


Abbildung 36: Die Gaußglockenfunktion $G(t,x)$ als Funktion von x, für verschiedene Werte von t. Der Flächeninhalt unter dem Graphen ist immer gleich Eins.

19.4 Beispiel: Wir greifen Beispiel 19.2 noch mal auf.

1. Es ist (zur Verdeutlichung setzen wir Extra–Klammern)

$$\begin{aligned}\int_{x=0}^{1} g(x)\,\mathrm{d}x &= \int_{x=0}^{1}\left(\int_{y=0}^{2} ye^{xy}\,\mathrm{d}y\right)\,\mathrm{d}x \qquad \Big| \quad \text{nutze (19.3)}\\ &= \int_{y=0}^{2}\left(\int_{x=0}^{1} ye^{xy}\,\mathrm{d}x\right)\,\mathrm{d}y = \int_{y=0}^{2} y\cdot\left(\int_{x=0}^{1} e^{xy}\,\mathrm{d}x\right)\,\mathrm{d}y\\ &= \int_{y=0}^{2} y\cdot\left(\frac{1}{y}e^{xy}\Big|_{x=0}^{x=1}\right)\,\mathrm{d}y = \int_{y=0}^{2} e^{xy}\Big|_{x=0}^{x=1}\mathrm{d}y = \int_{y=0}^{2}(e^y-1)\,\mathrm{d}y\\ &= (e^y-y)\Big|_{y=0}^{y=2} = e^2-2-1 = e^2-3.\end{aligned}$$

2. Wir betrachten die Funktion $u(t,x)$, die gegeben ist durch

$$u(t,x) = \frac{1}{\sqrt{4\pi t}}\int_{-\infty}^{\infty} e^{-\frac{|x-y|^2}{4t}}u_0(y)\,\mathrm{d}y.$$

Die Gesamt–Wärmemenge zur Zeit t berechnen wir dann wie folgt:

$$\begin{aligned}
\int_{x=-\infty}^{\infty} u(t,x)\,\mathrm{d}x &= \int_{x=-\infty}^{\infty} \left(\frac{1}{\sqrt{4\pi t}} \int_{y=-\infty}^{\infty} e^{-\frac{|x-y|^2}{4t}} u_0(y)\,\mathrm{d}y \right) \mathrm{d}x \qquad \Big| \quad \text{nutze (19.3)} \\
&= \int_{y=-\infty}^{\infty} \left(\int_{x=-\infty}^{\infty} \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x-y|^2}{4t}} u_0(y)\,\mathrm{d}x \right) \mathrm{d}y \\
&= \int_{y=-\infty}^{\infty} u_0(y) \left(\int_{x=-\infty}^{\infty} \frac{1}{\sqrt{4\pi t}} e^{-\frac{|x-y|^2}{4t}}\,\mathrm{d}x \right) \mathrm{d}y \qquad \Big| \quad \text{subst. } s = \frac{x-y}{\sqrt{2t}} \\
&= \int_{y=-\infty}^{\infty} u_0(y) \underbrace{\left(\int_{s=-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}}\,\mathrm{d}s \right)}_{=1 \quad \text{(Gauß–Glocke, s.u.)}} \mathrm{d}y \\
&= \int_{y=-\infty}^{\infty} u_0(y)\,\mathrm{d}y,
\end{aligned}$$

und das ist aber die Gesamtwärmemenge zur Zeit 0. Auf diesem Wege haben wir die Erhaltung der Energie durch Nachrechnen wiederentdeckt!

19.5 Differentiation von Parameterintegralen: Sei $f \in\colon \mathbb{R}^2 \to \mathbb{R}$ und seien a, $b \in \mathbb{R}$. Außerdem sei wieder für jedes $x \in \mathbb{R}$ die Funktion $y \mapsto f(x,y)$ stückweise stetig, sodass das Parameterintegral

$$g(x) = \int_{y=a}^{b} f(x,y)\,\mathrm{d}y$$

existiert. Wenn zusätzlich f nach x stetig partiell differenzierbar ist, dann ist auch die Funktion g differenzierbar, und es ist

$$g'(x) = \int_{y=a}^{b} \frac{\partial f}{\partial x}(x,y)\,\mathrm{d}y,$$

also dürfen Differentiation nach dem Parameter x und Integration bezüglich der Variablen y vertauscht werden.

19.6 Beispiel:

1. Bereits im letzten Beispiel verwendet wurde das wichtige Integral

$$\int_{-\infty}^{\infty} e^{-\frac{s^2}{2}} ds = \sqrt{2\pi}.$$

Da $e^{-\frac{x^2}{2}}$ keine elementare Stammfunktion (zusammengesetzt aus $e, \sin, \cos ...$) besitzt, kann das Integral **nicht** elementar gelöst werden. Man kann sich mit folgendem Trick behelfen: Setze

$$f(x) := \left(\int_0^x e^{-\frac{t^2}{2}} dt \right)^2, \quad g(x) := -2 \int_0^1 \frac{e^{-\frac{x^2}{2}(1+s^2)}}{1+s^2} ds.$$

Die Integranden sind stetig partiell differenzierbar. Wir erhalten als Ableitungen nach der Kettenregel:

$$f'(x) = 2 \left(\int_0^x e^{-\frac{t^2}{2}} dt \right) e^{-\frac{x^2}{2}},$$

wobei $(\int_0^x h(t)dt)' = h(x)$ verwendet wird, und

$$g'(x) := 2\int_0^1 \frac{x(1+s^2)e^{-\frac{x^2}{2}(1+s^2)}}{1+s^2}ds = 2\left(\int_0^1 e^{-\frac{x^2s^2}{2}}xds\right)e^{-\frac{x^2}{2}}$$

und nach Substitution $xs = t$ im Integral

$$g'(x) = 2\left(\int_0^x e^{-\frac{t^2}{2}}dt\right)e^{-\frac{x^2}{2}} = f'(x).$$

Dann ist $g(x) = f(x) + C$ mit einer Konstanten C. Berechne C: Speziell für $x = 0$ ist $f(0) = \int_0^0 \cdots = 0$ und

$$g(0) = -2\int_0^1 \frac{ds}{1+s^2} = -2\arctan s|_{s=0}^1 = -2(\frac{\pi}{4} - 0) = -\frac{\pi}{2}.$$

Also $C = g(0) - f(0) = -\frac{\pi}{2}$. Für große a ist $e^a \geq a$ bzw. $e^{-a} \leq \frac{1}{a}$. Für große x also

$$|g(x)| \leq 2\int_0^1 \frac{2}{x^2(1+s^2)^2}ds = \frac{1}{x^2}\int_0^1 \frac{4}{(1+s^2)^2}ds.$$

Das Integral ist eine Konstante, es folgt $\lim_{x\to\infty} g(x) = 0$. Dann ist $\lim_{x\to\infty} f(x) = \lim_{x\to\infty} g(x) - C = 0 - C = \frac{\pi}{2}$. Wenn wir uns erinnern, wie f definiert war, folgt

$$\int_0^\infty e^{\frac{-s^2}{2}}dt = \sqrt{\frac{\pi}{2}}$$

und daraus (aus Symmetriegründen)

$$\int_{-\infty}^\infty e^{-\frac{s^2}{2}}ds = 2\int_0^\infty e^{-\frac{s^2}{2}}ds = \sqrt{2\pi}.$$

2. Es sei

$$g(x) := \int_{y=1}^2 \frac{e^{xy} - e^y}{y}\,\mathrm{d}y.$$

Da $f(x,y) = \frac{e^{xy}-e^y}{y}$ stetig partiell nach x differenzierbar ist (beachte, dass wir wegen $1 \leq y \leq 2$ keine Gefahr laufen, durch Null zu dividieren), gilt für $x \neq 0$, dass

$$\begin{aligned} g'(x) &= \int_{y=1}^2 \frac{\partial}{\partial x}\left(\frac{e^{xy}-e^y}{y}\right)\mathrm{d}y = \int_{y=1}^2 \frac{1}{y}\cdot y\cdot e^{xy}\,\mathrm{d}y = \int_{y=1}^2 e^{xy}\,\mathrm{d}y \qquad \Big| \text{ benutze jetzt } x \neq 0 \\ &= \frac{1}{x}e^{xy}\Big|_{y=1}^{y=2} = \frac{1}{x}\left(e^{2x} - e^x\right). \end{aligned}$$

19.7 Variable Integrationsgrenzen: Jetzt sei

$$g(x) = \int_{y=x+1}^{y=x^2} \frac{e^{xy}-e^y}{y}\,\mathrm{d}y, \quad x > 0,$$

und wir suchen $g'(x)$. Wir beobachten, dass das Integral auf der rechten Seite an drei Stellen vom Parameter x abhängt:

- im Integranden,
- in der oberen Grenze,
- in der unteren Grenze.

Zur Ableitung der Funktion g benutzen wir eine Hilfsfunktion h, gegeben durch

$$h(x,u,v) := \int_{y=u}^{y=v} \frac{e^{xy}-e^y}{y}\,\mathrm{d}y,$$

und dann ist $g(x) = h(x, x+1, x^2)$. Gemäß Kettenregel gilt dann

$$g'(x) = \nabla h(x,x+1,x^2) \cdot \begin{bmatrix} (x)' \\ (x+1)' \\ (x^2)' \end{bmatrix} = \nabla h(x,x+1,x^2) \cdot \begin{bmatrix} 1 \\ 1 \\ 2x \end{bmatrix}.$$

Hierbei bedeutet $\nabla h(x, x+1, x^2)$, dass wir zuerst ∇ auf $h(x,u,v)$ anwenden, also

$$\nabla h(x,u,v) = \begin{bmatrix} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial u} & \frac{\partial h}{\partial v} \end{bmatrix}$$

ermitteln, und danach setzen wir $(x, x+1, x^2)$ ein für (x,u,v). Das machen wir jetzt. Wegen Beispiel 19.6 haben wir

$$\begin{aligned} \frac{\partial h}{\partial x}(x,u,v) &= \int_{y=u}^{v} \frac{\partial}{\partial x}\left(\frac{e^{xy}-e^y}{y}\right)\mathrm{d}y = \int_{y=u}^{v} \frac{1}{y}\cdot y \cdot e^{xy}\,\mathrm{d}y = \int_{y=u}^{v} e^{xy}\,\mathrm{d}y \\ &= \frac{1}{x}e^{xy}\Big|_{y=u}^{y=v} = \frac{e^{vx}-e^{ux}}{x}. \end{aligned}$$

Nach Hauptsatz der Differential– und Integralrechnung ist

$$\frac{\partial h}{\partial u}(x,u,v) = -\frac{e^{xu}-e^u}{u}, \qquad \frac{\partial h}{\partial v}(x,u,v) = +\frac{e^{xv}-e^v}{v}.$$

Und jetzt setzen wir anstatt (x,u,v) wieder $(x, x+1, x^2)$ ein:

$$\begin{aligned} g'(x) &= \nabla h(x,x+1,x^2) \cdot \begin{bmatrix} 1 \\ 1 \\ 2x \end{bmatrix} \\ &= \begin{bmatrix} \frac{e^{x^2\cdot x}-e^{(x+1)\cdot x}}{x} & -\frac{e^{x\cdot(x+1)}-e^{x+1}}{x+1} & \frac{e^{x\cdot x^2}-e^{x^2}}{x^2} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 2x \end{bmatrix} \\ &= \frac{1}{x}\left(e^{x^3}-e^{x^2}e^x\right) - \frac{1}{x+1}\left(e^{x^2}e^x - e^x e\right) + \frac{2}{x}\left(e^{x^3}-e^{x^2}\right) \\ &= \frac{1}{x}\left(3e^{x^3}-e^{x^2}(e^x+2)\right) - \frac{e^x}{x+1}\left(e^{x^2}-e\right). \end{aligned}$$

20 Kurvenintegrale

20.1 Motivation: Wenn ein Teilchen entlang einer geraden Strecke mit gleichförmiger Geschwindigkeit gezogen wird und wenn dabei gegen eine konstante Reibungskraft gearbeitet werden muss, dann ist bekanntlich die geleistete mechanische Arbeit gleich dem Produkt aus Kraft und Weglänge.
Jetzt wollen wir dies verallgemeinern, siehe Bild 37:

- der Weg braucht nicht mehr eine gerade Strecke sein — auch eine geschwungene Kurve ist jetzt erlaubt;
- die zu überwindende Kraft braucht nicht mehr konstant sein — jetzt wollen wir auch eine variable Kraft zulassen.

Anschaulich bedeutet dies: wir werden die Kraftfunktion F entlang einer Kurve integrieren.
Anschaulich ist auch klar: eine Kurve entspricht einem verbogenen Stück des $\mathbb{R}$, das man in den $\mathbb{R}^2$ oder den $\mathbb{R}^n$ hineingelegt hat.

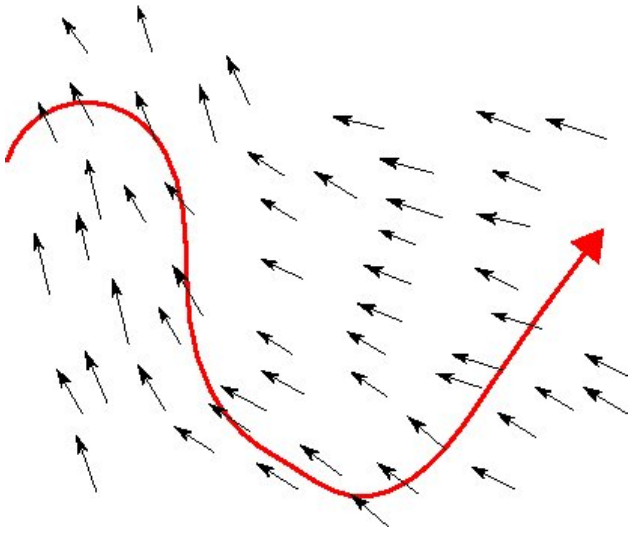


Abbildung 37: Ein Teilchen soll entlang des roten Wegs gezogen werden, und dabei ist die angegebene Kraft zu überwinden.

20.2 Kurven: Eine Funktion $\gamma\colon [a,b] \to \mathbb{R}^n$, die stetig und stückweise differenzierbar ist, heißt *Kurve im* $\mathbb{R}^n$.
Die Punktmenge

$$\operatorname{spur}(\gamma) := \{\gamma(t)\colon a \le t \le b\}$$

heißt *Spur der Kurve* (zuvor auch *Bild der Kurve* genannt), und $\gamma(a) \in \mathbb{R}^n$ heißt *Anfangspunkt der Kurve*. Entsprechend heißt $\gamma(b) \in \mathbb{R}^n$ *Endpunkt der Kurve*.
Man erhält eine gute Vorstellung, wenn man t als die Zeitvariable liest und $\gamma(t)$ als den Ort innerhalb des $\mathbb{R}^n$, an dem ein Teilchen zur Zeit t ist.
Was ist der Unterschied zwischen einer Kurve γ und $\operatorname{spur}(\gamma)$? Wir überlegen uns, dass jedes mathematische Objekt einen Typ hat (wir kennen schon die Typen „reelle Zahl“, „komplexe Zahl“, „Vektor“, „Matrix“ usw.). Die Kurve γ hat den Typ „Funktion von

$[a, b]$ in den $\mathbb{R}^n$“, und spur(γ) hat den Typ „Punktmenge des $\mathbb{R}^n$“. Wir stellen uns ein Auto vor, dass mit nassen Reifen über eine Straße fährt. Dann ist $[a, b]$ das Zeitintervall, und γ gibt an, wo das Auto wann ist. Im Gegensatz dazu ist spur(γ) die Spur der nassen Reifen auf dem Asphalt. Die Information, wann das Auto wo war, ist in spur(γ) nicht mehr enthalten.

20.3 Beispiel:

1. Seien $\vec{x}$, $\vec{y} \in \mathbb{R}^n$. Dann beschreibt die Funktion $\gamma\colon [0, 1] \to \mathbb{R}^n$, gegeben durch die Gleichung
$$\gamma(t) := (1 - t)\vec{x} + t\vec{y},$$
die *Verbindungsstrecke* von $\vec{x}$ nach $\vec{y}$. Das Teilchen beginnt seine Fahrt zur Zeit 0 im Punkt $\vec{x}$ und erreicht den Punkt $\vec{y}$ zur Zeit 1.

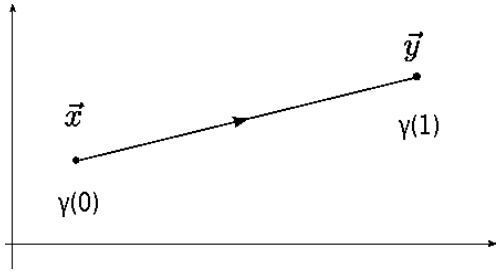


Abbildung 38: Eine Verbindungsstrecke.

2. Die Funktion
$$\gamma\colon [0, 2\pi] \to \mathbb{R}^2, \qquad \gamma(t) := \begin{bmatrix} r\cos(t) \\ r\sin(t) \end{bmatrix}$$
beschreibt einen im Gegenuhrzeigersinn durchlaufenen Kreis vom Radius r. Wir haben $\gamma(0) = [r, 0]^T = \gamma(2\pi)$.

   Eine Kurve $\gamma\colon [a, b] \to \mathbb{R}^n$ mit $\gamma(a) = \gamma(b)$ heißt *geschlossene Kurve*.

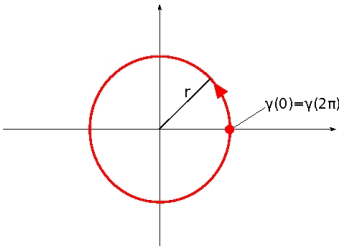


Abbildung 39: Ein Kreis vom Radius r.

3. Die Funktion
$$\gamma\colon [0, \pi] \to \mathbb{R}^2, \qquad \gamma(t) := \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}$$
beschreibt den Einheitshalbkreis in der oberen Halbebene, denn es ist $(\cos(t))^2 + (\sin(t))^2 = 1$ und $\sin(t) \geq 0$ für $t \in [0, \pi]$.

   Die Funktion
$$\hat{\gamma}(t)\colon [-1, 1] \to \mathbb{R}^2, \qquad \hat{\gamma}(t) := \begin{bmatrix} t \\ \sqrt{1 - t^2} \end{bmatrix}$$

beschreibt ebenfalls den Einheitshalbkreis in der oberen Halbebene, denn es ist $t^2 + (\sqrt{1-t^2})^2 = 1$ und $\sqrt{1-t^2} \geq 0$. Wir erkennen also

$$\operatorname{spur}(\gamma) = \operatorname{spur}(\hat{\gamma}),$$

aber die Kurven γ und $\hat{\gamma}$ sind sehr unterschiedlich.

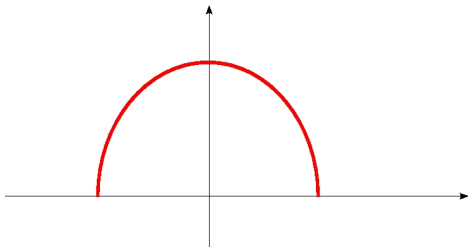


Abbildung 40: Der Einheitshalbkreis in der oberen Halbebene.

4. Die Funktion

$$\gamma\colon [0, 10\pi] \to \mathbb{R}^3, \qquad \gamma(t) := \begin{bmatrix} r\cos(t) \\ r\sin(t) \\ ct \end{bmatrix}$$

beschreibt eine Schraubenlinie im $\mathbb{R}^3$, siehe Bild 41.

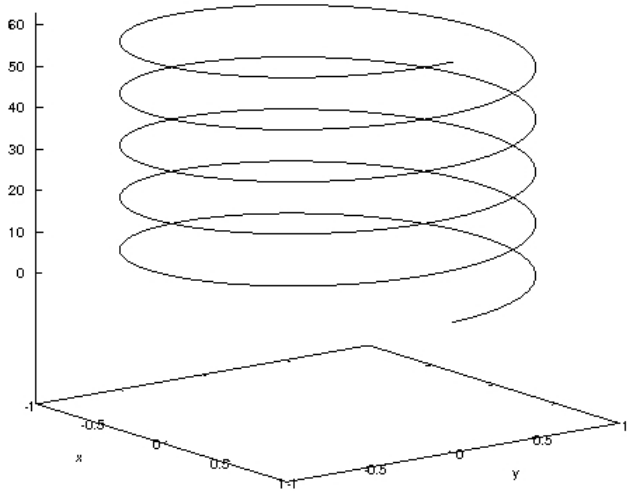


Abbildung 41: Eine Schraubenlinie mit $r = 1$ und $c = 2$.

20.4 Tangente an eine Kurve: Sei $\gamma\colon [a,b] \to \mathbb{R}^n$ eine Kurve und $t_0 \in (a,b)$ ein Zeitpunkt im Innneren des Intervalls $[a,b]$.
Die lineare Approximation im Punkt $\gamma(t_0)$ durch ein Taylorpolynom ist

$$\gamma(t_0) + D\gamma(t_0) \cdot (t - t_0) = \underbrace{\gamma(t_0)}_{\text{Aufpunkt}} + (t - t_0) \cdot \underbrace{D\gamma(t_0)}_{\text{Richtungsvektor}} .$$

Hierbei ist $D\gamma$ die totale Ableitung der Funktion γ. Anstatt $D\gamma(t_0)$ schreiben wir überlicherweise $\gamma'(t_0)$, und es ist

$$\gamma'(t_0) = \begin{bmatrix} \gamma_1'(t_0) \\ \gamma_2'(t_0) \\ \vdots \\ \gamma_n'(t_0) \end{bmatrix} .$$

Die **Anschauung** dazu ist: für das Teilchen, das auf spur(γ) entlang fliegt, ist zur Zeit t_0 der Geschwindigkeitsvektor gleich $\gamma'(t_0)$, und der Betrag der Geschwindigkeit ist $\|\gamma'(t_0)\|$. Der Geschwindigkeitsvektor verläuft tangential an spur(γ).

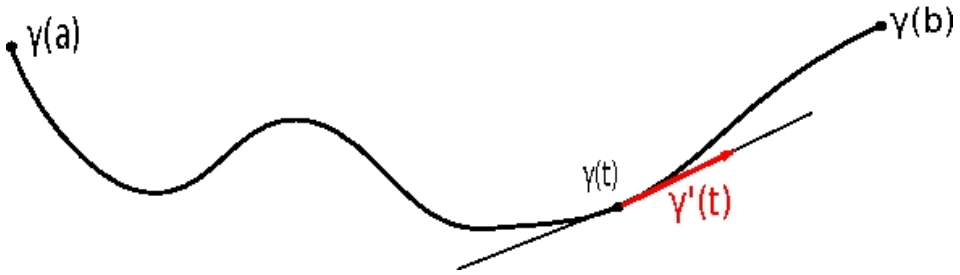


Abbildung 42: Die Spur einer Kurve mit einem Tangentenvektor $\gamma'(t)$ am Punkt $\gamma(t)$.

20.5 Beispiel:

1. Für die Verbindungsstrecke von $\vec{x}$ und $\vec{y} \in \mathbb{R}^n$ ist $\gamma(t) = (1-t)\vec{x} + t\vec{y}$. Dann folgt

$$\gamma'(t) = -\vec{x} + \vec{y} = \vec{y} - \vec{x}.$$

   Dieser Ausdruck hängt von der Zeit t gar nicht ab, es handelt sich also um eine gleichförmige lineare Bewegung. Innerhalb von einer Zeiteinheit wandert das Teilchen von $\vec{x}$ nach $\vec{y}$, also ist der Betrag der Geschwindigkeit tatsächlich $\frac{\|\vec{y}-\vec{x}\|}{1} = \|\gamma'(t)\|$, wie vorhergesagt (bei gleichförmigen linearen Bewegungen ergibt sich der Geschwindigkeitsbetrag tatsächlich durch Division von Gesamtweglänge durch Gesamtzeit).

2. Wir betrachten ein Teilchen auf dem Einheitskreis. Dann ist $\gamma\colon [0, 2\pi] \to \mathbb{R}^2$ mit

$$\gamma(t) = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}, \qquad \text{also} \qquad \gamma'(t) = \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix}.$$

   Die Länge des Tangentenvektors $\gamma'(t)$ ist dann

$$\|\gamma'(t)\| = \sqrt{\sin^2(t) + \cos^2(t)} = 1.$$

   Es handelt sich um eine gleichmäßige Kreisbewegung. Die Richtung, in die $\gamma'(t)$ zeigt, ist für jedes t eine andere.

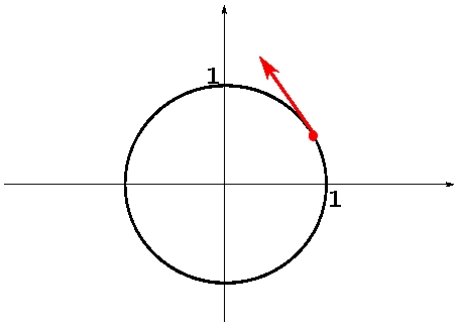


Abbildung 43: Der Einheitskreis mit einem Tangentenvektor.

20.6 Länge einer Kurve: Sei $\gamma\colon [a, b] \to \mathbb{R}^n$ eine Kurve. Die zurückgelegte Strecke ist das Integral des Geschwindigkeits**betrags** über der Zeit, also ist die Länge der Spur von γ gleich

$$L(\gamma) := \int_a^b \|\gamma'(t)\| \, \mathrm{d}t.$$

Achtung: wenn man von der Geschwindigkeit redet, dann meint man im Allgemeinen den Geschwindigkeitsvektor $\gamma'(t) \in \mathbb{R}^n$. Hier jedoch integrieren wir über den Geschwindigkeitsbetrag, welcher eine Zahl (≥ 0) ist.

20.7 Beispiel:

1. Für die Kreisbahn mit Radius r haben wir $\gamma\colon [0,2\pi] \to \mathbb{R}^2$ mit

$$\gamma(t) = \begin{bmatrix} r\cos(t) \\ r\sin(t) \end{bmatrix}, \quad r > 0.$$

Dann folgt

$$\begin{aligned} L(\gamma) &= \int_0^{2\pi} \|\gamma'(t)\| \, \mathrm{d}t = \int_0^{2\pi} \left\| \begin{bmatrix} -r\sin t \\ r\cos t \end{bmatrix} \right\| \mathrm{d}t = \int_0^{2\pi} \sqrt{r^2\sin^2(t) + r^2\cos^2(t)}\,\mathrm{d}t \\ &= \int_0^{2\pi} r\sqrt{\sin^2(t)+\cos^2(t)}\,\mathrm{d}t = \int_0^{2\pi} r\,\mathrm{d}t = 2\pi r. \end{aligned}$$

Das entspricht unseren Erwartungen.

2. Wir betrachten die Schraubenlinie auf dem Zeitintervall $[0,\ell]$. Dann ist $\gamma\colon [0,\ell] \to \mathbb{R}^3$ mit

$$\gamma(t) = \begin{bmatrix} r\cos(t) \\ r\sin(t) \\ ct \end{bmatrix}, \qquad \gamma'(t) = \begin{bmatrix} -r\sin(t) \\ r\cos(t) \\ c \end{bmatrix},$$

wobei wir natürlich ℓ, r, $c > 0$ haben. Es folgt dann

$$\begin{aligned} L(\gamma) &= \int_0^{\ell} \|\gamma'(t)\| \, \mathrm{d}t = \int_0^{\ell} \sqrt{r^2\sin^2(t) + r^2\cos^2(t) + c^2}\,\mathrm{d}t \\ &= \int_0^{\ell} \sqrt{r^2+c^2}\,\mathrm{d}t = \ell\cdot\sqrt{r^2+c^2}. \end{aligned}$$

Dies hätten wir auch herausfinden können, indem wir den Zylindermantel, auf dem die Schraubenlinie aufgewickelt ist, verzerrungsfrei abwickeln.

20.8 Arbeitsintegral: Sei $M \subseteq \mathbb{R}^n$, $F\colon M \to \mathbb{R}^n$ ein stückweise stetiges Vektorfeld, und sei $\gamma\colon [a,b] \to M$ eine Kurve. Dann heißt

$$\int_\gamma F(\vec{x})\,\mathrm{d}\vec{x} := \int_a^b \langle F(\gamma(t)), \gamma'(t)\rangle \, \mathrm{d}t$$

das *Arbeitsintegral von F längs* $\mathrm{spur}(\gamma)$ oder auch *Kurvenintegral von F längs* $\mathrm{spur}(\gamma)$. Wir stellen uns f als Kraftfeld vor und $\mathrm{spur}(\gamma)$ als den Weg des Teilchens. Siehe auch Bild 37.

20.9 Bemerkungen:

1. Es ist auch die Notation

$$\int_a^b F(\gamma(t)) \cdot \gamma'(t)\,\mathrm{d}t$$

gebräuchlich. Hierbei bezeichnet der Multiplikationspunkt $\cdot$ das Skalarprodukt zweier Vektoren mit n Einträgen.

2. Der Wert $\int_\gamma F(\vec{x})\,\mathrm{d}\vec{x}$ ändert sich nicht, wenn wir spur(γ) mit anderer Beschleunigung aber in derselben Richtung durchlaufen.

   Mathematisch genau: ist $\psi : [c : d] \to [a, b]$ stetig differenzierbar mit $\psi(c) = a$, $\psi(d) = b$ und $\psi'(s) > 0$ für $s \in (c, d)$, dann ist $\hat{\gamma} = \gamma \circ \psi : [c : d] \to \mathbb{R}^n$ eine Kurve mit spur(γ) = spur($\hat{\gamma}$) und $\int_\gamma F(\vec{x})\,\mathrm{d}\vec{x} = \int_{\hat{\gamma}} F(\vec{x})\,\mathrm{d}\vec{x}$. Konkret z.B. $\gamma(t)$ und $\gamma(t^5)$, wenn $\gamma : [-1, 1] \to \mathbb{R}^n$.

3. Dagegen gilt $\int_\gamma F(\vec{x})\,\mathrm{d}\vec{x} = -\int_{\hat{\gamma}} F(\vec{x})\,\mathrm{d}\vec{x}$, wenn spur($\gamma$) mit irgendeiner Beschleunigung in anderer Richtung durchlaufen wird.

   Mathematisch genau: ist $\psi : [c : d] \to [a, b]$ stetig differenzierbar mit $\psi(c) = b$, $\psi(d) = a$ und $\psi'(s) < 0$ für $s \in (c, d)$, dann ist $\hat{\gamma} = \gamma \circ \psi : [c : d] \to \mathbb{R}^n$ eine Kurve mit spur(γ) = spur($\hat{\gamma}$) und $\int_\gamma F(\vec{x})\,\mathrm{d}\vec{x} = -\int_{\hat{\gamma}} F(\vec{x})\,\mathrm{d}\vec{x}$. Konkret z.B. $\gamma(t)$ und $\gamma(-t)$, wenn $\gamma : [-1, 1] \to \mathbb{R}^n$.

Wir kommen zu einem neuen Gedanken.

20.10 Beispiel: Wir integrieren ein Vektorfeld F entlang zweier unterschiedlicher Kurven γ und $\hat{\gamma}$, für die auch deren Spuren unterschiedlich sind. Siehe Bild 44.

1. Sei
$$F(x, y) := \begin{bmatrix} 2xy \\ x^2 + y^2 \end{bmatrix}, \qquad \gamma(t) := \begin{bmatrix} t\cos(t) \\ t\sin(t) \end{bmatrix}, \qquad t \in \left[0, \frac{\pi}{2}\right].$$
   Dann haben wir mit $\cos^2 t + \sin^2 t = 1$
$$\begin{aligned}
\int_\gamma F(\vec{x})\,\mathrm{d}x &= \int_0^{\pi/2} \langle F(\gamma(t)), \gamma'(t)\rangle\ \mathrm{d}t \\
&= \int_0^{\pi/2} \left\langle \begin{bmatrix} 2t^2\cos(t)\sin(t) \\ t^2\cos^2(t) + t^2\sin^2(t) \end{bmatrix}, \begin{bmatrix} \cos(t) - t\sin(t) \\ \sin(t) + t\cos(t) \end{bmatrix} \right\rangle \mathrm{d}t \\
&= \int_0^{\pi/2} \left(2t^2\cos^2(t)\sin(t) - 2t^3\cos(t)\sin^2(t) + t^2\sin(t) + t^3\cos(t)\right)\ \mathrm{d}t \\
&= (\text{man rechnet einige Zeit}) = \frac{\pi^3}{24}.
\end{aligned}$$

2. Wir nehmen dieselbe Funktion F, jetzt allerdings die Kurve $\hat{\gamma}$ mit $\hat{\gamma}(t) := [0, t]^T$ für $t \in [0, \frac{\pi}{2}]$. Wir beobachten $\hat{\gamma}(0) = [0, 0]^T = \gamma(0)$ und $\hat{\gamma}(\frac{\pi}{2}) = [0, \frac{\pi}{2}]^T = \gamma(\frac{\pi}{2})$. Es haben also spur($\hat{\gamma}$) und spur($\gamma$) dieselben Anfangs- und Endpunkte. Dann rechnen wir wie folgt:
$$\int_{\hat{\gamma}} F(\vec{x})\,\mathrm{d}\vec{x} = \int_0^{\pi/2} \langle F(\hat{\gamma}(t)), \hat{\gamma}'(t)\rangle\ \mathrm{d}t = \int_0^{\pi/2} \left\langle \begin{bmatrix} 0 \\ t^2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\rangle \mathrm{d}t = \int_0^{\pi/2} t^2\,\mathrm{d}t = \frac{\pi^3}{24}.$$

Die Gleichheit der Integrale ist kein Zufall.

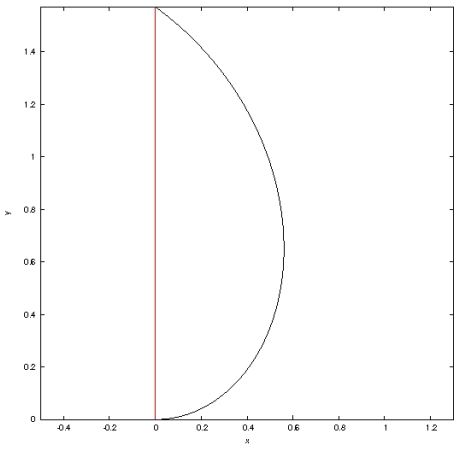


Abbildung 44: spur(γ) ist der schwarze Bogen, spur($\hat{\gamma}$) die rote Strecke.

20.11 Potential: Sei $M \subseteq \mathbb{R}^n$ und $F\colon M \to \mathbb{R}^n$ ein Vektorfeld. Dann heißt ein Skalarfeld $\varphi : M \to \mathbb{R}$ *Potential von* f, falls gilt:

- φ ist stetig partiell differenzierbar,
- $\nabla\varphi(\vec{x}) = F(\vec{x})^T$ für jedes $\vec{x} \in M$.

In diesem Falle nennen wir F ein *Potentialfeld* oder ein *konservatives Feld* oder ein *Gradientenfeld*.

20.12 Kurvenintegrale von Potentialfeldern: Sei $M \subseteq \mathbb{R}^n$ und $F : M \to \mathbb{R}^n$ ein Potentialfeld mit Potential $\varphi : M \to \mathbb{R}$, das heißt $F^T = \nabla\varphi$. Dann gilt für jede Kurve $\gamma : [a, b] \to M$, dass

$$\Big(\varphi(\gamma(t))\Big)' \overset{\text{Kettenregel}}{=} \nabla\varphi(\gamma(t)) \cdot \gamma'(t) = F(\gamma(t))^T \cdot \gamma'(t) = \langle F(\gamma(t)), \gamma'(t)\rangle\,,$$

also ergibt sich dann

$$\int_\gamma F(\vec{x})\,\mathrm{d}\vec{x} = \int_a^b \langle F(\gamma(t)), \gamma'(t)\rangle\;\mathrm{d}t = \int_a^b \Big(\varphi(\gamma(t))\Big)'\,\mathrm{d}t \overset{\text{Hauptsatz}}{=} \varphi(\gamma(b)) - \varphi(\gamma(a)).$$

Das bedeutet: Wenn F ein Potentialfeld ist, dann hängt der Wert des Kurvenintegrals nur von der Lage der Anfangs– und Endpunkte der Kurve ab, aber nicht vom Verlauf der Spur dazwischen!
Man sagt auch: das Potential φ ist die *Stammfunktion* des Potentialfelds F.
Ein **wichtiger Spezialfall:** wenn F ein Potentialfeld ist und γ geschlossen, so ist

$$\int_\gamma F(\vec{x})\,\mathrm{d}\vec{x} = 0.$$

20.13 Beispiel:

1. Wir kehren zurück zu Beispiel 20.10. Dort war

$$F(x, y) = \begin{bmatrix} 2xy \\ x^2 + y^2 \end{bmatrix},$$

und dies ist ein Potentialfeld, denn für die (im Moment noch erratene) Funktion $\varphi(x,y) = x^2y + \frac{1}{3}y^3$ gilt tatsächlich

$$\nabla\varphi(x,y) = [2xy, x^2+y^2] = F(x,y)^T.$$

Dies erklärt, warum die beiden Integrale gleich waren.

2. Auf der Menge $M := \mathbb{R}^2 \setminus \{\vec{0}\}$ betrachten wir die Funktion

$$F(x,y) = \begin{bmatrix} \frac{-y}{x^2+y^2} \\ \frac{x}{x^2+y^2} \end{bmatrix}.$$

Wir haben bei der Wahl von M den Nullpunkt ausgeschlossen, weil wir ansonsten durch Null dividiert hätten. Die Spur der Kurve sei der im Gegenuhrzeigersinn durchlaufene Einheitskreis, also

$$\gamma(t) = \begin{bmatrix} \cos(t) \\ \sin(t) \end{bmatrix}, \qquad t \in [0, 2\pi].$$

Es ist also γ eine geschlossene Kurve. Trotzdem ist unser Kurvenintegral nicht gleich Null, denn

$$\begin{aligned} \int_\gamma F(\vec{x})\,\mathrm{d}\vec{x} &= \int_0^{2\pi} \langle F(\gamma(t)), \gamma'(t)\rangle \,\mathrm{d}t \\ &= \int_0^{2\pi} \begin{bmatrix} \frac{-\sin(t)}{\cos^2(t)+\sin^2(t)} & \frac{\cos(t)}{\cos^2(t)+\sin^2(t)} \end{bmatrix} \cdot \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} \mathrm{d}t \\ &= \int_0^{2\pi} \Big(\sin^2(t) + \cos^2(t) \Big)\,\mathrm{d}t = 2\pi \neq 0 \ ! \end{aligned}$$

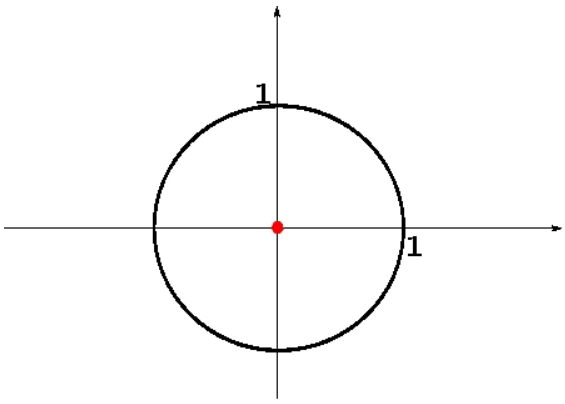


Abbildung 45: Der Integrationsweg für Beispiel 20.13 (b).

20.14 Notwendige Bedingung für die Existenz eines Potentials: Sei $M \subseteq \mathbb{R}^n$ und $F = [f_1, \ldots, f_n]^T \colon M \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld. Wenn F ein Potential φ besitzt, dann ist für jegliche $i, j = 1, \ldots, n$

$$f_i = \frac{\partial\varphi}{\partial x_i}, \qquad f_j = \frac{\partial\varphi}{\partial x_j}.$$

Der Satz von Schwarz liefert dann (falls $i \neq j$)

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial^2 \varphi}{\partial x_i \partial x_j} = \frac{\partial^2 \varphi}{\partial x_j \partial x_i} = \frac{\partial f_j}{\partial x_i}.$$

Eine **Mindestvoraussetzung** für die Existenz eines Potentials ist also die **Integrabilitätsbedingung**

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}, \qquad \text{für alle } i \neq j, \qquad \text{bzw. die Jacobi–Matrix } \quad J_f \quad \text{ist symmetrisch.}$$

Umgangssprachlich: „die Ableitungen über Kreuz müssen gleich sein“!

20.15 Beispiel:

1. Wir kehren zu Beispiel 20.10 zurück. Da hatten wir

   $$F(x,y) = \begin{bmatrix} 2xy \\ x^2 + y^2 \end{bmatrix} = \begin{bmatrix} f_1(x,y) \\ f_2(x,y) \end{bmatrix}.$$

   Wir prüfen die Integrabilitätsbedingung nach:

   $$\frac{\partial f_1}{\partial y} \stackrel{?}{=} \frac{\partial f_2}{\partial x},$$

   und dies stimmt tatsächlich, denn es ist $2x = 2x$. Also ergibt es Sinn, ein Potential φ zu suchen (von dem wir aber noch nicht wissen, ob es überhaupt existiert). Es soll gelten $\frac{\partial \varphi}{\partial x} = 2xy$, also folgt

   $$\varphi(x,y) = \int 2xy \, \mathrm{d}x + g(y) = x^2 y + c + g(y).$$

   Hierbei sind c eine wählbare Konstante und g eine (noch) wählbare Funktion. Dann haben wir tatsächlich $\frac{\partial \varphi}{\partial x} = 2xy$, und wir wollen haben

   $$x^2 + y^2 \stackrel{!}{=} \frac{\partial \varphi}{\partial y} = x^2 + 0 + g'(y).$$

   Also muss $g'(y) = y^2$ sein, d.h. $g(y) = \frac{1}{3}y^3 + C$. Zusammengefasst bekommen wir z.B.

   $$\varphi(x,y) = x^2 y + \frac{1}{3}y^3,$$

   wenn wir $c = C = 0$ setzen.

2. Potentiale sind etwas Seltenes und Kostbares! Sei z.B.

   $$F(x,y) = \begin{bmatrix} x \\ x + y \end{bmatrix}.$$

   Dann ist $\frac{\partial f_1}{\partial y} = 0 \neq 1 = \frac{\partial f_2}{\partial x}$, also kann diese harmlos ausschauende Funktion gar kein Potential besitzen.

3. Wir setzen Beispiel 20.13 (b) fort. Es war

$$F(x,y) = \begin{bmatrix} \frac{-y}{x^2+y^2} \\ \frac{x}{x^2+y^2} \end{bmatrix} = \begin{bmatrix} f_1(x,y) \\ f_2(x,y) \end{bmatrix}.$$

Wir testen, ob die Integrabilitätsbedingung erfüllt ist:

$$\frac{\partial f_1}{\partial y} = \frac{-(x^2+y^2)+2y^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2},$$
$$\frac{\partial f_2}{\partial x} = \frac{x^2+y^2-2x^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2},$$

und dieser Test ist erfolgreich. Also suchen wir ein Potential φ auf ähnlichem Rechenweg wie eben, und es ensteht

$$\varphi(x,y) = \arctan\left(\frac{y}{x}\right) \quad \text{für } x \neq 0.$$

Wir haben jetzt ein Potential φ, aber trotzdem haben wir $\int_\gamma F(\vec{x})\,\mathrm{d}\vec{x} \neq 0$ ausgerechnet. Der Grund ist, dass der Einheitskreis, den wir beim Integrieren entlang γ ablaufen, zwei Punkte besitzt mit $x = 0$, und dort haben wir leider kein Potential.

Der Integrand F besitzt im Ursprung $[0,0]^T$ eine Polstelle (auch Singularität genannt), und diese Singularität „strahlt aus“.

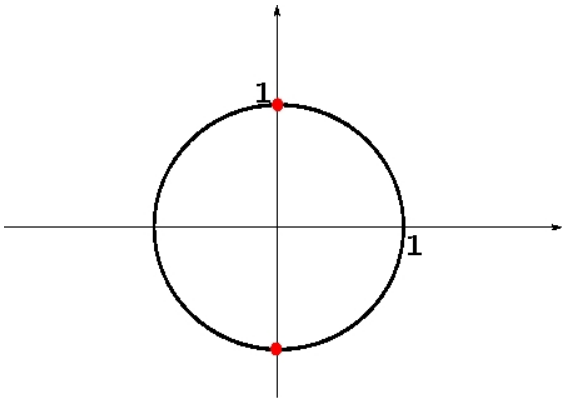


Abbildung 46: Der Integrationsweg mit zwei Stellen ohne Potential.

Wir erkennen: die Integrabilitätsbedingung ist zwar notwendig dafür, dass das Vektorfeld F ein Potential besitzt überall auf $\mathrm{spur}(\gamma)$, aber nicht hinreichend. Eine Abhilfe geben wir im nächsten Satz.

20.16 Potential im Fall M sternförmig: Sei $M \subseteq \mathbb{R}^n$ und das Vektorfeld $F\colon M \to \mathbb{R}^n$ erfülle die Integrabilitätsbedingung. Wenn die Menge M offen ist und *sternförmig* ist, dann besitzt F auf ganz M ein Potential.
Hierbei bedeutet „sternförmig“, dass es einen Punkt $\vec{p} \in M$ gibt, so dass die gerade Verbindungsstrecke $t\vec{p} + (1-t)\vec{x}$, $t \in [0,1]$, zu jedem andere Punkt $\vec{x} \in M$, vollständig in M verläuft.

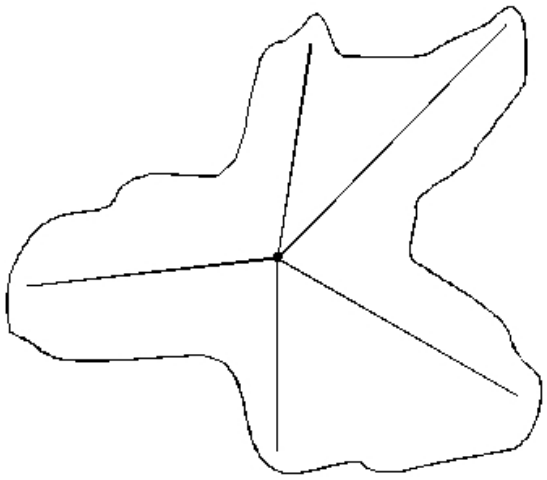


Abbildung 47: Sternförmig: Vom Punkt in der Mitte von M kann jeder andere Punkt in M durch ein Geradenstück erreicht werden, das vollständig in M liegt. Nicht sternförmig ist beispielsweise die offene Kreisscheibe $K^\circ = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 < 1\} \backslash (0, 0)$ ohne Mittelpunkt. Egal welchen Punkt $\vec{p} \in K^\circ$ wir wählen, die Verbindungsstrecke zu $-\vec{p}$ verläuft nicht in K°

21 Integrale im $\mathbb{R}^n$

Unser **Ziel** ist die Integration von Funktionen in mehreren Variablen. Ein **Problem** kann dabei darin liegen, dass die Integrationsgebiete (also diejenigen Mengen, in denen die Integrationsvariable $\vec{x}$ lebt) relativ wild sein können.

21.1 Intervalle im $\mathbb{R}^n$: Seien $\vec{a}$, $\vec{b} \in \mathbb{R}^n$ mit $a_i \leq b_i$ für alle $i = 1, \ldots, n$. Dann heißt

$$I := [\vec{a}, \vec{b}] := \{\vec{x} \in \mathbb{R}^n \colon a_i \leq x_i \leq b_i \text{ für alle } i = 1, \ldots, n\}$$

abgeschlossenes Intervall im $\mathbb{R}^n$.
Dann ist $I = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n]$.

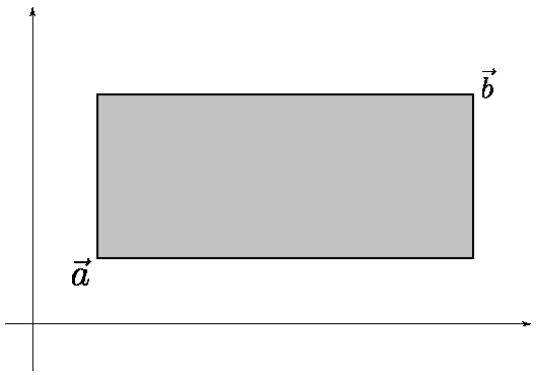


Abbildung 48: Ein Intervall im $\mathbb{R}^2$.

21.2 Integrale über Intervalle: Sei $I = [\vec{a}, \vec{b}]$ ein Intervall im $\mathbb{R}^n$, und sei $f \colon I \to \mathbb{R}$ stückweise stetig. Dann legen wir fest

$$\int_I f(\vec{x})\,\mathrm{d}\vec{x} := \int_{a_1}^{b_1} \int_{a_2}^{b_2} \ldots \int_{a_n}^{b_n} f(x_1, x_2, \ldots, x_n)\,\mathrm{d}x_n \ldots \mathrm{d}x_2\,\mathrm{d}x_1.$$

Hierbei ist die Integrationsreihenfolge beliebig, also z.B. für $n = 2$:

$$\int_I f(x,y)\,\mathrm{d}(x,y) = \int_a^b \int_c^d f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_c^d \int_a^b f(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

wenn $I = [a, b] \times [c, d]$.

21.3 Beispiel:

1. Wenn wir die Funktion $f(\vec{x}) = 1$ wählen, dann liefert das Integral uns das Volumen des Intervalls I:

$$\begin{aligned}\int_I 1\,\mathrm{d}\vec{x} &= \int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} 1\,\mathrm{d}x_n \ldots \mathrm{d}x_1 = \int_{a_1}^{b_1} \ldots \int_{a_{n-1}}^{b_{n-1}} (b_n - a_n)\,\mathrm{d}x_{n-1} \ldots \mathrm{d}x_1 \\ &= (b_1 - a_1) \cdot (b_2 - a_2) \cdot \ldots \cdot (b_n - a_n).\end{aligned}$$

2. Ein schöner Spezialfall ergibt sich, wenn die Funktion f die Gestalt $f(x,y) = g(x)\cdot h(y)$ hat. Dann entsteht

$$\int_a^b \int_c^d f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_a^b \int_c^d g(x)h(y)\,\mathrm{d}y\,\mathrm{d}x = \int_a^b g(x)\left(\int_c^d h(y)\,\mathrm{d}y\right)\mathrm{d}x$$
$$= \left(\int_a^b g(x)\,\mathrm{d}x\right)\cdot\left(\int_c^d h(y)\,\mathrm{d}y\right).$$

3. Sei $f(x,y) = 2x\cos(x^2-y)$, und das Intervall sei $I = [0,1]\times[-1,1]$. Dann können wir einerseits substituieren wie folgt (wobei wir darauf achten, die Integrationsgrenzen mitzusubstituieren):

$$\int_{y=-1}^{1}\int_{x=0}^{1} 2x\cos(x^2-y)\,\mathrm{d}x\,\mathrm{d}y \qquad \Big|\ \text{substituiere } u = x^2,\quad \frac{\mathrm{d}u}{\mathrm{d}x} = 2x$$
$$= \int_{y=-1}^{1}\int_{u=0}^{1}\cos(u-y)\,\mathrm{d}u\,\mathrm{d}y = \int_{y=-1}^{1}\sin(u-y)\Big|_{u=0}^{u=1}\mathrm{d}y$$
$$= \int_{y=-1}^{1}\Big(\sin(1-y)-\sin(-y)\Big)\,\mathrm{d}y = \Big(\cos(1-y)-\cos(-y)\Big)\Big|_{y=-1}^{y=1}$$
$$= \cos(0)-\cos(2)-\cos(-1)+\cos(1) = 1-\cos(2)-\cos(1)+\cos(1) = 1-\cos(2).$$

Andererseits können wir aber auch die Integrale tauschen:

$$\int_{x=0}^{1}\int_{y=-1}^{1} 2x\cos(x^2-y)\,\mathrm{d}y\,\mathrm{d}x = \int_{x=0}^{1} 2x\left(\int_{y=-1}^{1}\cos(x^2-y)\,\mathrm{d}y\right)\mathrm{d}x$$
$$= (\ldots\text{ähnliche Rechnung wie eben}\ldots) = 1-\cos(2).$$

Wir wollen jetzt auch über unregelmäßig berandeten Gebieten G integrieren, nicht bloß über rechteckigen Intervallen I.

21.4 Gebietsintegral: Sei $G\subset\mathbb{R}^n$ kompakt mit stückweise glattem Rand, und die Funktion $f\colon G\to\mathbb{R}$ sei stückweise stetig. Wir nennen G das *Integrationsgebiet*.
Wir wählen ein Intervall $I\subset\mathbb{R}^n$ mit $G\subseteq I$ und setzen

$$\tilde f(\vec x) := \begin{cases} f(\vec x) & :\ \text{wenn } \vec x\in G,\\ 0 & :\ \text{wenn } \vec x\notin G,\end{cases}$$

und diese Funktion $\tilde f$ nennen wir *Fortsetzung von f durch Null*. Dann definieren wir

$$\int_G f(\vec x)\,\mathrm{d}\vec x := \int_I \tilde f(\vec x)\,\mathrm{d}\vec x.$$

Das Integral auf der linken Seite nennen wir *Gebietsintegral*.
Dies ist erstmal nur eine Definition zum Zwecke der Begriffsfestlegung. Die tatsächliche Berechnung eines solchen Integrals verlangt uns noch einige Arbeit ab, mit der wir jetzt loslegen.
Im Folgenden betrachten wir nur den Fall $n = 2$; höhere Dimensionen können analog betrachtet werden.

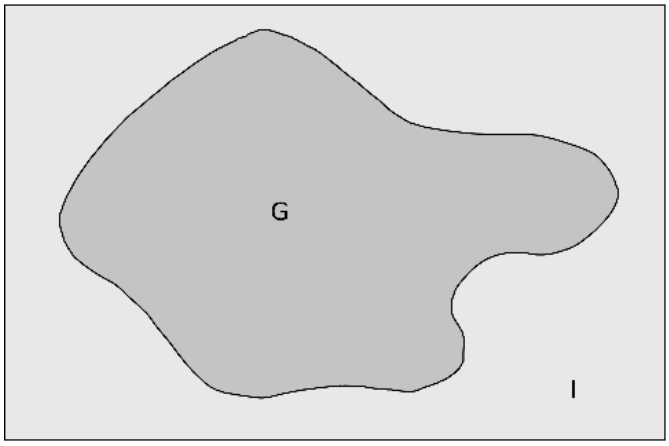


Abbildung 49: Ein beliebiges Integrationsgebiet G im $\mathbb{R}^2$.

21.5 Projizierbare Mengen: Sei $G \subset \mathbb{R}^2$ eine gegebene Menge.

1. Wenn es stetige Funktionen $c, d\colon [a,b] \to \mathbb{R}$ gibt mit $c(x) \le d(x)$ für jedes $x \in [a,b]$ und der Eigenschaft

$$G = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \colon a \le x \le b \quad \text{und} \quad c(x) \le y \le d(x) \right\},$$

dann heißt G *y–projizierbar*.

In diesem Fall haben wir

$$\int_G f(x,y)\,\mathrm{d}(x,y) = \int_{x=a}^{b} \int_{y=c(x)}^{d(x)} f(x,y)\,\mathrm{d}y\,\mathrm{d}x.$$

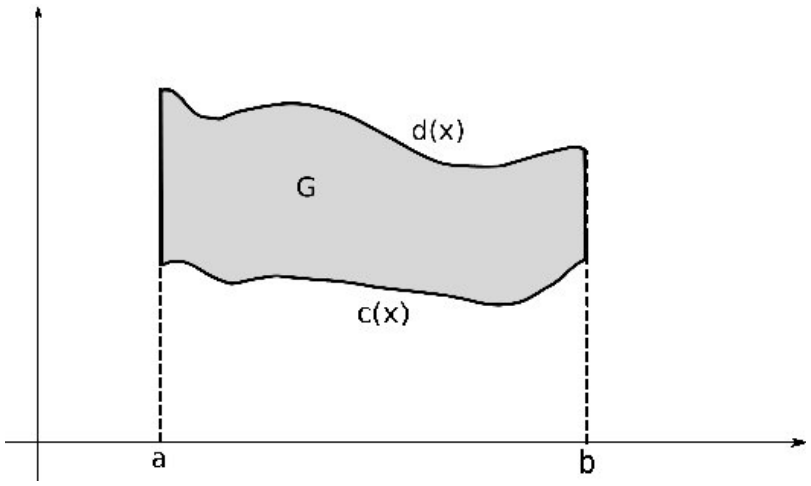


Abbildung 50: Eine y–projizierbare Menge G.

2. Wenn es stetige Funktionen $a, b\colon [c,d] \to \mathbb{R}$ gibt mit $a(y) \le b(y)$ für jedes $y \in [c,d]$ und der Eigenschaft

$$G = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \colon c \le y \le d \quad \text{und} \quad a(y) \le x \le b(y) \right\},$$

dann heißt G *x–projizierbar*.

In diesem Fall haben wir

$$\int_G f(x,y)\,\mathrm{d}(x,y) = \int_{y=c}^{d} \int_{x=a(y)}^{b(y)} f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

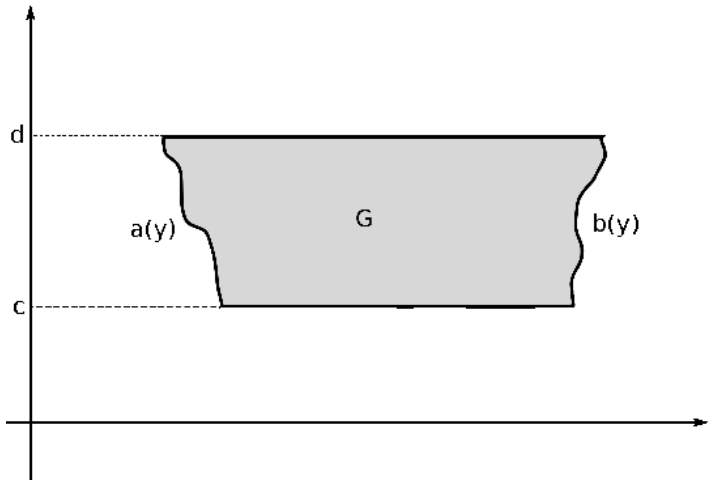


Abbildung 51: Eine x–projizierbare Menge G.

21.6 Beispiel:

a) Wir betrachten die obere Halbkreisscheibe mit Radius $r > 0$ (Abbildung 52):

$$G := \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \colon x^2 + y^2 \le r^2 \quad \text{und} \quad y \ge 0 \right\} =$$

$$= \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \colon -r \le x \le r \quad \text{und} \quad 0 \le y \le \sqrt{r^2 - x^2} \right\}.$$

Also ist diese Menge G y–projizierbar und wir haben

$$\int_G f(x,y)\,\mathrm{d}(x,y) = \int_{x=-r}^{r} \int_{y=0}^{\sqrt{r^2-x^2}} f(x,y)\,\mathrm{d}y\,\mathrm{d}x.$$

Zum Beispiel für $f(x,y) = 2x^2y$ ergibt sich dann:

$$\begin{aligned}
\int_G 2x^2y\,\mathrm{d}(x,y) &= \int_{x=-r}^{r} \int_{y=0}^{\sqrt{r^2-x^2}} 2x^2y\,\mathrm{d}y\,\mathrm{d}x = \int_{x=-r}^{r} x^2y^2 \Big|_{y=0}^{y=\sqrt{r^2-x^2}}\,\mathrm{d}x \\
&= \int_{x=-r}^{r} x^2(r^2-x^2)\,\mathrm{d}x = \int_{x=-r}^{r} \left(r^2x^2 - x^4\right)\mathrm{d}x \\
&= \left(\frac{1}{3}r^2x^3 - \frac{1}{5}x^5\right)\Big|_{x=-r}^{x=r} \\
&= \frac{1}{3}r^5 + \frac{1}{3}r^5 - \frac{1}{5}r^5 - \frac{1}{5}r^5 = \left(\frac{2}{3} - \frac{2}{5}\right)r^5 = \frac{4}{15}r^5.
\end{aligned}$$

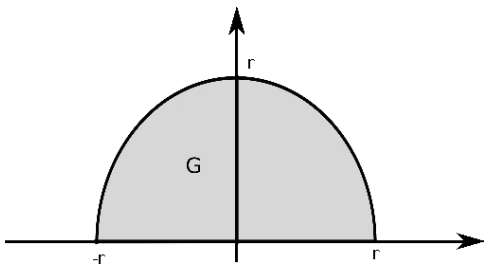


Abbildung 52: Eine Halbkreisscheibe.

b) Wir nehmen G als dasjenige gleichschenklige Dreieck in der Ebene, das begrenzt wird von den Geraden $y = x$, $y = 2 - x$ und $y = 0$ (Abbildung 53). Sei $f(x, y) = \frac{1}{1+x}$. Diese Menge G ist x–projizierbar und y–projizierbar. Wir können uns also für die leichtere Rechnung entscheiden:

$$\begin{aligned}\int_G f(x,y)\,\mathrm{d}(x,y) &= \int_{y=0}^{1}\int_{x=y}^{x=2-y} \frac{1}{1+x}\,\mathrm{d}x\,\mathrm{d}y = \int_{y=0}^{1} \ln(1+x)\Big|_{x=y}^{x=2-y}\,\mathrm{d}y \\ &= \int_{y=0}^{1} \Big(\ln(3-y) - \ln(1+y)\Big)\,\mathrm{d}y = \cdots = 3\ln(3) - 4\ln(2).\end{aligned}$$

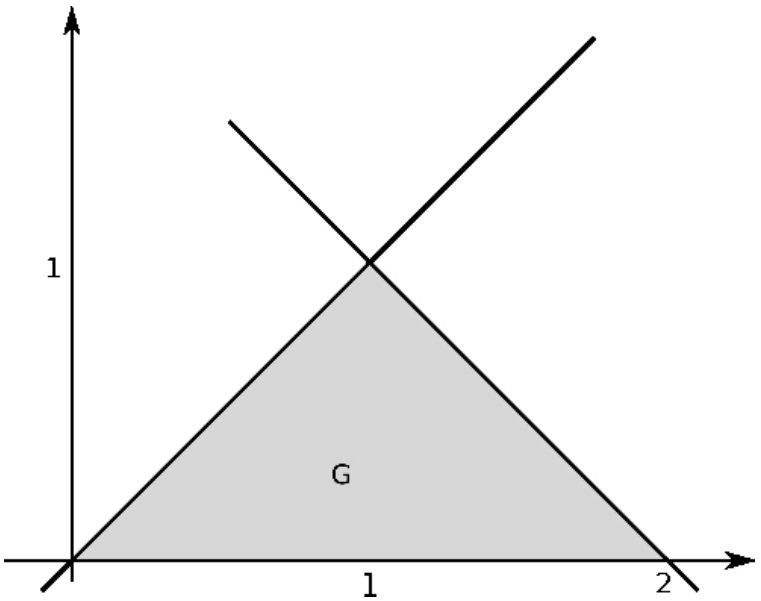


Abbildung 53: Ein Dreieck.

c) Wir betrachten einen Kreisring mit innerem Radius 1 und äußerem Radius 4 (Abbildung 54):

$$G := \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : 1 \le x^2 + y^2 \le 4 \right\}.$$

Leider ist diese Menge weder x–projizierbar noch y–projizierbar. Es gibt aber mindestens zwei Auswege:

- Wir zerlegen G in einen „oberen“ und einen „unteren“ Halbring:

$$G = \left(G \cap \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \colon y \geq 0 \right\}\right) \cup \left(G \cap \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 \colon y \leq 0 \right\}\right) = G_1 \cup G_2,$$

  und dann berechnen wir

$$\int_G f(x,y)\,\mathrm{d}(x,y) = \int_{G_1} f(x,y)\,\mathrm{d}(x,y) + \int_{G_2} f(x,y)\,\mathrm{d}(x,y).$$

  Beachte, dass G_1 und G_2 beide y–projizierbar sind.

- Wir schreiben G als Differenzmenge zweier Kreisscheiben, also $G = K_4(\vec{0}) \setminus K_1(\vec{0})$, und dann berechnen wir

$$\int_G f(x,y)\,\mathrm{d}(x,y) = \int_{K_4(\vec{0})} f(x,y)\,\mathrm{d}(x,y) - \int_{K_1(\vec{0})} f(x,y)\,\mathrm{d}(x,y).$$

Ein analoges Vorgehen ermöglicht für viele Mengen G die Rückführung auf projizierbare Mengen.

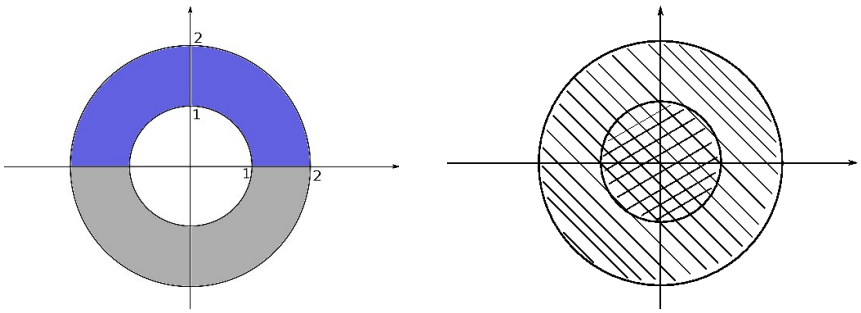


Abbildung 54: Wir zerlegen einen Kreisring in zwei Halbringe, oder wir interpretieren ihn als Differenzmenge zweier Kreisscheiben.

Ab jetzt gehen wir vom $\mathbb{R}^2$ wieder zurück zum $\mathbb{R}^n$ mit beliebigem n.

21.7 Substitutionsregel: Es seien G, $\hat{G} \subset \mathbb{R}^n$ Integrationsgebiete, $f \colon G \to \mathbb{R}$ sei stückweise stetig, und die Funktion $H \colon \hat{G} \to G$ sei bijektiv und stetig differenzierbar. Dann gilt

$$\int_G f(\vec{x})\,\mathrm{d}\vec{x} = \int_{\hat{G}} f(H(\vec{y}))\,|\det(J_H(\vec{y}))|\,\mathrm{d}\vec{y}.$$

Hierbei ist J_H die Jacobi–Matrix der Abbildung H.
Der Nutzen dieses Satzes wird für uns im wesentlichen darin liegen, ein schwierig zu beschreibendes Integrationsgebiet G in ein leichter zu beschreibendes Gebiet $\hat{G}$ umzurechnen.

21.8 Umrechnung in Polarkoordinaten: Wir erinnern daran, dass wir im $\mathbb{R}^2$ Polarkoordinaten eingeführt hatten gemäß

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r\cos(\varphi) \\ r\sin(\varphi) \end{bmatrix}.$$

Dann ist $H\colon [0,\infty)\times[0,2\pi)\to\mathbb{R}^2$ mit

$$H(r,\varphi) := \begin{bmatrix} r\cos(\varphi) \\ r\sin(\varphi) \end{bmatrix}$$

die Substitutionsfunktion bei der Umrechnung in Polarkoordinaten, und es ergibt sich

$$\int_G f(x,y)\,\mathrm{d}(x,y) = \int_{\hat{G}} f(H(r,\varphi))\,|\det J_H(r,\varphi)|\,\mathrm{d}(r,\varphi),$$

wobei

$$\det J_H(r,\varphi) = \det \begin{bmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{bmatrix} = r\cos^2(\varphi) + r\sin^2(\varphi) = r,$$

was wir in folgende Merkregel gießen:

$$\mathrm{d}x\,\mathrm{d}y = r\,\mathrm{d}r\,\mathrm{d}\varphi.$$

21.9 Beispiel:

1. Sei G ein Kreis um den Ursprung mit Radius R, also

$$G = K_R(\vec{0}) = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2\colon x^2+y^2 \le R^2 \right\}.$$

Dieser Menge G entspricht in Polarkoordinaten die Menge

$$\hat{G} = \{(r,\varphi)\in[0,\infty)\times[0,2\pi)\colon 0\le r\le R\} = [0,R]\times[0,2\pi),$$

und dieses $\hat{G}$ ist erfreulicherweise ein Intervall, was das Verwalten der Integrationsgrenzen deutlich vereinfacht. Also ist

$$\begin{aligned}\int_{K_R(\vec{0})} f(x,y)\,\mathrm{d}(x,y) &= \int_{\hat{G}} f(r\cos(\varphi), r\sin(\varphi))r\,\mathrm{d}(r,\varphi) \\ &= \int_{\varphi=0}^{2\pi}\int_{r=0}^{R} f(r\cos(\varphi), r\sin(\varphi))r\,\mathrm{d}r\,\mathrm{d}\varphi.\end{aligned}$$

Zum Beispiel für $f(x,y)=1$ sollte sich der Flächeninhalt des Kreises ergeben. Das prüfen wir nach:

$$\begin{aligned}\int_{K_R(\vec{0})} 1\,\mathrm{d}(x,y) &= \int_{\varphi=0}^{2\pi}\int_{r=0}^{R} r\,\mathrm{d}r\,\mathrm{d}\varphi = \left(\int_{\varphi=0}^{2\pi}\mathrm{d}\varphi\right)\cdot\left(\int_{r=0}^{R} r\,\mathrm{d}r\right) \\ &= 2\pi\cdot\frac{1}{2}r^2\Big|_{r=0}^{r=R} = 2\pi\cdot\frac{1}{2}R^2 = \pi R^2.\end{aligned}$$

2. Wir greifen Beispiel 21.6 (c) auf:

$$G = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2\colon 1\le x^2+y^2\le 4 \right\}, \qquad f(x,y) = 1+y.$$

In Polarkoordinaten wird dies zu $\hat{G} = \{(r,\varphi)\colon 1 \le r \le 2, \quad 0 \le \varphi \le 2\pi\}$ und $f(r\cos(\varphi), r\sin(\varphi)) = 1 + r\sin(\varphi)$, also

$$\begin{aligned}
\int_G (1+y)\,\mathrm{d}(x,y) &= \int_{\varphi=0}^{2\pi}\int_{r=1}^{2} (1+r\sin(\varphi))r\,\mathrm{d}r\,\mathrm{d}\varphi \\
&= \int_{\varphi=0}^{2\pi}\int_{r=1}^{2} r\,\mathrm{d}r\,\mathrm{d}\varphi + \int_{\varphi=0}^{2\pi}\int_{r=1}^{2} r^2\sin(\varphi))\,\mathrm{d}r\,\mathrm{d}\varphi \\
&= \left(\int_{\varphi=0}^{2\pi}\mathrm{d}\varphi\right)\cdot\left(\int_{r=1}^{2} r\,\mathrm{d}r\right) + \underbrace{\left(\int_{\varphi=0}^{2\pi}\sin(\varphi)\,\mathrm{d}\varphi\right)}_{=0} \quad \cdot \quad \underbrace{\left(\int_{r=1}^{2} r^2\,\mathrm{d}r\right)}_{\text{braucht man nicht berechnen}} \\
&= 2\pi\cdot\frac{1}{2}r^2\Big|_{r=1}^{r=2} \\
&= 2\pi\cdot\left(2-\frac{1}{2}\right) = 3\pi.
\end{aligned}$$

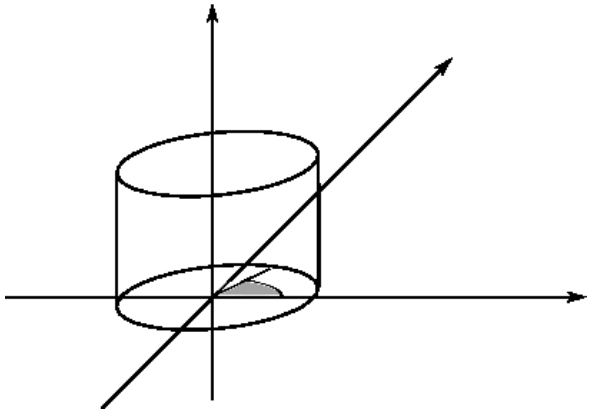


Abbildung 55: Der Winkel φ bei Zylinderkoordinaten.

21.10 Umrechnung in Zylinderkoordinaten: Dies sind Koordinaten, die zu einer Zylindersymmetrie im $\mathbb{R}^3$ gut passen. Wir rechnen um

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} r \\ \varphi \\ z \end{bmatrix}$$

gemäß der Vorschrift

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} r\cos(\varphi) \\ r\sin(\varphi) \\ z \end{bmatrix} = H(r,\varphi,z).$$

Für die Umrechnung der Differentiale benötigen wir den Betrag der Determinante der Jacobi–Matrix:

$$|\det J_H(r,\varphi,z)| = \left|\det\begin{bmatrix} \cos(\varphi) & -r\sin(\varphi) & 0 \\ \sin(\varphi) & r\cos(\varphi) & 0 \\ 0 & 0 & 1 \end{bmatrix}\right| = 1\cdot\left|\det\begin{bmatrix} \cos(\varphi) & -r\sin(\varphi) \\ \sin(\varphi) & r\cos(\varphi) \end{bmatrix}\right| = r,$$

analog zur Umrechnung in Polarkoordinaten. Wir erhalten also die Merkregel

$$\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = r\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}z.$$

21.11 Volumen eines Kegels: Wir bestimmen das Volumen eines Kegels. Dieser steht auf der Spitze, hat die Höhe 1 und den Radius 1. Es ist also

$$G = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \colon x^2 + y^2 \leq z^2, \quad 0 \leq z \leq 1 \right\}.$$

In Zylinderkoordinaten wird daraus dann $0 \leq r \leq z$, $0 \leq \varphi \leq 2\pi$ und $0 \leq z \leq 1$. Das Volumen des Kegels bestimmt sich dann zu (hierbei müssen wir aufpassen, weil die Grenzen für r von z abhängen, weshalb das r–Integral nach Innen gehört):

$$\begin{aligned} \int_G 1 \,\mathrm{d}(x,y,z) &= \int_{z=0}^{1} \int_{\varphi=0}^{2\pi} \int_{r=0}^{z} 1 \cdot r \,\mathrm{d}r \,\mathrm{d}\varphi \,\mathrm{d}z = \int_{z=0}^{1} \int_{\varphi=0}^{2\pi} \frac{1}{2} r^2 \Big|_{r=0}^{r=z} \mathrm{d}\varphi \,\mathrm{d}z \\ &= \int_{z=0}^{1} \int_{\varphi=0}^{2\pi} \frac{1}{2} z^2 \,\mathrm{d}\varphi \,\mathrm{d}z = \frac{1}{2} \cdot 2\pi \cdot \int_{z=0}^{1} z^2 \,\mathrm{d}z \\ &= \pi \cdot \frac{1}{3} z^3 \Big|_{z=0}^{z=1} = \frac{1}{3}\pi. \end{aligned}$$

Das entspricht unseren Erwartungen („Grundflächeninhalt mal Höhe durch drei“).

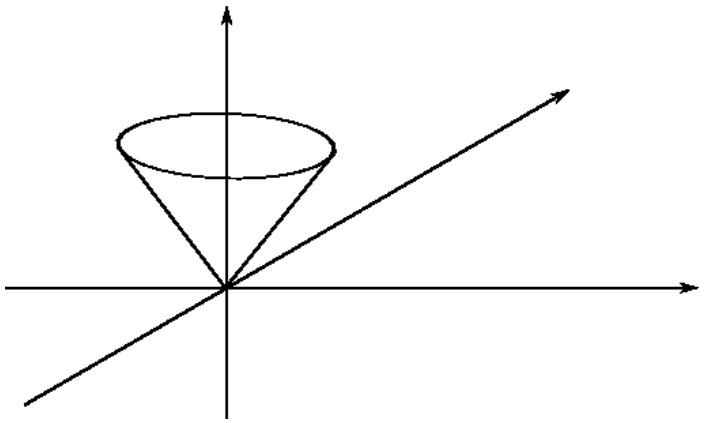


Abbildung 56: Wir suchen das Volumen dieses Kegels.

21.12 Umrechnung in Kugelkoordinaten: Dies sind Koordinaten, die zu einer Kugelsymmetrie im $\mathbb{R}^3$ gut passen. Wir nehmen $r \geq 0$ als den Abstand von $[x,y,z]^\top$ zum Ursprung, $\varphi \in [0, 2\pi)$ als den Winkel von $[x,y]^\top$ zur positiven x–Achse (auch als geographische Länge bezeichnet), und $\theta \in [0,\pi]$ ist der Winkel von $[x,y,z]^\top$ zur positiven z–Achse. Für $\theta = 0$ ergibt sich der Nordpol, für $\theta = \pi$ der Südpol, und für $\theta = \pi/2$ der Äquator. (Achtung: in manchen Büchern ist $\theta \in [-\pi/2, \pi/2]$, wobei $\theta = 0$ den Äquator beschreibt. Dann ändern sich einige Formeln.)
Die Wortbeschreibung ergibt dann die Formeln

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = H(r,\varphi,\theta) = \begin{bmatrix} r\cos(\varphi)\sin(\theta) \\ r\sin(\varphi)\sin(\theta) \\ r\cos(\theta) \end{bmatrix}.$$

Für die Umrechung der Differentiale benötigen wir die Determinante der Jacobi–Matrix:

$$\begin{aligned}\det J_H(r,\varphi,\theta) &= \begin{vmatrix} \cos(\varphi)\sin(\theta) & -r\sin(\varphi)\sin(\theta) & r\cos(\varphi)\cos(\theta) \\ \sin(\varphi)\sin(\theta) & r\cos(\varphi)\sin(\theta) & r\sin(\varphi)\cos(\theta) \\ \cos(\theta) & 0 & -r\sin(\theta) \end{vmatrix} \\ &= (\dots \text{einige Rechnungen} \dots) = -r^2\sin(\theta).\end{aligned}$$

Die Substitutionsformel verlangt jetzt allerdings den *Betrag* der Determinante der Jacobi–Matrix, also bekommen wir die Merkregel

$$\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = r^2\sin(\theta)\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}\theta.$$

21.13 Beispiel: Wir bestimmen das Volumen der Einheitskugel im $\mathbb{R}^3$, also betrachten wir

$$G = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \colon x^2 + y^2 + z^2 \le 1 \right\}.$$

In Kugelkoordinaten wird dies zu $0 \le r \le 1, \quad 0 \le \varphi \le 2\pi, \quad 0 \le \theta \le \pi$. Das Volumen der Kugel berechnet sich dann so:

$$\begin{aligned}\int_G 1\,\mathrm{d}(x,y,z) &= \int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi}\int_{r=0}^{1} 1\cdot r^2\sin(\theta)\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}\theta \\ &= \left(\int_{\theta=0}^{\pi}\sin(\theta)\,\mathrm{d}\theta\right)\cdot\left(\int_{\varphi=0}^{2\pi}\mathrm{d}\varphi\right)\cdot\left(\int_{r=0}^{1} r^2\,\mathrm{d}r\right) \\ &= -\cos(\theta)\Big|_{\theta=0}^{\theta=\pi}\cdot 2\pi\cdot\frac{1}{3}r^3\Big|_{r=0}^{r=1} \\ &= 2\cdot 2\pi\cdot\frac{1}{3} = \frac{4}{3}\pi.\end{aligned}$$

Das ist das erwartete Ergebnis.

22 Vektoranalysis

Unser **Ziel** ist es, Vektorfelder besser zu beschreiben. Wir denken dabei an elektrische oder magnetische Felder, an das Gravitationsfeld, an Strömungen oder an elastische Verformungen: sei z.B. ein Stahlträger gegeben, der sich unter Last ein wenig verbiegt. Das können wir uns so vorstellen, dass wir den Träger im Ruhezustand hernehmen und an jeden seiner Punkte einen Vektor anheften, der diesen Punkt im Ruhezustand mit dem neuen Punkt unter Last verbindet. Auf diesem Wege entsteht ein Vektorfeld, das *Verschiebungsfeld* genannt wird. Wie können wir dies mathematisch beschreiben ?

22.1 Divergenz: Sei $M \subseteq \mathbb{R}^n$ offen, und $F = (f_1, \ldots, f_n)^T \colon M \to \mathbb{R}^n$ sei ein stetig differenzierbares Vektorfeld. Dann nennen wir

$$\operatorname{div} F(\vec{x}) := \sum_{j=1}^{n} \frac{\partial f_j}{\partial x_j}(\vec{x}) = \frac{\partial f_1}{\partial x_1}(\vec{x}) + \frac{\partial f_2}{\partial x_2}(\vec{x}) + \ldots + \frac{\partial f_n}{\partial x_n}(\vec{x}), \qquad \vec{x} \in M,$$

die *Divergenz von* F. Beachte dabei: es ist $F \colon M \to \mathbb{R}^n$ ein Vektorfeld, aber $\operatorname{div} F \colon M \to \mathbb{R}$ ein Skalarfeld.

22.2 Anschauliche Bedeutung: Die Funktion $\operatorname{div} F$ misst den aus einer Volumeneinheit austretenden Fluss:

- Punkte $\vec{x} \in M$ mit $\operatorname{div} F(\vec{x}) > 0$ heißen *Quellen*,
- Punkte $\vec{x} \in M$ mit $\operatorname{div} F(\vec{x}) < 0$ heißen *Senken*.
- Wenn $\operatorname{div} F(\vec{x}) = 0$ für alle $\vec{x} \in M$ gilt, so nennt man das Vektorfeld F *quellenfrei*.

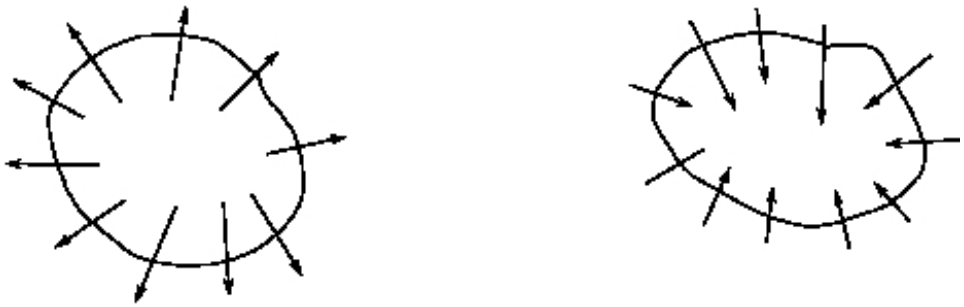


Abbildung 57: Links ist eine Quelle mit $\operatorname{div} F > 0$, und rechts ist eine Senke mit $\operatorname{div} F < 0$.

22.3 Rotation: Sei $M \subset \mathbb{R}^3$ offen, und sei $F = (f_1, f_2, f_3)^T \colon M \to \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld. Dann heißt

$$\operatorname{rot} F(\vec{x}) := \begin{bmatrix} \frac{\partial f_3}{\partial x_2}(\vec{x}) - \frac{\partial f_2}{\partial x_3}(\vec{x}) \\ \frac{\partial f_1}{\partial x_3}(\vec{x}) - \frac{\partial f_3}{\partial x_1}(\vec{x}) \\ \frac{\partial f_2}{\partial x_1}(\vec{x}) - \frac{\partial f_1}{\partial x_2}(\vec{x}) \end{bmatrix}$$

Rotation von F. Beachte dabei: $F \colon M \to \mathbb{R}^3$ und $\operatorname{rot} F \colon M \to \mathbb{R}^3$ sind beides Vektorfelder. Die Rotation gibt es nur im $\mathbb{R}^3$, aber nicht im $\mathbb{R}^2$.

22.4 Bemerkung:

1. Mit dem n–komponentigen *Nabla–Operator*

$$\vec{\nabla}_n := \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix}$$

haben wir formal

$$\operatorname{div} F(\vec{x}) = \underbrace{\vec{\nabla}_n^T \cdot F(\vec{x})}_{\text{Skalarprodukt}}, \qquad \operatorname{rot} F(\vec{x}) = \underbrace{(\vec{\nabla}_3 \times F)(\vec{x})}_{\text{Kreuzprodukt}}.$$

2. Wir erinnern an die *Integrabilitätsbedingung* aus dem Abschnitt über Potenziale:

$$\frac{\partial f_i}{\partial x_j}(\vec{x}) = \frac{\partial f_j}{\partial x_i}(\vec{x}) \qquad \text{für alle } i \neq j.$$

Diese Bedingung ist also erfüllt für $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ genau dann, wenn $\operatorname{rot} F(\vec{x}) = \vec{0}$ für alle $\vec{x} \in \mathbb{R}^3$ gilt.

3. Es beschreibt $\operatorname{rot} F$ die Wirbeldichte des Feldes.

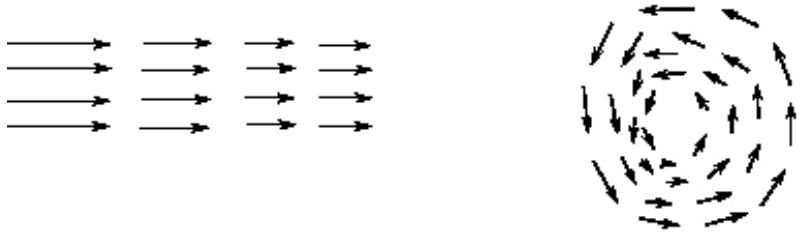


Abbildung 58: Links ist ein Vektorfeld F mit $\operatorname{rot} F = \vec{0}$, rechts ist ein Vektorfeld F mit $\operatorname{rot} F \neq \vec{0}$.

- Gradientenfelder F haben immer $\operatorname{rot} F = \vec{0}$, denn: sei $\varphi\colon \mathbb{R}^3 \to \mathbb{R}$ ein zweimal stetig differenzierbares Skalarfeld. Dann gilt

$$\operatorname{rot}(\nabla\varphi) = \operatorname{rot} \begin{bmatrix} \varphi_x \\ \varphi_y \\ \varphi_z \end{bmatrix} = \begin{bmatrix} \varphi_{zy} - \varphi_{yz} \\ \varphi_{xz} - \varphi_{zx} \\ \varphi_{yx} - \varphi_{xy} \end{bmatrix} = \vec{0}.$$

- Wenn $\operatorname{rot} F \neq \vec{0}$ ist, so kann F kein Gradientenfeld sein. Typische Beispiele sind Magnetfelder.

Wir merken uns die anschauliche Interpretation: „rot F misst die Abweichung des Vektorfeldes F von einem Gradientenfeld“.

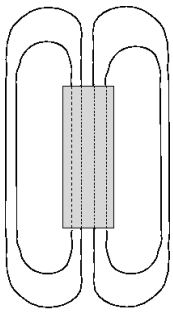


Abbildung 59: Die Feldlinien zum Magnetfeld eines Stabmagneten (sehr schematisch).

22.5 Bemerkung: Wir kombinieren die Operatoren ∇, div und rot:

- Für $\varphi\colon \mathbb{R}^3 \to \mathbb{R}$ gilt $\operatorname{rot}(\nabla\varphi) = \vec{0}$.
- Für $F\colon \mathbb{R}^3 \to \mathbb{R}^3$ gilt $\operatorname{div}(\operatorname{rot} F) = 0$.
- Für $\varphi\colon \mathbb{R}^n \to \mathbb{R}$ gilt

$$\begin{aligned}\triangle\,\varphi := \operatorname{div}(\nabla\varphi) = \operatorname{div}(\operatorname{grad}\varphi) &= \operatorname{div}\begin{bmatrix}\frac{\partial\varphi}{\partial x_1} & \frac{\partial\varphi}{\partial x_2} & \dots & \frac{\partial\varphi}{\partial x_n}\end{bmatrix}\\ &= \sum_{j=1}^{n}\frac{\partial}{\partial x_j}\frac{\partial\varphi}{\partial x_j} = \sum_{j=1}^{n}\frac{\partial^2\varphi}{\partial x_j^2},\end{aligned}$$

  und wir nennen $\triangle := \sum_{j=1}^{n}\frac{\partial^2}{\partial x_j^2}$ den *Laplace*–Operator.

22.6 Flächen im $\mathbb{R}^3$: Sei $D \subset \mathbb{R}^2$ kompakt, $G \supset D$ offen, und sei $\Phi\colon G \to \mathbb{R}^3$ injektiv und stetig differenzierbar, sodass die Jacobi–Matrix von Φ, das heißt

$$J_\Phi(u,v) = \begin{bmatrix}\frac{\partial\Phi_1}{\partial u}(u,v) & \frac{\partial\Phi_1}{\partial v}(u,v)\\ \frac{\partial\Phi_2}{\partial u}(u,v) & \frac{\partial\Phi_2}{\partial v}(u,v)\\ \frac{\partial\Phi_3}{\partial u}(u,v) & \frac{\partial\Phi_3}{\partial v}(u,v)\end{bmatrix},$$

für alle $(u,v) \in G$ vollen Rang hat, also $\operatorname{rang} J_\Phi(u,v) = 2$.
Dann heißt Φ *reguläre Fläche (über D)*, und

$$\mathcal{F} := \Phi(D) = \left\{\Phi(u,v)\colon \begin{bmatrix}u\\v\end{bmatrix} \in D\right\}$$

heißt *reguläres Flächenstück*.

22.7 Beispiel: Wir wollen von einer Kugel (deren Mittelpunkt im Ursprung liegt) mit Radius R die obere und untere Halbkugelfläche beschreiben. Dazu wählen wir das Parametergebiet

$$G := \left\{\begin{bmatrix}u\\v\end{bmatrix} \in \mathbb{R}^2\colon u^2+v^2 < R^2\right\},$$

was eine offene Teilmenge des $\mathbb{R}^2$ ist. Für die Halbkugelflächen wählen wir zwei Funktionen Φ_+ und Φ_-,

$$\Phi_\pm(u,v) = \begin{bmatrix}u\\v\\ \pm\sqrt{R^2-u^2-v^2}\end{bmatrix},$$

und diese sind stetig differenzierbar auf G. Wir haben die Jacobi–Matrizen

$$J_{\Phi_\pm}(u,v) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\mp u}{\sqrt{R^2-u^2-v^2}} & \frac{\mp v}{\sqrt{R^2-u^2-v^2}} \end{bmatrix},$$

was für alle $[u,v]^T \in G$ immer den Rang zwei besitzt.
Auf dem Rand von G, also für $u^2+v^2=R^2$, bekämen wir aber in der letzten Zeile von $J_{\Phi_\pm}$ zwei Divisionen durch Null. Weil G eine offene Menge des $\mathbb{R}^2$ ist, gehört der Rand von G aber sowieso nicht zu G. Wir wollen einen kleinen Sicherheitsabstand ε zum Rand von G einhalten. Deshalb definieren wir

$$D := \left\{ \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^2 \colon u^2+v^2 \le (R-\varepsilon)^2 \right\}, \qquad \varepsilon > 0.$$

Die Funktionen $\Phi_\pm$ sind reguläre Flächen über D, und die zugehörigen Flächenstücke $\mathcal{F}_+ := \Phi_+(D)$ sowie $\mathcal{F}_- := \Phi_-(D)$ sind dann (abgesehen von einer sehr schmalen Bauchbinde entlang des Äquators) die obere Halbkugelfläche und die untere Halbkugelfläche.

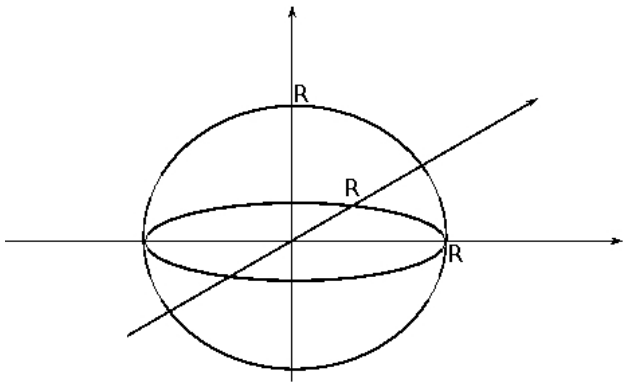


Abbildung 60: Die Kugel mit Radius R.

Jetzt können wir Flächenstücke beschreiben, und als nächstes schauen wir uns deren Tangententialebenen an.

22.8 Tangentialebene und Normalenvektor einer Fläche: Sei $D \subset \mathbb{R}^2$ kompakt, $G \supset D$ offen, $\Phi\colon G \to \mathbb{R}^3$ eine reguläre Fläche. Dann heißt für $[u_0,v_0]^T \in D$ die Ebene

$$T(u_0,v_0) := \left\{ \Phi(u_0,v_0) + \lambda \frac{\partial \Phi}{\partial u}(u_0,v_0) + \mu \frac{\partial \Phi}{\partial v}(u_0,v_0) \colon \lambda,\mu \in \mathbb{R} \right\}$$

Tangentialebene an die Fläche Φ *im Punkt* $\Phi(u_0,v_0)$.
Weiterhin heißt

$$\vec{n}(u_0,v_0) := \frac{\frac{\partial \Phi}{\partial u}(u_0,v_0) \times \frac{\partial \Phi}{\partial v}(u_0,v_0)}{\left\| \frac{\partial \Phi}{\partial u}(u_0,v_0) \times \frac{\partial \Phi}{\partial v}(u_0,v_0) \right\|}$$

Normalenvektor an die Fläche Φ *im Punkt* $\Phi(u_0,v_0)$.
Man beachte: bei der Beschreibung der Tangentialebene haben wir den Aufpunkt $\Phi(u_0,v_0)$ genommen, und dort haben wir die beiden Vektoren $\lambda\partial_u\Phi$ und $\mu\partial_v\Phi$ angehängt. Diese beiden Vektoren zeigen in unterschiedliche Richtungen des $\mathbb{R}^3$, denn sie sind Vielfache

der beiden Spalten der Jacobi–Matrix J_Φ, und diese Spalten sind linear unabhängig, weil wir vorausgesetzt hatten, dass die Jacobi–Matrix J_Φ den Rang zwei hat (und der Rang ist die Anzahl der linear unabhänigen Spalten). Also ist $T(u_0, v_0)$ tatsächlich die Beschreibung einer Ebene.
Wenn wir das Kreuzprodukt $\partial_u\Phi \times \partial_v\Phi$ der beiden die Tangentialebene aufspannenden Vektoren bilden, dann entsteht (gemäß der Eigenschaften des Kreuzprodukts) ein Produktvektor, der auf jedem der beiden Faktoren $\partial_u\Phi(u_0, v_0)$ und $\partial_v\Phi(u_0, v_0)$ senkrecht steht. Also steht $\vec{n}(u_0, v_0)$ tatsächlich auf der Tangentialebene $T(u_0, v_0)$ senkrecht. Das rechtfertigt den Namen *Normalenvektor*.

22.9 Flächeninhalt: Sei $G \subseteq \mathbb{R}^2$ offen, $D \subset G$ kompakt, und $\Phi\colon G \to \mathbb{R}^3$ sei eine reguläre Fläche über D. Dann ist der Flächeninhalt von $\Phi(D)$ gegeben durch

$$\int_D \left\| \frac{\partial\Phi}{\partial u}(u,v) \times \frac{\partial\Phi}{\partial v}(u,v) \right\| \mathrm{d}(u,v). \qquad (\spadesuit)$$

Das probieren wir anhand einer Kugeloberfläche aus, siehe Beispiel 22.7:

22.10 Beispiel: Wir haben

$$\Phi_\pm(u,v) = \begin{bmatrix} u \\ v \\ \pm\sqrt{R^2-u^2-v^2} \end{bmatrix}, \qquad J_{\Phi_\pm}(u,v) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\mp u}{\sqrt{R^2-u^2-v^2}} & \frac{\mp v}{\sqrt{R^2-u^2-v^2}} \end{bmatrix},$$

und das Kreuzprodukt der beiden Spalten von J_Φ berechnen wir so:

$$\frac{\partial\Phi}{\partial u}(u,v) \times \frac{\partial\Phi}{\partial v}(u,v) = \begin{bmatrix} 0 \pm \frac{u}{\sqrt{R^2-u^2-v^2}} \\ 0 \pm \frac{v}{\sqrt{R^2-u^2-v^2}} \\ 1-0 \end{bmatrix} = \frac{\pm 1}{\sqrt{R^2-u^2-v^2}} \begin{bmatrix} u \\ v \\ \pm\sqrt{R^2-u^2-v^2} \end{bmatrix}.$$

Davon berechnen wir jetzt die Länge:

$$\left\| \frac{\partial\Phi}{\partial u}(u,v) \times \frac{\partial\Phi}{\partial v}(u,v) \right\| = \frac{1}{\sqrt{R^2-u^2-v^2}} \sqrt{u^2+v^2+(R^2-u^2-v^2)} = \frac{R}{\sqrt{R^2-u^2-v^2}}.$$

Wenn wir das Kreuzprodukt auf Länge Eins bringen, dann entsteht der Normalenvektor:

$$\vec{n}(u,v) = \frac{\frac{\partial\Phi}{\partial u}(u,v) \times \frac{\partial\Phi}{\partial v}(u,v)}{\left\| \frac{\partial\Phi}{\partial u}(u,v) \times \frac{\partial\Phi}{\partial v}(u,v) \right\|} = \frac{1}{R} \begin{bmatrix} u \\ v \\ \pm\sqrt{R^2-u^2-v^2} \end{bmatrix}.$$

Wir berechnen jetzt den Flächeninhalt der Halbkugelfläche, wobei wir $D := \{[u,v]^T \in \mathbb{R}^2 \colon u^2+v^2 \le R^2\}$ nehmen und in Polarkordinaten umrechnen:

$$\begin{aligned}
&\int_D \left\| \frac{\partial \Phi}{\partial u}(u,v) \times \frac{\partial \Phi}{\partial v}(u,v) \right\| \mathrm{d}(u,v) = \int_D \frac{R}{\sqrt{R^2-u^2-v^2}} \,\mathrm{d}(u,v) \\
&= \int_{\varphi=0}^{2\pi} \int_{r=0}^{R} \frac{R}{\sqrt{R^2-r^2}} r \,\mathrm{d}r \,\mathrm{d}\varphi \\
&= 2\pi \int_{r=0}^{r=R} \frac{R}{\sqrt{R^2-r^2}} r \,\mathrm{d}r \qquad \Big| \quad \text{substituiere } t := R^2-r^2, \quad \frac{\mathrm{d}t}{\mathrm{d}r} = -2r \\
&= 2\pi \cdot \int_{t=R^2}^{t=0} \frac{R}{\sqrt{t}} \cdot \frac{-1}{2} \,\mathrm{d}t \\
&= -R\pi \cdot \int_{t=R^2}^{t=0} \frac{1}{\sqrt{t}} \,\mathrm{d}t \\
&= +R\pi \cdot \int_{t=0}^{t=R^2} \frac{1}{\sqrt{t}} \,\mathrm{d}t \\
&= R\pi \cdot 2\sqrt{t}\Big|_{t=0}^{t=R^2} = 2\pi R^2.
\end{aligned}$$

Die Gesamtkugeloberfläche ist also $4\pi R^2$, wie zu erwarten.
Wir führen eine Schreibweise ein:

$$\mathrm{d}\sigma := \left\| \frac{\partial \Phi}{\partial u}(u,v) \times \frac{\partial \Phi}{\partial v}(u,v) \right\| \mathrm{d}(u,v).$$

Dann können wir das Integral (♠) auch als $\int_D 1 \,\mathrm{d}\sigma$ schreiben. Das wollen wir jetzt etwas verallgemeinern:

22.11 Oberflächenintegral: Sei Φ eine reguläre Fläche über D und $\mathcal{F} := \Phi(D)$ das zugehörige reguläre Flächenstück. Sei $f \colon \mathcal{F} \to \mathbb{R}$ eine stetige skalare Funktion. Dann definieren wir

$$\int_{\mathcal{F}} f(\vec{x}) \,\mathrm{d}\sigma(\vec{x}) := \int_D f(\Phi(u,v)) \left\| \frac{\partial \Phi}{\partial u}(u,v) \times \frac{\partial \Phi}{\partial v}(u,v) \right\| \mathrm{d}(u,v),$$

und dieses Integral nennen wir *Oberflächenintegral von f über der Fläche $\mathcal{F}$*. Beachte, dass wir genau den Flächeninhalt bekommen, falls $f = 1$.

22.12 Beispiel: Wir nehmen $f(x,y,z) = z$ und $\Phi \colon D \to \mathbb{R}^3$ mit

$$D := \left\{ \begin{bmatrix} u \\ v \end{bmatrix} : u^2+v^2 \le R^2 \right\}, \qquad \Phi(u,v) = \begin{bmatrix} u \\ v \\ \sqrt{R^2-u^2-v^2} \end{bmatrix},$$

wie eben. Dann berechnen wir

$$\begin{aligned}
&\int_{\mathcal{F}} f(x,y,z) \,\mathrm{d}\sigma = \int_D f(u,v,\sqrt{R^2-u^2-v^2}) \left\| \frac{\partial \Phi}{\partial u}(u,v) \times \frac{\partial \Phi}{\partial v}(u,v) \right\| \mathrm{d}(u,v) \qquad \Big| \quad \text{siehe oben} \\
&= \int_D \sqrt{R^2-u^2-v^2} \cdot \frac{R}{\sqrt{R^2-u^2-v^2}} \,\mathrm{d}(u,v) = R \int_D \mathrm{d}(u,v) \qquad \Big| \quad \text{Polarkoordinaten} \\
&= R \int_{\vec{p}=0}^{2\pi} \int_{r=0}^{R} r \,\mathrm{d}r \,\mathrm{d}\vec{p} = 2\pi R \cdot \frac{1}{2} r^2 \Big|_{r=0}^{r=R} = \pi R^3.
\end{aligned}$$

Der nächste Satz ist ein berühmtes Theorem, das uns die Möglichkeit gibt, bestimmte Integrale über dreidimensionale Mengen umzuwandeln in Oberflächenintegrale über den Rand dieser Mengen.

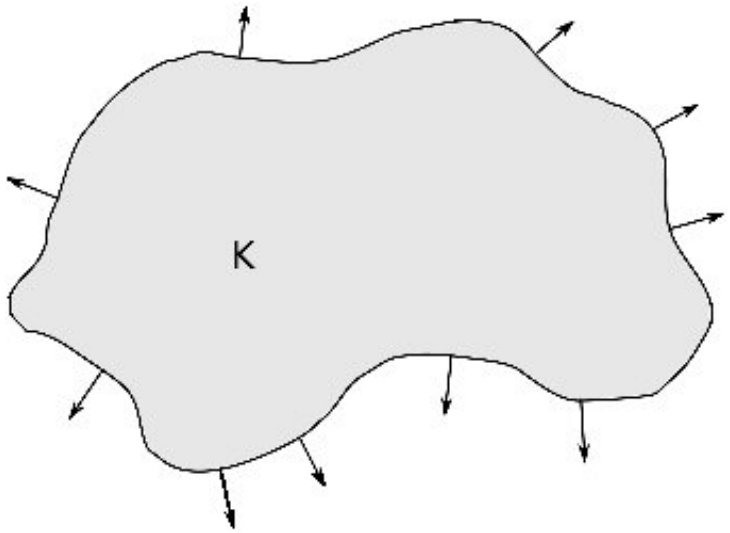


Abbildung 61: Eine Menge K mit einigen Normalenvektoren $\vec{\nu}(\vec{x})$. Diese Normalenvektoren stehen immer senkrecht auf dem Rand ∂K, und sie haben immer die Länge Eins.

22.13 Gaußscher Integralsatz: Es sei $K \subset \mathbb{R}^3$ eine offene beschränkte Menge mit „genügend glattem“ Rand ∂K (also insbesondere kein Fraktal oder ähnliches). Sei $G \supset \overline{K}$ eine offene Menge, und sei $F\colon G \to \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld. Für jeden Punkt $\vec{x} \in \partial K$ bezeichne $\vec{\nu}(\vec{x})$ den nach Außen zeigenden Normalenvektor an die Fläche ∂K.
Dann ist

$$\underbrace{\int_K \operatorname{div} F(\vec{x})\,\mathrm{d}\vec{x}}_{\text{Gebietsintegral}} = \underbrace{\int_{\partial K} \langle F(\vec{x}), \vec{\nu}(\vec{x})\rangle\ \mathrm{d}\sigma(\vec{x})}_{\text{Oberflächenintegral}}.$$

22.14 Geometrische Interpretation: Wir interpretieren den Vektor $F(\vec{x})$ als die vektorielle Fließgeschwindigkeit einer Flüssigkeit am Ortspunkt $\vec{x}$. Was bedeuten dann die beiden Integrale?
Wir beginnen mit dem Oberflächenintegral auf der rechten Seite. Wir nehmen einen Punkt $\vec{x}$ auf der Oberfläche ∂K. Dann ist $\langle F(\vec{x}), \vec{\nu}(\vec{x})\rangle$ das Skalarprodukt der beiden Vektoren $F(\vec{x})$ und $\vec{\nu}(\vec{x})$, also

$$\langle F(\vec{x}), \vec{\nu}(\vec{x})\rangle = \|F(\vec{x})\| \cdot \|\vec{\nu}(\vec{x})\| \cdot \cos\angle\left(F(\vec{x}), \vec{\nu}(\vec{x})\right).$$

Wenn der Vektor $F(\vec{x})$ aus der Menge K herauszeigt, dann fließt dort die Flüssigkeit aus K heraus, und der Winkel zwischen $F(\vec{x})$ und $\vec{\nu}(\vec{x})$ ist kleiner als $\pi/2$, der Cosinus also positiv. Um das Skalarprodukt besser zu verstehen, nehmen wir uns eine Menge $D \subset \mathbb{R}^2$ und eine reguläre Fläche $\Phi\colon D \to \mathbb{R}^3$, die in einer Umgebung von $\vec{x}$ den Rand ∂K beschreibt. Dann gilt

$$\vec{\nu}(\vec{x}) = \frac{\frac{\partial\Phi}{\partial u}(u,v) \times \frac{\partial\Phi}{\partial v}(u,v)}{\left\|\frac{\partial\Phi}{\partial u}(u,v) \times \frac{\partial\Phi}{\partial v}(u,v)\right\|},$$

also folgt dann

$$\begin{aligned}\langle F(\vec{x}), \vec{\nu}(\vec{x})\rangle \, \mathrm{d}\sigma(\vec{x}) &= \langle F(\vec{x}), \vec{\nu}(\vec{x})\rangle \left\| \frac{\partial \Phi}{\partial u}(u,v) \times \frac{\partial \Phi}{\partial v}(u,v) \right\| \mathrm{d}(u,v) \\ &= \underbrace{\left\langle F(\vec{x}), \frac{\partial \Phi}{\partial u}(u,v) \times \frac{\partial \Phi}{\partial v}(u,v) \right\rangle}_{\text{Spatprodukt}} \mathrm{d}(u,v),\end{aligned}$$

und dieses Spatprodukt ist gleich dem Volumen des von den drei Vektoren $F(\vec{x})$, $\partial_u \Phi(u,v)$ und $\partial_v \Phi(u,v)$ aufgespannten Parallelepipeds (Spats), also gleich dem Flüssigkeitsvolumen, das durch das Oberflächenelement $\mathrm{d}\sigma(\vec{x})$ strömt.
Das bedeutet insgesamt: das Oberflächenintegral im Gaußschen Satz entspricht dem Gesamtfluss des Vektorfelds F durch die Fläche ∂K. Dabei zählt der Fluss von innen nach außen positiv, und der Fluss von außen nach innen zählt negativ.

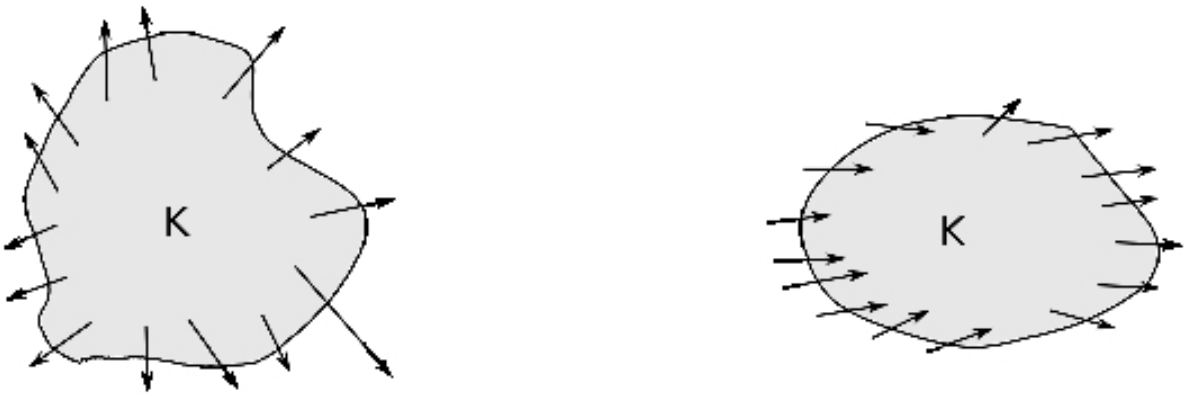


Abbildung 62: Links haben wir einen positiven Gesamtfluss über den Rand ∂K, rechts haben wir den Gesamtfluss gleich Null über ∂K.

Jetzt überlegen wir uns, was das Gebietsintegral auf der linken Seite des Integralsatzes von Gauß bedeutet. Wir erinnern daran, dass $\operatorname{div} F(\vec{x})$ angibt, ob die Flüssigkeitsströmung im Punkt $\vec{x}$ eine Quelle oder eine Senke besitzt. Dann ist $\int_K \operatorname{div} F(\vec{x})\,\mathrm{d}\vec{x}$ also die Gesamtsumme des in den Quellen entstehenden Flusses, vermindert um die Gesamtsumme des in den Senken versickernden Flusses.
Der Integralsatz von Gauß besagt damit: der Gesamtfluss über den Rand ∂K ist gleich der Gesamtmenge des im Innern erzeugten oder verschwundenen Flusses.

22.15 Beispiel: Gesucht ist $\int_{\partial K} \langle F(\vec{x}), \vec{\nu}(\vec{x})\rangle \,\mathrm{d}\sigma(\vec{x})$ für

$$F(x,y,z) = \begin{bmatrix} xy^2 \\ x^2 y \\ y \end{bmatrix} \qquad \text{und} \qquad K = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \colon x^2 + y^2 \le 1, \quad -1 \le z \le 1 \right\}.$$

Wir integrieren also über einen aufrecht stehenden Zylinder. Der Gauß–Satz liefert

$$\begin{aligned} &\int_{\partial K} \langle F(\vec{x}), \vec{\nu}(\vec{x})\rangle \,\mathrm{d}\sigma(\vec{x}) = \int_K \operatorname{div} F(\vec{x})\,\mathrm{d}\vec{x} = \int_K \left(\frac{\partial f_1}{\partial x_1}(\vec{x}) + \frac{\partial f_2}{\partial x_2}(\vec{x}) + \frac{\partial f_3}{\partial x_3}(\vec{x}) \right) \mathrm{d}\vec{x} \\ &= \int_K \left(y^2 + x^2 + 0 \right) \mathrm{d}(x,y,z) \qquad \Big| \quad \text{Zylinderkoordinaten} \\ &= \int_{z=-1}^{1} \int_{\varphi=0}^{2\pi} \int_{r=0}^{1} r^2 \cdot r \,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}z = 2 \cdot 2\pi \cdot \frac{1}{4} r^4 \Big|_{r=0}^{r=1} = 4\pi \cdot \frac{1}{4} = \pi. \end{aligned}$$

22.16 Integralsatz von Stokes: Sei $D \subset \mathbb{R}^2$ kompakt, $G \supset D$ offen und $\Phi \colon G \to \mathbb{R}^3$ zweimal stetig differenzierbar und eine auf D reguläre Fläche. Sei $\mathcal{F}$ das durch Φ gegebene reguläre Flächenstück und R sein Rand. Weiter sei γ eine Kurve, die einmal gegen den Uhrzeigersinn entlang R läuft. Schließlich sei $U \subseteq \mathbb{R}^3$ offen mit $\mathcal{F} \subseteq U$ und $F \colon U \to \mathbb{R}^3$ stetig differenzierbar. Dann gilt

$$\underbrace{\int_{\mathcal{F}} \langle \operatorname{rot} F, \vec{\nu}(\vec{x})\rangle \,\mathrm{d}\sigma(\vec{x})}_{\text{Oberflächenintegral}} = \underbrace{\int_{\gamma} F(\vec{x})\,\mathrm{d}\vec{x}}_{\text{Wegintegral}}.$$

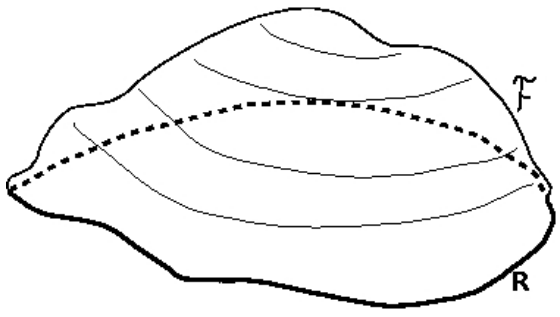


Abbildung 63: Ein Flächenstück $\mathcal{F}$ mit seinem Rand R (untere fette Linie).

22.17 Beispiel: Wir betrachten erneut $\mathcal{F}$ als obere Halbkugel mit Radius R. Wie üblich sei also

$$G = \left\{ \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^2 \colon u^2 + v^2 < R^2 \right\},$$

und die Fläche darüber sei

$$\Phi(u,v) = \begin{bmatrix} u \\ v \\ \sqrt{R^2 - u^2 - v^2} \end{bmatrix}.$$

Wir nehmen als Vektorfeld

$$F(x,y,z) = \begin{bmatrix} -y \\ x \\ 1 \end{bmatrix},$$

und darauf wenden wir den Satz von Stokes an. Es ist also

$$\operatorname{rot} F(\vec{x}) = \begin{bmatrix} \frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z} \\ \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x} \\ \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \end{bmatrix} = \begin{bmatrix} 0-0 \\ 0-0 \\ 1+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

Wir rechnen jetzt das Oberflächenintegral aus:

$$\begin{aligned}
\int_{\mathcal{F}} \langle \operatorname{rot} F(\vec{x}), \vec{\nu}(\vec{x}) \rangle \, \mathrm{d}\sigma(\vec{x}) &= \int_G \langle \operatorname{rot} F(\Phi(u,v)), \vec{\nu}(\Phi(u,v)) \rangle \, \|\partial_u \Phi(u,v) \times \partial_v \Phi(u,v)\| \, \mathrm{d}(u,v) \\
&= \int_G \left\langle \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \frac{\partial_u \Phi(u,v) \times \partial_v \Phi(u,v)}{\|\partial_u \Phi(u,v) \times \partial_v \Phi(u,v)\|} \right\rangle \|\partial_u \Phi(u,v) \times \partial_v \Phi(u,v)\| \, \mathrm{d}(u,v) \\
&= \int_G \left\langle \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} u/\sqrt{R^2 - u^2 - v^2} \\ v/\sqrt{R^2 - u^2 - v^2} \\ 1 \end{bmatrix} \right\rangle \mathrm{d}(u,v) \\
&= \int_G 2 \, \mathrm{d}(u,v) = \int_{\varphi=0}^{2\pi} \int_{r=0}^{R} 2r \, \mathrm{d}r \, \mathrm{d}\varphi = 2\pi \cdot r^2 \Big|_{r=0}^{r=R} = 2\pi R^2.
\end{aligned}$$

Und jetzt kommen wir zur Berechnung des Wegintegrals: Der Rand von $\mathcal{F}$ ist der Äquator und demnach gegeben durch

$$\gamma(t) = \begin{bmatrix} R\cos(t) \\ R\sin(t) \\ 0 \end{bmatrix}, \qquad t \in [0, 2\pi],$$

und somit folgt dann

$$\begin{aligned}
\int_\gamma F(\vec{x}) \, \mathrm{d}\vec{x} &= \int_{t=0}^{2\pi} \langle F(\gamma(t)), \gamma'(t) \rangle \, \mathrm{d}t \qquad \Big| \quad F(x,y,z) = [-y, x, 1]^T \\
&= \int_{t=0}^{2\pi} \left\langle \begin{bmatrix} -R\sin(t) \\ R\cos(t) \\ 1 \end{bmatrix}, \begin{bmatrix} -R\sin(t) \\ R\cos(t) \\ 0 \end{bmatrix} \right\rangle \mathrm{d}t \\
&= \int_{t=0}^{2\pi} \left(R^2 \sin^2(t) + R^2 \cos^2(t) \right) \mathrm{d}t = 2\pi R^2.
\end{aligned}$$

Das wollten wir haben.