

Höhere Mathematik II

Sommersemester 2025

R. Haller, S. Roch

Version vom 20. August 2025

Inhaltsverzeichnis

1	Matrixrechnung und Eigenwerttheorie	1
1.1	Wiederholung: Matrizen, Gaußverfahren und Determinanten	1
1.2	Eigenwerte und Eigenvektoren	12
2	Mehrdimensionale Differentialrechnung	25
2.1	Einleitung	25
2.2	Folgen und Stetigkeit	27
2.3	Ableiten von reellwertigen Funktionen	33
2.4	Ableiten von Vektor-wertigen Funktionen	39
2.5	Rechnen mit Ableitungen	42
2.6	Bestimmen von Extrema in mehreren Variablen	45
2.7	Extremwertprobleme mit Nebenbedingung	52
3	Mehrdimensionale Integration	59
3.1	Zweidimensionale Integration	59
3.2	Dreidimensionale Integration	66
3.3	Kurvenintegrale	71
4	Gewöhnliche Differentialgleichungen	77
4.1	Erste Beispiele und Begriffe	77
4.2	Differentialgleichungen mit getrennten Variablen	80
4.3	Lineare Differentialgleichungen 1. Ordnung	83
4.4	Lineare Differentialgleichungen 2. Ordnung	86
	Index	91

1 Matrixrechnung und Eigenwerttheorie

Dieses Kapitel schließt sich dem Kapitel „Elemente der linearen Algebra“ der Vorlesung *Höhere Mathematik I* an. Es beginnt mit einer Zusammenfassung der dortigen Erkenntnisse über Matrizen, lineare Gleichungssysteme und Determinanten.

1.1 Wiederholung: Matrizen, Gaußverfahren und Determinanten

1.1.1 Lineare Gleichungssysteme und Matrizen

Definition 1.1 (Matrix). Sind a_{ij} für $i = 1, 2, \dots, n$ und $j = 1, 2, \dots, m$ reelle Zahlen, so nennt man das Schema mit n Zeilen und m Spalten

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} \end{pmatrix}$$

eine $n \times m$ -Matrix. Mit $M_{n,m}(\mathbb{R})$ oder $\mathbb{R}^{n \times m}$ bezeichnen wir die Menge aller $n \times m$ -Matrizen.

Beispiel 1.2.

$$2 \times 3\text{-Matrix: } \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad 3 \times 3\text{-Matrix: } \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Definition 1.3 (Koeffizientenmatrix). Die Koeffizienten des linearen Gleichungssys-

1 Matrixrechnung und Eigenwerttheorie

tems (LGS)

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m &= b_n \end{aligned} \tag{1.1}$$

bilden die Koeffizientenmatrix

$$A := \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} \end{pmatrix}$$

des Systems.

Führen wir die (Spalten-)Vektoren $b := (b_1, \dots, b_n)^T$ und $x = (x_1, \dots, x_m)^T$ ein, so können wir mit der Koeffizientenmatrix das Gleichungssystem in (1.1) als $Ax = b$ mit der Unbekannten $x \in \mathbb{R}^m$ schreiben. Das Gleichungssystem wird dann durch die *erweiterte Koeffizientenmatrix*

$$(A \mid b) := \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ a_{31} & a_{32} & \dots & a_{3m} & b_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{array} \right)$$

dargestellt. Diese kann mit dem weiter unten erklärten Gauß-Verfahren immer so weit vereinfacht werden, dass die Matrix in sogenannter Zeilenstufenform vorliegt.

Definition 1.4. Eine Matrix ist in Zeilenstufenform, wenn jede Zeile entweder

- mehr aufeinanderfolgende Nullen am Zeilenanfang hat als die direkt darüber stehende Zeile oder
- eine Nullzeile ist.

Der erste von 0 verschiedene Eintrag einer Zeile wird auch als Pivotelement bezeichnet.

Eine Matrix in Zeilenstufenform hat damit die folgende Form, in der Sterne beliebige reelle Einträge und die Einträge \mathbf{X} reelle Zahlen ungleich Null darstellen. Letztere sind

1.1 Wiederholung: Matrizen, Gaußverfahren und Determinanten

genau die Pivotelemente.

$$\begin{pmatrix} 0 & \dots & 0 & \mathbf{X} & * & \dots & * & * & * & \dots & * & \dots & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & \mathbf{X} & * & \dots & * & \dots & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & \mathbf{X} & \dots & * & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \mathbf{X} & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & \dots & 0 \end{pmatrix}$$

Beispiel 1.5. Die in den folgenden Matrizen eingerahmten Zahlen sind jeweils die Pivotelemente.

$$\begin{pmatrix} 0 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 4 & 5 \end{pmatrix} \quad \text{Keine Zeilenstufenform}$$

$$\begin{pmatrix} \boxed{1} & 2 & 3 & 4 & 5 \\ 0 & \boxed{2} & 3 & 4 & 5 \\ 0 & 0 & 0 & \boxed{4} & 5 \end{pmatrix} \quad \text{Zeilenstufenform}$$

$$\begin{pmatrix} 0 & 1 & 3 & 4 & 5 \\ 0 & 1 & 3 & 4 & 5 \\ 0 & 1 & 0 & 4 & 5 \end{pmatrix} \quad \text{Keine Zeilenstufenform}$$

$$\begin{pmatrix} 0 & \boxed{1} & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & \boxed{5} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{Zeilenstufenform}$$

Wir erinnern uns, dass es für eine Matrix in Zeilenstufenform besonders einfach ist, das zugehörige Gleichungssystem zu lösen.

Satz 1.6 (Lösbarkeit LGS). *Sei A die Koeffizientenmatrix und $(A \mid b)$ die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems. Ist A in Zeilenstufenform, so gilt:*

- (a) *Das Gleichungssystem hat **keine** Lösung, wenn für mindestens ein i die i -te Zeile von A eine Nullzeile und $b_i \neq 0$ ist.*
- (b) *Hat das Gleichungssystem eine Lösung (siehe (a)), und gibt es in jeder Spalte von A ein Pivotelement, dann besitzt das Gleichungssystem **genau eine** Lösung.*
- (c) *Hat das Gleichungssystem eine Lösung (siehe (a)), und existiert eine Spalte ohne Pivotelement, dann besitzt das Gleichungssystem **unendlich viele** Lösungen.*

1 Matrixrechnung und Eigenwerttheorie

Beispiel 1.7. (a) Das LGS

$$(A | b) := \left(\begin{array}{ccccc|c} \boxed{1} & 2 & 3 & 4 & 5 & 2 \\ 0 & \boxed{2} & 4 & 1 & 5 & 1 \\ 0 & 0 & 0 & \boxed{1} & 5 & 3 \end{array} \right)$$

hat unendlich viele Lösungen, denn in der dritten und fünften Spalte von A gibt es kein Pivotelement.

(b) Das LGS

$$(A | b) := \left(\begin{array}{ccccc|c} \boxed{1} & 2 & 3 & 4 & 5 & 2 \\ 0 & \boxed{2} & 3 & 4 & 5 & 1 \\ 0 & 0 & 0 & \boxed{4} & 5 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

hat keine Lösung, denn die vierte Zeile von A ist eine Nullzeile und $b_4 = 1 \neq 0$.

(c) Das LGS

$$(A | b) := \left(\begin{array}{ccccc|c} \boxed{1} & 2 & 3 & 4 & 5 & 2 \\ 0 & \boxed{2} & 3 & 4 & 5 & 1 \\ 0 & 0 & \boxed{3} & 4 & 5 & 1 \\ 0 & 0 & 0 & \boxed{4} & 5 & 1 \\ 0 & 0 & 0 & 0 & \boxed{5} & 1 \end{array} \right)$$

hat eine eindeutige Lösung, denn in allen Spalten von A gibt es Pivotelemente.

Methode 1.8 (Lösen eines LGS in Zeilenstufenform). Sei A die Koeffizientenmatrix und $(A | b)$ die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems $Ax = b$ mit unbekanntem x . Ist A in Zeilenstufenform, so können wir die Lösungsmenge L des Gleichungssystems durch folgende Schritte bestimmen:

- Ist für ein i die i -te Zeile von A eine Nullzeile und gleichzeitig $b_i \neq 0$, dann gilt $L = \emptyset$.
- Für jede Spalte von A ohne Pivotelement führen wir einen freien Parameter t_j ein. Bezeichnet k die Anzahl solcher Spalten, erhalten wir also freie Parameter t_1, \dots, t_k .
- Für jede von einer Nullzeile verschiedene Zeile von $(A | b)$ schreiben wir die zugehörige lineare Gleichung auf und ersetzen dabei für jeden freien Parameter die Unbekannte x_j durch t_j . Für die i -te Zeile und im Fall von einem freien Parameter t_j ergibt das also beispielsweise

$$a_{i1}x_1 + \dots + a_{ij-1}x_{j-1} + a_{ij}t_j + a_{ij+1}x_{j+1} + \dots + a_{im}x_m = b_i. \quad (1.2)$$

1.1 Wiederholung: Matrizen, Gaußverfahren und Determinanten

- (d) Fangen wir von unten an, so können wir in den Gleichungen (1.2) durch iterative Rücksubstitutionen die Unbekannten x_i , die nicht durch freie Parameter ersetzt wurden, der Reihe nach mittels der freien Parameter ausdrücken.

Beispiel 1.9. Beispielfür das Ablesen der Lösungsmenge beim Vorliegen von freien Parametern schauen wir uns das LGS aus Beispiel 1.7 (a) an:

$$(A | b) := \left(\begin{array}{ccccc|c} \boxed{1} & 2 & 3 & 4 & 5 & 2 \\ 0 & \boxed{2} & 4 & 1 & 5 & 1 \\ 0 & 0 & 0 & \boxed{1} & 5 & 3 \end{array} \right).$$

Es gibt dort die beiden freien Parameter $x_3 = s$ und $x_5 = t$. Die dritte Zeile liefert dann $x_4 + 5t = 3$, also $x_4 = 3 - 5t$. Geht man damit in die zweite Zeile, bekommt man

$$\begin{aligned} 1 &= 2x_2 + 4x_3 + x_4 + 5x_5 = 2x_2 + 4s + (3 - 5t) + 5t = 2x_2 + 4s + 3 \\ \iff 2x_2 &= -2 - 4s \iff x_2 = -1 - 2s. \end{aligned}$$

In gleicher Weise liefert dann die erste Zeile

$$\begin{aligned} 2 &= x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = x_1 + 2(-1 - 2s) + 3s + 4(3 - 5t) + 5t \\ &= x_1 - s - 15t + 10 \iff x_1 = -8 + s + 15t. \end{aligned}$$

Zusammen ergibt sich damit die Lösungsmenge als

$$\begin{aligned} L &= \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \middle| Ax = b \right\} = \left\{ \begin{pmatrix} -8 + s + 15t \\ -1 - 2s \\ s \\ 3 - 5t \\ t \end{pmatrix} \middle| s, t \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} -8 \\ -1 \\ 0 \\ 3 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 15 \\ 0 \\ 0 \\ -5 \\ 1 \end{pmatrix} \middle| s, t \in \mathbb{R} \right\}. \end{aligned}$$

Wir betrachten nun ein beliebiges lineares Gleichungssystem mit erweiterter Koeffizientenmatrix $(A | b)$ und beantworten die Frage, wie man dieses auf Zeilenstufenform bringt, damit Methode 1.8 angewendet werden kann. Dazu erinnern wir uns an den *Gauß-Algorithmus*. Dieser verwendet die folgenden Umformungen

Definition 1.10 (Elementare Matrixumformungen). Sei $A \in M_{n,m}(\mathbb{R})$ eine Matrix. Wir bezeichnen die folgenden Umformungen von A als elementar:

(G1) Zwei Zeilen von A werden vertauscht: $\boxed{Z_j \leftrightarrow Z_k}$.

1 Matrixrechnung und Eigenwerttheorie

(G2) Multiplikation einer Zeile von A mit einer Zahl $c \neq 0$: $cZ_j \rightarrow Z_j$.

(G3) Zu einer Zeile von A wird das Vielfache einer anderen Zeile addiert:

$$Z_j + cZ_k \rightarrow Z_j \text{ für } j \neq k.$$

Tatsächlich lässt sich durch mehrfaches Ausführen der elementaren Umformungen (G1)–(G3) jede Matrix $A \in M_{n,m}(\mathbb{R})$ auf Zeilenstufenform bringen.

Methode 1.11 (Gauß-Algorithmus). Sei $A \in M_{n,m}(\mathbb{R})$. Um A auf Zeilenstufenform zu bringen, führt man die folgenden Schritte nacheinander für $i = 1$ (erste Zeile), dann für $i = 2$ (zweite Zeile) usw. aus:

- Falls ab der i -ten Zeile nur noch Nullzeilen auftreten, beende die Durchführung.
- Tausche sofern nötig die i -te Zeile mit einer der darunter stehenden Zeilen, so dass die Anzahl der führenden Nullen in der (neuen) i -ten Zeile kleinstmöglich wird.
- Hat die $(i+1)$ -te Zeile ein Pivotelement d , das unter dem Pivotelement c der i -ten Zeile steht, so addiere das $\frac{-d}{c}$ -fache der i -ten Zeile zur $(i+1)$ -ten Zeile, sodass die Zahl unter dem Pivotelement der i -ten Zeile nachher gleich 0 ist; verfähre genauso nicht nur für die $(i+1)$ -te Zeile, sondern auch für die $(i+2)$ -te Zeile, die $(i+3)$ -te Zeile und so weiter. (Es wird sich jedoch immer auf das Pivotelement der i -ten Zeile bezogen!)

Der entscheidende Punkt ist, dass elementare (Zeilen-)Umformungen der erweiterten Koeffizientenmatrix die Lösungsmenge des Gleichungssystems unverändert lassen. Man kann also den Gauß-Algorithmus auf die erweiterte Koeffizientenmatrix $(A \mid b)$ eines LGS anwenden, um die Matrix A auf Zeilenstufenform zu bringen und dann die Lösungsmenge des LGS anhand Methode 1.8 bestimmen.

Methode 1.12 (Lösen eines LGS mit dem Gauß-Algorithmus). Sei A die Koeffizientenmatrix und $(A \mid b)$ die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems $Ax = b$ mit unbekanntem x .

- Bringe mit dem Gauß-Algorithmus aus Methode 1.11 die Matrix $(A \mid b)$ auf Zeilenstufenform.
- Bestimme die Lösungsmenge L dieses Systems mit Methode 1.8.

1.1.2 Inverse Matrix und Determinanten

Als entscheidende Eigenschaft von Matrizen hatten wir erkannt, dass man mit ihnen (fast) wie mit Zahlen rechnen kann. Wir können Matrizen gleicher Größe addieren und, wenn die Ausmaße der Matrizen richtig zusammenpassen auch zwei Matrizen multiplizieren. Für quadratische Matrizen $A \in M_{n,n}(\mathbb{R})$, d. h. gleich viele Zeilen und

1.1 Wiederholung: Matrizen, Gaußverfahren und Determinanten

Spalten, kann man manchmal auch die sogenannte inverse Matrix A^{-1} bilden und damit auch durch Matrizen „teilen“. Dazu gibt es die folgenden Begriffe.

Definition 1.13 (Einheitsmatrix und Inverse). (a) Die Matrix

$$I_n := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

heißt Einheitsmatrix.

(b) Eine Matrix $A \in M_{n,n}(\mathbb{R})$ heißt invertierbar oder regulär, wenn es eine Matrix $C \in M_{n,n}(\mathbb{R})$ gibt mit der Eigenschaft

$$AC = CA = I_n.$$

Die Matrix C heißt dann inverse Matrix von A und wird mit A^{-1} bezeichnet. Ist eine Matrix $A \in M_{n,n}(\mathbb{R})$ nicht invertierbar, so heißt sie singulär.

Methode 1.14 (Berechnung von A^{-1}). (a) Betrachte das LGS $AX = I_n$ mit unbekannter Matrix X und der Einheitsmatrix I_n als rechter Seite, d. h. die erweiterte Koeffizientenmatrix $(A \mid I_n)$,

(b) Wende die elementaren Umformungen (G1)–(G3) des Gauß–Verfahrens aus Definition 1.10 an, bis das resultierende LGS die erweiterte Koeffizientenmatrix $(I_n \mid C)$ mit einer Matrix C hat. Dann ist C die inverse Matrix A^{-1} von A .

Einige wichtige Eigenschaften einer Matrix lassen sich bereits an einer einzigen Zahl ablesen – ihrer Determinante. Sie haben in der Höheren Mathematik I bereits gesehen wie man diese für 2×2 - und 3×3 -Matrizen bestimmt. Für 2×2 -Matrizen nach

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

und im Fall von 3×3 -Matrizen nach der *Regel von Sarrus*

$$\begin{aligned} \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} &= \begin{array}{c} \left| \begin{array}{ccc|cc} a_1 & b_1 & c_1 & a_1 & b_1 \\ & \searrow & \times & \times & \nearrow \\ a_2 & b_2 & c_2 & a_2 & b_2 \\ & \nearrow & \times & \times & \searrow \\ a_3 & b_3 & c_3 & a_3 & b_3 \end{array} \right| \\ = \boxed{\sum \searrow - \sum \nearrow} \\ = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_3b_2c_1 - b_3c_2a_1 - c_3a_2b_1. \end{array} \end{aligned}$$

1 Matrixrechnung und Eigenwerttheorie

Wir schauen uns jetzt weitere Berechnungsmöglichkeiten für die Determinante an, die insbesondere auch Matrizen mit mehr als drei Zeilen und Spalten zugänglich machen. Der Schlüssel dazu ist die folgende Entwicklungsformel, die die Berechnung einer $n \times n$ -Determinante auf die Berechnung von bis zu n kleineren Determinanten der Größe $(n-1) \times (n-1)$ zurückführt.

Definition 1.15 (Entwicklung von Determinanten). (a) Sei $A \in M_{n,n}(\mathbb{R})$. Streichen wir aus A die i -te Zeile und j -te Spalte, so erhalten wir die Streichungsmatrix $A'_{ij} \in M_{n-1,n-1}(\mathbb{R})$.

$$A'_{ij} := \begin{pmatrix} a_{11} & \dots & a_{1j-1} & a_{1j} & a_{1j+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2j-1} & a_{2j} & a_{2j+1} & \dots & a_{2n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{i-11} & \dots & a_{i-1j-1} & a_{i-1j} & a_{i-1j+1} & \dots & a_{i-1n} \\ a_{i+11} & \dots & a_{i+1j-1} & a_{i+1j} & a_{i+1j+1} & \dots & a_{i+1n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{n-11} & \dots & a_{n-1j-1} & a_{n-1j} & a_{n-1j+1} & \dots & a_{n-1n} \\ a_{n1} & \dots & a_{nj-1} & a_{nj} & a_{nj+1} & \dots & a_{nn} \end{pmatrix}.$$

(b) Ist $A \in M_{n,n}(\mathbb{R})$, so ist

$$\det A := \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A'_{ij} \quad (\text{Entwicklungsformel für die } i\text{-te Zeile})$$

für jedes $i = 1, 2, \dots, n$.

Obige Formel sieht wahrscheinlich unverständlich und abschreckend aus. Die Ausführung ist vielleicht ein wenig mühsam aber doch recht einprägsam. Man sucht sich eine Zeile Nummer i der Matrix aus (sinnvollerweise die mit den meisten Nullen) geht diese Zeile durch und schreibt für jeden Eintrag a_{ij} dieser Zeile hin:

- Ein wechselndes Vorzeichen $(-1)^{i+j}$, das beim ersten Eintrag in ungeraden Zeilen mit Plus und sonst mit Minus startet (Schachbrettmuster) und dann bei jedem weiteren Eintrag wechselt mal
- Den gerade betrachteten Eintrag a_{ij} mal
- Die Determinante der Streichungsmatrix A'_{ij} , also der Matrix die durch Streichen der Zeile und Spalte entsteht, die sich im gerade betrachteten Eintrag schneiden.

Dann summiert man alle diese Produkte auf.

Beispiel 1.16. (a) Durch Entwicklung jeweils nach der ersten Zeile bekommt man leicht für jedes $n \in \mathbb{N}$

$$\det(I_n) = (-1)^2 \cdot 1 \cdot \det(I_{n-1}) + 0 = \det(I_{n-1}) = \dots = \det(I_2) = 1 \cdot 1 - 0 \cdot 0 = 1.$$

1.1 Wiederholung: Matrizen, Gaußverfahren und Determinanten

(b) Tatsächlich liefert die Regel von Sarrus

$$\det \begin{pmatrix} 1 & 2 & 0 \\ 4 & 5 & 6 \\ 0 & 2 & 3 \end{pmatrix} = \begin{array}{c|cc|cc} 1 & 2 & 0 & 1 & 2 \\ \searrow & \times & \times & \nearrow & \\ 4 & 5 & 6 & 4 & 5 \\ \nearrow & \times & \times & \searrow & \\ 0 & 2 & 3 & 0 & 2 \end{array} = 15 - 12 - 24 = -21$$

dasselbe Ergebnis wie die Entwicklung nach der ersten Zeile

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 0 \\ 4 & 5 & 6 \\ 0 & 2 & 3 \end{pmatrix} &= (-1)^{1+1} \cdot 1 \cdot \det \begin{pmatrix} 5 & 6 \\ 2 & 3 \end{pmatrix} + (-1)^{1+2} \cdot 2 \cdot \det \begin{pmatrix} 4 & 6 \\ 0 & 3 \end{pmatrix} + 0 \\ &= 15 - 12 - 2 \cdot 12 = -21. \end{aligned}$$

Die folgenden beiden Rechenregeln für Determinanten sind oft praktisch und in keiner Weise offensichtlich.

Rechenregel 1.17 (für Determinanten). Es seien $A, B \in M_{n,n}(\mathbb{R})$. Dann gilt

- (a) $\det(A^T) = \det(A)$,
- (b) $\det(AB) = \det(A) \cdot \det(B)$.

Bemerkung 1.18. Die Rechenregel $\det(A^T) = \det(A)$ bedeutet, dass man in einer Matrix die Spalten zu Zeilen machen kann und umgekehrt ohne die Determinante zu ändern. Insbesondere kann man damit jede Rechnung, die man in den Zeilen macht, genauso auch in den Spalten machen. Das ist oft sehr praktisch. Beispielsweise gilt die Entwicklungsformel automatisch auch für ein Entwickeln nach den Spalten der Matrix:

$$\det A := \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A'_{ij} \quad (\text{Entwicklungsformel für die } j\text{-te Spalte})$$

für $j = 1, 2, \dots, n$.

Beispiel 1.19. Tatsächlich bekommen wir für die Matrix aus Beispiel 1.16 (b) zum dritten Mal dasselbe Ergebnis heraus, wenn wir die Determinante beispielsweise nach der dritten Spalte entwickeln:

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 0 \\ 4 & 5 & 6 \\ 0 & 2 & 3 \end{pmatrix} &= 0 + (-1)^{2+3} \cdot 6 \cdot \det \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} + (-1)^{3+3} \cdot 3 \cdot \det \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \\ &= -12 - 9 = -21. \end{aligned}$$

Das Verfahren zum Entwickeln bleibt genau wie nach Definition 1.15 beschrieben und auch die Vorzeichen gehen nach demselben Schachbrettmuster.

1 Matrixrechnung und Eigenwerttheorie

Die zweite Rechenregel in 1.17 (b) liefert für jede invertierbare Matrix sofort

$$1 = \det(I_n) = \det(AA^{-1}) = \det(A) \cdot \det(A^{-1}).$$

Also kann für eine invertierbare Matrix A die Determinante nicht Null sein und es gilt dann $\det(A^{-1}) = 1/\det(A)$. Tatsächlich gilt der folgende Satz, der insbesondere besagt, dass man die Determinante nutzen kann, um zu überprüfen, ob eine gegebene Matrix invertierbar ist, ohne gleich die Inverse ausrechnen zu müssen.

Satz 1.20. *Eine Matrix $A \in M_{n,n}(\mathbb{R})$ ist genau dann invertierbar, wenn $\det(A) \neq 0$ ist. In diesem Fall gilt*

$$\det(A^{-1}) = \frac{1}{\det(A)} = \det(A)^{-1}.$$

Rechenregel 1.21 (Elementarumformungen bei Determinanten). Die Elementarumformungen (G1)–(G3) des Gauß-Verfahrens aus Definition 1.10 können auch benutzt werden, um die Berechnung von Determinanten zu vereinfachen, denn sie ändern den Wert der Determinante gar nicht bis wenig. Im Einzelnen gelten die folgenden Regeln für die Determinante einer Matrix $A \in M_{n,n}(\mathbb{R})$:

- (a) Vertauscht man nach (G1) zwei Zeilen von A , so wird die Determinante mit (-1) multipliziert.
- (b) Multipliziert man nach (G2) eine Zeile von A mit einer Zahl $c \neq 0 \in \mathbb{R}$, so ändert sich auch die Determinante um den Faktor c .
- (c) Addiert man nach (G3) das Vielfache einer Zeile zu einer anderen Zeile, so ändert sich die Determinante nicht.
- (d) Alle drei bisherigen Regeln gelten genauso auch für Spalten statt Zeilen.
- (e) Ist $c \in \mathbb{R}$, so gilt

$$\det(cA) = c^n \det(A).$$

Die Begründung der ersten drei Regeln ergibt sich jeweils aus der Entwicklungsformel. Hier folgen kurz die Ideen für die Begründungen aller Regeln:

- (a) Der Entwicklungsformel kann man entnehmen, dass das Vertauschen zweier benachbarter Zeilen, die Determinante um den Faktor (-1) verändert, denn nach der Vertauschung werden beim Entwickeln Summanden addiert, die sich nach dem Schachbrettmuster genau im Vorzeichen unterscheiden: Vertauscht man Zeile i mit Zeile $i+1$, so erhält man für die Determinante durch eine Entwicklung

1.1 Wiederholung: Matrizen, Gaußverfahren und Determinanten

nach Zeile $i + 1$:

$$\det \begin{pmatrix} \text{---} & a_1 & \text{---} \\ & \vdots & \\ \text{---} & a_{i+1} & \text{---} \\ \text{---} & a_i & \text{---} \\ & \vdots & \\ \text{---} & a_n & \text{---} \end{pmatrix} = \sum_{j=1}^n (-1)^{i+1+j} a_{ij} A'_{ij}$$

$$= - \sum_{j=1}^n (-1)^{i+j} a_{ij} A'_{ij} = - \det A.$$

Für den allgemeinen Fall überlegt man sich dann, dass man die Vertauschung von zwei beliebigen Zeilen immer durch eine *ungerade* Anzahl von Vertauschen von benachbarten Zeilen erreicht. Also bleibt immer ein Faktor (-1) übrig.

(b) Eine Entwicklung nach der i -ten Zeile, die mit c multipliziert wurde, ergibt direkt

$$\det \begin{pmatrix} \text{---} & a_1 & \text{---} \\ & \vdots & \\ \text{---} & ca_i & \text{---} \\ & \vdots & \\ \text{---} & a_n & \text{---} \end{pmatrix} = \sum_{j=1}^n (-1)^{i+j} ca_{ij} A'_{ij} = c \sum_{j=1}^n (-1)^{i+j} a_{ij} A'_{ij} = c \det A.$$

(c) Wir ersetzen die Zeile a_i durch die neue Zeile $a_i + ca_k$. Dann ergibt eine Entwicklung nach der i -ten Zeile

$$\det \begin{pmatrix} \text{---} & a_1 & \text{---} \\ & \vdots & \\ \text{---} & a_{i-1} & \text{---} \\ \text{---} & a_i + ca_k & \text{---} \\ \text{---} & a_{i+1} & \text{---} \\ & \vdots & \\ \text{---} & a_n & \text{---} \end{pmatrix} = \sum_{j=1}^n (-1)^{i+j} (a_{ij} + ca_{ik}) A'_{ij}$$

$$= \sum_{j=1}^n (-1)^{i+j} a_{ij} A'_{ij} + c \sum_{j=1}^n (-1)^{i+j} a_{ik} A'_{ij}$$

$$= \det A + c \det C,$$

wobei man die Matrix $C \in M_{n,n}(\mathbb{R})$ erhält, indem man in A die i -te Zeile durch a_i ersetzt. Damit sind zwei verschiedene Zeilen von C identisch und, indem wir diese beiden Zeilen vertauschen, bekommen wir aus der Regel in (a), dass $\det(C) = -\det(C)$ ist, was nichts anderes bedeutet als $\det(C) = 0$. Also bleibt in der Rechnung $\det(A)$ übrig.

1 Matrixrechnung und Eigenwerttheorie

- (d) Das liegt wieder an $\det(A) = \det(A^T)$, vgl. Bemerkung 1.18.
- (e) Die Matrix cA entsteht aus A , indem *jede* Zeile von A mit dem Faktor c multipliziert wird. Nach der Regel in (b) bekommt man also für jede Zeile einen Faktor c .

Bemerkung 1.22 (Determinantenberechnung durch Gauß–Algorithmus). Für $n > 3$ ist für die Berechnung der Determinante meist eine Kombination aus Elementarumformungen nach den Regeln 1.21 und Entwicklung, vgl. Definition 1.15, sinnvoll. Das Ziel ist es dabei, die Matrix A durch die Elementarumformungen aus dem Gauß–Verfahren auf Zeilenstufenform (siehe Definition 1.4) zu bringen.

Für eine Matrix B in Zeilenstufenform

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & \cdots & b_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{nn} \end{pmatrix},$$

gilt nämlich

$$\det B = b_{11} \cdot b_{22} \cdots b_{nn},$$

indem man nacheinander immer wieder nach der ersten Zeile entwickelt. So eine Matrix, die nur auf und oberhalb der Diagonalen von Null verschiedene Einträge hat, nennt man auch *obere Dreiecksmatrix*

Beispiel 1.23. Wir bestimmen auf diese Weise die Determinante von A für

$$A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 5 & -3 \\ -3 & 2 & -4 \end{pmatrix}.$$

Es gilt

$$\begin{pmatrix} 1 & 3 & -2 \\ 2 & 5 & -3 \\ -3 & 2 & -4 \end{pmatrix} \xrightarrow{\substack{z_2 - 2z_1 \\ z_3 + 3z_1}} \begin{pmatrix} 1 & 3 & -2 \\ 0 & -1 & 1 \\ 0 & 11 & -10 \end{pmatrix} \xrightarrow{z_3 + 11z_2} \begin{pmatrix} 1 & 3 & -2 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} =: B.$$

Wir haben somit die Matrix A durch drei Ausführungen vom Gaußschritt (G3) in B umgeformt. Da (G3) nach Regel 1.21 (c) die Determinante nicht ändert, folgt

$$\det A = \det B = 1 \cdot (-1) \cdot 1 = -1.$$

1.2 Eigenwerte und Eigenvektoren

Die Determinante einer Matrix ist wichtig für eine Vielzahl von Berechnungen, kann aber offensichtlich nur einen kleinen Teil der Information wiedergeben, die in der

Matrix enthalten ist. Weitere wichtige Eigenschaften einer Matrix lassen sich an ihren *Eigenwerten* ablesen.

Definition 1.24 (Eigenwert und Eigenvektor). Sei $A \in M_{n,n}(\mathbb{R})$ eine Matrix. Eine Zahl $\lambda \in \mathbb{R}$ heißt Eigenwert von A , wenn es einen Vektor $v \in \mathbb{R}^n \setminus \{0\}$ gibt, sodass gilt

$$Av = \lambda v.$$

Der Vektor v heißt dann Eigenvektor zum Eigenwert λ von A .

Eigenvektoren beschreiben solche Richtungen in \mathbb{R}^n , die durch Multiplikation mit der Matrix erhalten bleiben. In diese Richtungen wirkt die Matrix in der einfachst möglichen Weise, indem sie den Vektor um den Faktor λ streckt/staucht (bzw. bei negativem λ auch umklappt).

Beispiel 1.25.

$$\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \text{also ist } \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ Eigenvektor zum Eigenwert } 2.$$

$$\begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad \text{also ist } \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ kein Eigenvektor.}$$

Bemerkung 1.26. Es sei $A \in M_{n,n}(\mathbb{R})$ eine Matrix und $\lambda \in \mathbb{R}$ eine Zahl.

- (a) Für den Nullvektor $v = 0$ gilt immer $Av = A0 = 0 = \lambda 0 = \lambda v$. Das liefert also überhaupt keine Information über A . Deshalb wird der Nullvektor als Eigenvektor grundsätzlich ausgeschlossen.
- (b) Ist $v \in \mathbb{R}^n$ ein Eigenvektor von A (also insbesondere $v \neq 0$), so gilt für jedes $c \in \mathbb{R}$

$$A(cv) = c(Av) = c(\lambda v) = \lambda(cv).$$

Das bedeutet, dass auch der Vektor cv für jedes $c \neq 0$ ein Eigenvektor von A zum selben Eigenwert λ ist. Eigenvektoren sind also nicht eindeutig.

Wie bestimmt man nun die Eigenwerte und Eigenvektoren einer gegebenen Matrix $A \in M_{n,n}(\mathbb{R})$? Wenn $v \neq 0$ ein Eigenvektor von A zum Eigenwert λ ist, so gilt $Av = \lambda v$, d. h.

$$0 = Av - \lambda v = Av - \lambda I_n v = (A - \lambda I_n)v.$$

Das lineare Gleichungssystem $(A - \lambda I_n)v = 0$ hat dann also mit v und 0 zwei definitiv verschiedene Lösungen. Für die Matrix $A - \lambda I_n$ bedeutet das, dass sie nicht invertierbar sein kann. Umgekehrt wissen wir, wenn die Matrix $A - \lambda I_n$ nicht invertierbar ist, dass das LGS auch nicht eindeutig lösbar ist. Da es lösbar ist (Der Nullvektor ist eine Lösung.), muss es mehrere Lösungen geben und damit auch eine Lösung v , die nicht Null ist. Für diese gilt dann, indem man obige Rechnung rückwärts macht, $Av = \lambda v$, wir haben also einen Eigenvektor.

1 Matrixrechnung und Eigenwerttheorie

Die Eigenwerte sind damit genau die Zahlen, für die die Matrix $A - \lambda I_n$ nicht invertierbar ist. Nach Satz 1.20 ist das genau dann der Fall, wenn die Determinante dieser Matrix Null ist. Dieses Ergebnis halten wir fest.

Satz 1.27. *Ist $A \in M_{n,n}(\mathbb{R})$, so ist $\lambda \in \mathbb{R}$ genau dann ein Eigenwert von A , wenn $\det(A - \lambda I_n) = 0$ ist.*

Berechnet man die Determinante von $A - \lambda I_n$, so stellt man fest, dass diese immer ein Polynom vom Grad n ist. Wir haben damit also das Problem der Bestimmung von Eigenwerten auf das Finden der Nullstellen eines Polynoms reduziert. Als nächstes bekommt dieses Polynom seine übliche Bezeichnung.

Definition 1.28 (Charakteristisches Polynom). *Sei $A \in M_{n,n}(\mathbb{R})$. Dann heißt*

$$p : \mathbb{R} \rightarrow \mathbb{R}, \quad p(\lambda) := \det(A - \lambda I_n)$$

charakteristisches Polynom *der Matrix A .*

Definition 1.29 (Eigenraum). *Sei $A \in M_{n,n}(\mathbb{R})$ eine Matrix und $\lambda \in \mathbb{R}$ ein Eigenwert von A . Der Vektorraum*

$$E_\lambda := \ker(A - \lambda I_n) = \{v \in \mathbb{R}^n : Av = \lambda v\}$$

heißt Eigenraum von A zum Eigenwert λ .

Methode 1.30 (Bestimmung von Eigenwerten und Eigenvektoren). *Gegeben sei eine Matrix $A \in M_{n,n}(\mathbb{R})$ und gesucht sind alle Eigenwerte und die dazugehörigen Eigenvektoren in Form der Eigenräume.*

- (a) Man bestimmt das charakteristische Polynom $p(\lambda) = \det(A - \lambda I_n)$ der Matrix A . Dabei entsteht ein Polynom vom Grad n . Bei der Bestimmung lohnt es sich so zu entwickeln, dass Linearfaktoren der Form $(\lambda - \lambda_0)$ abgespalten werden. Denn diese liefern im nächsten Schritt ohne weitere Berechnungen Informationen zu Nullstellen.
- (b) Man bestimmt alle Nullstellen von p . Das sind die Eigenwerte.
- (c) Für jeden Eigenwert λ löst man das lineare Gleichungssystem $(A - \lambda I_n)x = 0$. Alle Lösungen x bilden zusammen den Eigenraum E_λ und alle Lösungen außer dem Nullvektor sind die Eigenvektoren zum Eigenwert λ .

Man beachte: Kommt in diesem Schritt nur der Nullvektor als Lösung des Systems heraus, ist irgendwo ein Fehler passiert. Entweder ist dann λ gar kein Eigenwert oder beim Lösen des Gleichungssystems stimmt etwas nicht. Nutzen Sie diesen eingebauten Kontrollmechanismus!

Beispiel 1.31. Wir bestimmen alle Eigenwerte und Eigenvektoren der Matrix

$$A := \begin{pmatrix} 2 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 3 & 1 \end{pmatrix} \in M_{3,3}(\mathbb{R}).$$

(a) Charakteristisches Polynom p berechnen:

$$p(\lambda) = \det(A - \lambda I_3) = \det \begin{pmatrix} 2 - \lambda & 1 & 0 \\ -1 & 0 - \lambda & 1 \\ 1 & 3 & 1 - \lambda \end{pmatrix} = \det \begin{pmatrix} 2 - \lambda & 1 & 0 \\ -1 & -\lambda & 1 \\ 1 & 3 & 1 - \lambda \end{pmatrix}$$

Addieren der dritten Spalte zur ersten ändert die Determinante nicht und liefert

$$= \det \begin{pmatrix} 2 - \lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 2 - \lambda & 3 & 1 - \lambda \end{pmatrix}.$$

Nun verwenden wir Rechenregel 1.21 (b), die es erlaubt, den Faktor $2 - \lambda$ aus der ersten Spalte herauszuziehen:

$$= (2 - \lambda) \det \begin{pmatrix} 1 & 1 & 0 \\ 0 & -\lambda & 1 \\ 1 & 3 & 1 - \lambda \end{pmatrix}.$$

Als nächstes wird die erste Zeile von der dritten abgezogen. Auch das ändert die Determinante nicht:

$$= (2 - \lambda) \det \begin{pmatrix} 1 & 1 & 0 \\ 0 & -\lambda & 1 \\ 0 & 2 & 1 - \lambda \end{pmatrix}.$$

Dann können wir schließlich nach der ersten Spalte entwickeln und die restliche 2×2 -Determinante leicht ausrechnen:

$$\begin{aligned} &= (2 - \lambda) \det \begin{pmatrix} -\lambda & 1 \\ 2 & 1 - \lambda \end{pmatrix} = (2 - \lambda)((-\lambda)(1 - \lambda) - 2 \cdot 1) \\ &= (2 - \lambda)(\lambda^2 - \lambda - 2). \end{aligned}$$

Man beachte, dass der Vorfaktor $2 - \lambda$ dabei ganz bewusst nicht angefasst wird. Insbesondere wäre es sehr ungeschickt diesen auszumultiplizieren, denn man kann hier direkt die erste Nullstelle ablesen!

(b) Nullstellen von p bestimmen: Es ist

$$\begin{aligned} p(\lambda) = 0 &\iff (2 - \lambda)(\lambda^2 - \lambda - 2) = 0 \\ &\iff (2 - \lambda) = 0 \quad \vee \quad \lambda^2 - \lambda - 2 = 0 \\ &\iff \lambda = 2 \quad \vee \quad \lambda = -1 \quad \vee \quad \lambda = 2. \end{aligned}$$

Wir haben also die zwei Eigenwerte $\lambda_1 = 2$ und $\lambda_2 = -1$.

1 Matrixrechnung und Eigenwerttheorie

- (c) Zur Bestimmung der Eigenvektoren müssen zwei lineare Gleichungssysteme gelöst werden. Wir beginnen mit $(A - 2I_3)x = 0$:

$$\begin{aligned} (A - 2I_3 \mid 0) &= \left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ -1 & -2 & 1 & 0 \\ 1 & 3 & -1 & 0 \end{array} \right) \xrightarrow{Z1 \leftrightarrow Z3} \left(\begin{array}{ccc|c} 1 & 3 & -1 & 0 \\ -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right) \\ &\xrightarrow{Z2+Z1} \left(\begin{array}{ccc|c} 1 & 3 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right) \xrightarrow{Z1-3 \cdot Z2} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right). \end{aligned}$$

Mit dem Parameter $x_3 = t$ folgt $x_2 = 0$ und $x_1 = x_3 = t$. Also gilt für den Eigenraum zum Eigenwert $\lambda_1 = 2$:

$$E_2 = \left\{ \begin{pmatrix} t \\ 0 \\ t \end{pmatrix} : t \in \mathbb{R} \right\} = \left\{ t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} : t \in \mathbb{R} \right\}.$$

Wir wenden uns dem zweiten Eigenwert $\lambda_2 = -1$ zu und lösen $(A - (-1)I_3)x = 0$:

$$\begin{aligned} (A + I_3 \mid 0) &= \left(\begin{array}{ccc|c} 3 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ 1 & 3 & 2 & 0 \end{array} \right) \xrightarrow{Z1 \leftrightarrow Z3} \left(\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ -1 & 1 & 1 & 0 \\ 3 & 1 & 0 & 0 \end{array} \right) \\ &\xrightarrow{\substack{Z2+Z1 \\ Z3-3 \cdot Z1}} \left(\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 0 & 4 & 3 & 0 \\ 0 & -8 & -6 & 0 \end{array} \right) \xrightarrow{Z3-2 \cdot Z2} \left(\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 0 & 4 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right). \end{aligned}$$

Mit dem Parameter $x_3 = t$ folgt $4x_2 = -3t$, also $x_2 = -3/4 \cdot t$ und mit der ersten Zeile:

$$0 = x_1 + 3x_2 + 2x_3 = x_1 - \frac{9}{4}t + 2t = x_1 - \frac{1}{4}t, \quad \text{d. h.} \quad x_1 = \frac{1}{4}t.$$

Zusammen ist damit

$$E_{-1} = \left\{ \begin{pmatrix} 1/4 \cdot t \\ -3/4 \cdot t \\ t \end{pmatrix} : t \in \mathbb{R} \right\} = \left\{ s \cdot \begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix} : s \in \mathbb{R} \right\},$$

wobei wir $s = t/4$ gesetzt haben, damit das Ergebnis schöner aussieht.

Definition 1.32. Seien $v_1, \dots, v_m \in \mathbb{R}^n$. Der Unterraum

$$\text{span}\{v_1, \dots, v_m\} := \{t_1v_1 + t_2v_2 + \dots + t_mv_m \mid t_1, \dots, t_m \in \mathbb{R}\} \subseteq \mathbb{R}^n$$

heißt *Spann oder lineare Hülle von v_1, \dots, v_m* .

Bemerkung 1.33. Eigenräume werden typischerweise als Spann von Eigenvektoren angegeben. Zum Beispiel die Eigenräume aus Beispiel 1.31:

$$E_2 = \left\{ t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} : t \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\},$$

$$E_{-1} = \left\{ s \cdot \begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix} : s \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} 1 \\ -3 \\ 4 \end{pmatrix} \right\}.$$

1.2.1 Diagonalisierung von Matrizen

Als eine Anwendung der Eigenwerte wollen wir uns zum Abschluss dieses Kapitels die Diagonalisierung von Matrizen anschauen, die beispielsweise nützlich ist bei der Berechnung von Potenzen einer Matrix.

Definition 1.34 (Diagonalisierbare Matrizen). (a) Eine Matrix $D \in M_{n,n}(\mathbb{R})$ der Gestalt

$$D = \begin{pmatrix} \mu_1 & 0 & 0 & \dots & 0 \\ 0 & \mu_2 & 0 & \dots & 0 \\ 0 & 0 & \mu_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \mu_n \end{pmatrix}$$

mit $\mu_1, \dots, \mu_n \in \mathbb{R}$ heißt Diagonalmatrix.

(b) Eine Matrix $A \in M_{n,n}(\mathbb{R})$ heißt diagonalisierbar, falls es eine invertierbare Matrix $S \in M_{n,n}(\mathbb{R})$ und eine Diagonalmatrix $D \in M_{n,n}(\mathbb{R})$ gibt mit $A = SDS^{-1}$.

Bevor es um die Frage geht, wie und ob man Matrizen diagonalisieren kann, d. h. wie man zu gegebenem vorgegebenem A die invertierbare Matrix S und die Diagonalmatrix D aus obiger Definition findet, schauen wir uns an, was Diagonalisierung mit der Bildung von Potenzen A^k einer Matrix A zu tun hat.

Bemerkung 1.35 (Potenzieren von diagonalisierbaren Matrizen). Sei A eine diagonalisierbare Matrix mit entsprechender invertierbarer Matrix S und Diagonalmatrix D , für die $A = SDS^{-1}$ gilt. Dann erhalten wir für $k \in \mathbb{N}$ und die Potenz A^k von A :

$$A^k = (SDS^{-1})^k = SDS^{-1}SDS^{-1}SDS^{-1} \dots SDS^{-1}SDS^{-1} = SD^kS^{-1}.$$

Was ist damit gewonnen? Die Potenzen einer Diagonalmatrix sind sehr leicht zu be-

1 Matrixrechnung und Eigenwerttheorie

stimmen, denn für

$$D = \begin{pmatrix} \mu_1 & 0 & 0 & \dots & 0 \\ 0 & \mu_2 & 0 & \dots & 0 \\ 0 & 0 & \mu_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \mu_n \end{pmatrix} \quad \text{gilt} \quad D^k = \begin{pmatrix} \mu_1^k & 0 & 0 & \dots & 0 \\ 0 & \mu_2^k & 0 & \dots & 0 \\ 0 & 0 & \mu_3^k & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \mu_n^k \end{pmatrix}.$$

Zur Berechnung der Potenz A^k muss man dann also nicht Unmengen Matrixmultiplikationen ausführen, sondern man muss nur alle Einträge von D als Zahlen potenzieren und dann noch zwei Matrixmultiplikationen machen um $A^k = SD^kS^{-1}$ auszurechnen.

Satz 1.36. Sei $A \in M_{n,n}(\mathbb{R})$. Sind v_1, \dots, v_n Eigenvektoren von A mit zugehörigen Eigenwerten $\lambda_1, \dots, \lambda_n$ und gilt $\text{span}\{v_1, v_2, \dots, v_n\} = \mathbb{R}^n$, so ist A diagonalisierbar mit

$$S := \left(\begin{array}{c|c|c|c} | & | & & | \\ v_1 & v_2 & \dots & v_n \\ | & | & & | \end{array} \right) \quad \text{und} \quad D = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}. \quad (1.3)$$

Wir glauben an dieser Stelle, dass aus der Bedingung $\text{span}\{v_1, \dots, v_n\} = \mathbb{R}^n$ folgt, dass S invertierbar ist, und rechnen nach, dass mit den so gewählten Matrizen S und D tatsächlich $A = SDS^{-1}$ gilt:

$$\begin{aligned} AS &= \left(\begin{array}{c|c|c|c} | & | & & | \\ Av_1 & Av_2 & \dots & Av_n \\ | & | & & | \end{array} \right) = \left(\begin{array}{c|c|c|c} | & | & & | \\ \lambda_1 v_1 & \lambda_2 v_2 & \dots & \lambda_n v_n \\ | & | & & | \end{array} \right) \\ &= \left(\begin{array}{c|c|c|c} | & | & & | \\ v_1 & v_2 & \dots & v_n \\ | & | & & | \end{array} \right) \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix} = SD. \end{aligned}$$

Multipliziert man diese Gleichheit von rechts mit S^{-1} folgt $A = SDS^{-1}$.

Damit ergibt sich folgende Methode zur Diagonalisierung von Matrizen.

Methode 1.37 (Diagonalisierung einer Matrix). Sei $A \in M_{n,n}(\mathbb{R})$ gegeben. Um festzustellen, ob A diagonalisierbar ist und S und D auszurechnen, kann man nach dem folgendem Schema vorgehen.

- (a) (Eigenwerte und Eigenvektoren) Bestimme alle Eigenwerte $\lambda_1, \dots, \lambda_m$ von A und die zugehörigen Eigenräume $E_{\lambda_1}, \dots, E_{\lambda_m}$ von A , vgl. Methode 1.30.

- (b) (Überprüfung Diagonalisierbarkeit) Hat man in (a) n verschiedene Eigenwerte (also genauso viele wie die Größe der Matrix) gefunden, ist A in jedem Fall diagonalisierbar und man kann gleich nach (c) springen. Hat man weniger als n verschiedene Eigenwerte, d. h. das charakteristische Polynom von A hatte mehrfache Nullstellen, so überprüfe man für jede mehrfache Nullstelle, ob die Vielfachheit der Nullstelle und die Anzahl der freien Parameter beim Lösen des LGS in der Bestimmung des zugehörigen Eigenraums übereinstimmen. Wenn das auch nur einmal schieft, ist die Matrix nicht diagonalisierbar.
- (c) (Bestimmung S und D) Wenn Sie hier angekommen sind, haben sie n Vektoren, die zusammen alle Eigenräume aufspannen. Diese bilden wie in (1.3) die Spalten der Matrix S und als D nehmen Sie die Diagonalmatrix, deren Eintrag in jeder Spalte genau der Eigenwert von A ist, der zum Eigenvektor in der entsprechenden Spalte von S gehört.

Beispiel 1.38. Wir diagonalisieren die Matrix $A := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

- (a) Zur Bestimmung der Eigenwerte benötigen wir zuerst das charakteristische Polynom von A :

$$\begin{aligned}
 p(\lambda) &= \det(A - \lambda I_n) = \det \begin{pmatrix} 1 - \lambda & 1 & 0 \\ 1 & 2 - \lambda & 1 \\ 0 & 1 & 1 - \lambda \end{pmatrix} \\
 &\stackrel{Z1-Z3}{=} \det \begin{pmatrix} 1 - \lambda & 0 & -(1 - \lambda) \\ 1 & 2 - \lambda & 1 \\ 0 & 1 & 1 - \lambda \end{pmatrix} = (1 - \lambda) \det \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 - \lambda & 1 \\ 0 & 1 & 1 - \lambda \end{pmatrix} \\
 &\stackrel{S3+S1}{=} (1 - \lambda) \det \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 - \lambda & 2 \\ 0 & 1 & 1 - \lambda \end{pmatrix}.
 \end{aligned}$$

Wir entwickeln nach der ersten Zeile und bekommen

$$\begin{aligned}
 &= (1 - \lambda) \det \begin{pmatrix} 2 - \lambda & 2 \\ 1 & 1 - \lambda \end{pmatrix} = (1 - \lambda)((2 - \lambda)(1 - \lambda) - 2) \\
 &= (1 - \lambda)(\lambda^2 - 3\lambda + 2 - 2) = (1 - \lambda)(\lambda^2 - 3\lambda) = \lambda(1 - \lambda)(\lambda - 3).
 \end{aligned}$$

Die Nullstellen, und damit die Eigenwerte von A , kann man nun sofort zu $\lambda_1 = 0$, $\lambda_2 = 1$ und $\lambda_3 = 3$ ablesen.

Als nächstes kommen die Eigenräume. Dazu lösen wir zunächst das lineare Gleichungssystem $(A - 0 \cdot I_3)x = 0$:

$$\left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array} \right) \xrightarrow{Z2-Z1} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array} \right) \xrightarrow{Z3-Z2} \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

1 Matrixrechnung und Eigenwerttheorie

Mit dem Parameter $t = x_3$ folgt $x_2 = -x_3 = -t$ und $x_1 = -x_2 = -(-t) = t$, also ist

$$E_0 = \left\{ \begin{pmatrix} t \\ -t \\ t \end{pmatrix} : t \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

Für $\lambda_2 = 1$ erhalten wir für das LGS $(A - 1 \cdot I_3)x = 0$:

$$\left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right) \xrightarrow{Z1 \leftrightarrow Z2} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right) \xrightarrow{\substack{Z3-Z2 \\ Z1-Z2}} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

Mit $x_3 = t$ folgt hier $x_2 = 0$ und $x_1 = -x_3 = -t$, also

$$E_1 = \left\{ \begin{pmatrix} -t \\ 0 \\ t \end{pmatrix} : t \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Schließlich ist für $\lambda_3 = 3$ mit dem LGS $(A - 3 \cdot I_3)x = 0$

$$\left(\begin{array}{ccc|c} -2 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 1 & -2 & 0 \end{array} \right) \xrightarrow{Z1 \leftrightarrow Z2} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 \end{array} \right) \xrightarrow{Z2+2 \cdot Z1} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 1 & -2 & 0 \end{array} \right) \\ \xrightarrow{\substack{Z1-Z2 \\ Z3+Z2}} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right).$$

Mit $x_3 = t$ folgt hier $x_2 = 2x_3 = 2t$ und $x_1 = x_3 = t$, also ist

$$E_3 = \left\{ \begin{pmatrix} t \\ 2t \\ t \end{pmatrix} : t \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}.$$

(b) Wir haben drei verschiedene Eigenwerte, also ist A diagonalisierbar.

(c) Nehmen wir die drei Eigenvektoren zusammen, so erhalten wir

$$S = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}$$

und mit der richtigen Zuordnung der Eigenwerte ist

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Wir haben also

$$A = SDS^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix}^{-1}$$

Berechnet man noch die Inverse von S , vgl. Methode 1.14, so findet man

$$S^{-1} = \frac{1}{6} \begin{pmatrix} 2 & -2 & 2 \\ -3 & 0 & 3 \\ 1 & 2 & 1 \end{pmatrix}.$$

Damit können wir jetzt beispielsweise A^{100} oder A^{1000} mit erträglichem Aufwand ausrechnen, denn es gilt nach unserer Überlegung in Bemerkung 1.35 für alle $k \in \mathbb{N}$

$$\begin{aligned} A^k &= SD^k S^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 2 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0^k & 0 & 0 \\ 0 & 1^k & 0 \\ 0 & 0 & 3^k \end{pmatrix} \cdot \frac{1}{6} \begin{pmatrix} 2 & -2 & 2 \\ -3 & 0 & 3 \\ 1 & 2 & 1 \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 0 & -1 & 3^k \\ 0 & 0 & 2 \cdot 3^k \\ 0 & 1 & 3^k \end{pmatrix} \cdot \begin{pmatrix} 2 & -2 & 2 \\ -3 & 0 & 3 \\ 1 & 2 & 1 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 3^k + 3 & 2 \cdot 3^k & 3^k - 3 \\ 2 \cdot 3^k & 4 \cdot 3^k & 2 \cdot 3^k \\ 3^k - 3 & 2 \cdot 3^k & 3^k + 3 \end{pmatrix}. \end{aligned}$$

Bemerkung 1.39. Nicht jede Matrix $A \in M_{n,n}(\mathbb{R})$ ist diagonalisierbar. Zum Beispiel ist für

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

das charakteristische Polynom

$$p(\lambda) = \det(A - \lambda I_2) = \det \begin{pmatrix} 1 - \lambda & 0 \\ 1 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2,$$

so dass $\lambda_1 = 1$ eine doppelte Nullstelle ist. In der Lösung des linearen Gleichungssystems $(A - I_2)x = 0$ ergibt sich aber nach

$$\left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

nur ein freier Parameter, so dass der Eigenraum E_1 nur von einem Eigenvektor aufgespannt wird. Das Diagonalisierbarkeitskriterium in Methode 1.37 (b) ist damit nicht erfüllt. Man hat schlicht nicht genug Eigenvektoren, um die Matrix S aufzustellen.

Das folgende Kriterium für Diagonalisierbarkeit ist sehr praktisch und in keiner Weise offensichtlich.

1 Matrixrechnung und Eigenwerttheorie

Satz 1.40 (Symmetrische Matrizen sind diagonalisierbar). *Ist $A \in M_{n,n}(\mathbb{R})$ so, dass $A^T = A$ gilt, dann ist A immer diagonalisierbar.*

Matrizen, für die $A^T = A$ gilt, heißen *symmetrisch*

Beispiel 1.41. Als Anwendung der Diagonalisierungs-Technik bestimmen wir zum Schluss die 2025-te Fibonacci-Zahl. Die Fibonacci-Zahlen beginnen mit 1, 1, 2, 3, 5, ... und die nächste Zahl dieser Folge ergibt sich immer als die Summe der beiden vorherigen. Es geht also los mit

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \dots$$

und wir wollen jetzt herausbekommen, welche Zahl an der 2025. Stelle in dieser Liste steht, natürlich ohne alle 2024 anderen Zahlen davor ausrechnen zu müssen.

Wir betrachten also die Folge (F_k) mit $F_1 = 1, F_2 = 1$, die der Rekursion $F_{k+1} = F_k + F_{k-1}$ für alle $k \geq 2$ genügt. Ferner setzen $F_0 := 0$ und für $k \geq 1$

$$v_k := \begin{pmatrix} F_k \\ F_{k-1} \end{pmatrix} \in \mathbb{R}^2.$$

Dann gilt für alle $k \geq 1$

$$v_{k+1} = \begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix} = \begin{pmatrix} F_k + F_{k-1} \\ F_k \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_k \\ F_{k-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} v_k$$

und damit

$$v_{k+1} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} v_k = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^2 v_{k-1} = \dots = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^k v_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^k \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Um F_{2025} als zweite Komponente von v_{2026} zu berechnen, brauchen wir also A^{2025} für die Matrix $A := \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Dazu diagonalisieren wir A nach dem Verfahren in Methode 1.37.

(a) Das charakteristische Polynom von A berechnet sich zu

$$p(\lambda) = \det(A - \lambda I_2) = \det \begin{pmatrix} 1 - \lambda & 1 \\ 1 & -\lambda \end{pmatrix} = (1 - \lambda)(-\lambda) - 1 = \lambda^2 - \lambda - 1$$

und die Nullstellen, d. h. die Eigenwerte von A sind nach der p - q -Formel

$$\lambda_{1,2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 1} = \frac{1}{2} \pm \sqrt{\frac{5}{4}} = \frac{1 \pm \sqrt{5}}{2}, \quad \text{also } \lambda_1 = \frac{1 + \sqrt{5}}{2} \text{ und } \lambda_2 = \frac{1 - \sqrt{5}}{2}.$$

Die Eigenräume findet man wieder durch Lösen der LGSe $(A - \lambda_{1,2} I_2)x = 0$. Wir führen das erste vor:

$$\left(\begin{array}{cc|c} 1 - \frac{1+\sqrt{5}}{2} & 1 & 0 \\ 1 & -\frac{1+\sqrt{5}}{2} & 0 \end{array} \right) = \left(\begin{array}{cc|c} \frac{1-\sqrt{5}}{2} & 1 & 0 \\ 1 & -\frac{1+\sqrt{5}}{2} & 0 \end{array} \right) \xrightarrow{Z2+1+\sqrt{5}/2Z1} \left(\begin{array}{cc|c} \frac{1-\sqrt{5}}{2} & 1 & 0 \\ 0 & 0 & 0 \end{array} \right).$$

1.2 Eigenwerte und Eigenvektoren

Mit $x_2 = t$ folgt $x_1 = -\frac{2}{1-\sqrt{5}}x_2 = \frac{2}{\sqrt{5}-1}t$ und damit

$$E_{\lambda_1} = \left\{ \begin{pmatrix} \frac{2}{\sqrt{5}-1}t \\ t \end{pmatrix} : t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 2s \\ (\sqrt{5}-1)s \end{pmatrix} : s \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} 2 \\ \sqrt{5}-1 \end{pmatrix} \right\}.$$

Genauso findet man

$$E_{\lambda_2} = \text{span} \left\{ \begin{pmatrix} -2 \\ \sqrt{5}+1 \end{pmatrix} \right\}.$$

(b) Wir haben zwei verschiedene Eigenwerte, also ist A diagonalisierbar.

(c) Mit

$$S = \begin{pmatrix} 2 & -2 \\ \sqrt{5}-1 & \sqrt{5}+1 \end{pmatrix} \quad \text{und} \quad D = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}$$

gilt nun $A = SDS^{-1}$ und $A^{2025} = SD^{2025}S^{-1}$. Wir benötigen als nächstes also S^{-1} . Es gilt

$$\begin{aligned} & \left(\begin{array}{cc|cc} 2 & -2 & 1 & 0 \\ \sqrt{5}-1 & \sqrt{5}+1 & 0 & 1 \end{array} \right) \xrightarrow{Z2-\sqrt{5}-1/2Z1} \left(\begin{array}{cc|cc} 2 & -2 & 1 & 0 \\ 0 & 2\sqrt{5} & -\frac{\sqrt{5}-1}{2} & 1 \end{array} \right) \\ & \xrightarrow{\substack{Z1/2 \\ Z2/2\sqrt{5}}} \left(\begin{array}{cc|cc} 1 & -1 & \frac{1}{2} & 0 \\ 0 & 1 & -\frac{\sqrt{5}-1}{4\sqrt{5}} & \frac{1}{2\sqrt{5}} \end{array} \right) \xrightarrow{Z1+Z2} \left(\begin{array}{cc|cc} 1 & 0 & \frac{\sqrt{5}+1}{4\sqrt{5}} & \frac{2}{4\sqrt{5}} \\ 0 & 1 & -\frac{\sqrt{5}-1}{4\sqrt{5}} & \frac{2}{4\sqrt{5}} \end{array} \right) \end{aligned}$$

und somit

$$\begin{pmatrix} 2 & -2 \\ \sqrt{5}-1 & \sqrt{5}+1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{\sqrt{5}+1}{4\sqrt{5}} & \frac{2}{4\sqrt{5}} \\ -\frac{\sqrt{5}-1}{4\sqrt{5}} & \frac{2}{4\sqrt{5}} \end{pmatrix}.$$

Es folgt

$$\begin{aligned} A &= \begin{pmatrix} 2 & -2 \\ \sqrt{5}-1 & \sqrt{5}+1 \end{pmatrix} \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{5}+1}{4\sqrt{5}} & \frac{2}{4\sqrt{5}} \\ -\frac{\sqrt{5}-1}{4\sqrt{5}} & \frac{2}{4\sqrt{5}} \end{pmatrix}, \\ A^k &= \begin{pmatrix} 2 & -2 \\ \sqrt{5}-1 & \sqrt{5}+1 \end{pmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^k \end{pmatrix} \begin{pmatrix} \frac{\sqrt{5}+1}{4\sqrt{5}} & \frac{2}{4\sqrt{5}} \\ -\frac{\sqrt{5}-1}{4\sqrt{5}} & \frac{2}{4\sqrt{5}} \end{pmatrix}. \end{aligned}$$

1 Matrixrechnung und Eigenwerttheorie

Wir können nun $v_{k+1} = A^k v_1$ berechnen:

$$\begin{aligned} A^k v_1 &= A^k \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ \sqrt{5}-1 & \sqrt{5}+1 \end{pmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^k \end{pmatrix} \begin{pmatrix} \frac{\sqrt{5}+1}{4\sqrt{5}} \\ -\frac{\sqrt{5}-1}{4\sqrt{5}} \end{pmatrix} \\ &= \begin{pmatrix} 2 & -2 \\ \sqrt{5}-1 & \sqrt{5}+1 \end{pmatrix} \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k \frac{\sqrt{5}+1}{4\sqrt{5}} \\ \left(\frac{1-\sqrt{5}}{2}\right)^k \frac{1-\sqrt{5}}{4\sqrt{5}} \end{pmatrix} \\ &= \begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^k \frac{\sqrt{5}+1}{2\sqrt{5}} - \left(\frac{1-\sqrt{5}}{2}\right)^k \frac{1-\sqrt{5}}{2\sqrt{5}} \\ \left(\frac{1+\sqrt{5}}{2}\right)^k \frac{1}{\sqrt{5}} - \left(\frac{1-\sqrt{5}}{2}\right)^k \frac{1}{\sqrt{5}} \end{pmatrix}. \end{aligned}$$

Schließlich folgt:

$$\begin{pmatrix} F_{2026} \\ F_{2025} \end{pmatrix} = v_{2026} = A^{2025} v_1 = \begin{pmatrix} \dots \\ \left(\frac{1+\sqrt{5}}{2}\right)^{2025} \frac{1}{\sqrt{5}} - \left(\frac{1-\sqrt{5}}{2}\right)^{2025} \frac{1}{\sqrt{5}} \end{pmatrix}$$

und somit

$$\begin{aligned} F_{2025} &= \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2}\right)^{2025} - \left(\frac{1-\sqrt{5}}{2}\right)^{2025} \right) \\ &= 7087392816114092863104543288308926461944946038998920616 \\ &\quad 72242360826040996582038473104532100347525076635712512 \\ &\quad 787887657147133880917647897624697727571194993510261758 \\ &\quad 622303698434275625929244269110857095599152999421410208 \\ &\quad 683245194925859326192653926929554791755473493205020144 \\ &\quad 894397011380434047669094673687672145460415348890314231 \\ &\quad 571156071926528489328749070735630792667485835932036147 \\ &\quad 913012910501318548192446375610424016293967650 \\ &\approx 7,0874 \cdot 10^{422}. \end{aligned}$$

2 Mehrdimensionale Differentialrechnung

2.1 Einleitung

In diesem Kapitel beschäftigen wir uns mit den Eigenschaften von Abbildungen

$$f : \mathbb{R}^m \rightarrow \mathbb{R}^n.$$

Insbesondere geht es uns darum, die in der Höheren Mathematik I betrachteten Methoden zur Bestimmung von Extremwerten auf höhere Dimensionen zu verallgemeinern.

- Beispiel 2.1.** (a) Die Abbildung $T : \mathbb{R}^3 \rightarrow \mathbb{R}$, die einem Punkt $(x, y, z)^T \in \mathbb{R}^3$ (in gegebenem Koordinatensystem) die Temperatur zuordnet. Typische Frage: wo ist die Temperatur minimal/maximal? Wie lässt sich ein Extremum bestimmen?
- (b) Die Abbildung $W : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die jedem Punkt die Windgeschwindigkeit oder eine elektrische Feldstärke zuordnet. Typische Frage: Wo sind Quellen des Felds?
- (c) Als abstrakteres Beispiel betrachten wir $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = x^2 + y^2$. Diese Funktion beschreibt das Quadrat des Abstands eines Punktes (x, y) zum Ursprung $(0, 0)$. Ein weiteres Beispiel: $g(x, y) = \sin(x) \cos(y)$. Die Graphen solcher Funktionen lassen sich als Flächen im \mathbb{R}^3 visualisieren, vgl. Abbildung 2.1.
- (d) Für jede Matrix $A \in M_{m,n}(\mathbb{R})$ definiert $f(v) = Av$ mit $v \in \mathbb{R}^n$ eine Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$.
- (e) Die Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben durch $f(x, y) = (-x, -y)^T$ und $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben durch $g(x, y) = (y, -x)^T$ sind Beispiele für sog. *Vektorfelder*. Zwei- und dreidimensionale Vektorfelder kann man durch Pfeile in einem Koordinatensystem darstellen, vgl. Abbildung 2.2.
- (f) Der Druck eines Gases ist näherungsweise gegeben durch die *van-der-Waals-Gleichung*

$$p(T, V) = \frac{RT}{v/n - b} - \frac{an^2}{V^2}$$

2 Mehrdimensionale Differentialrechnung

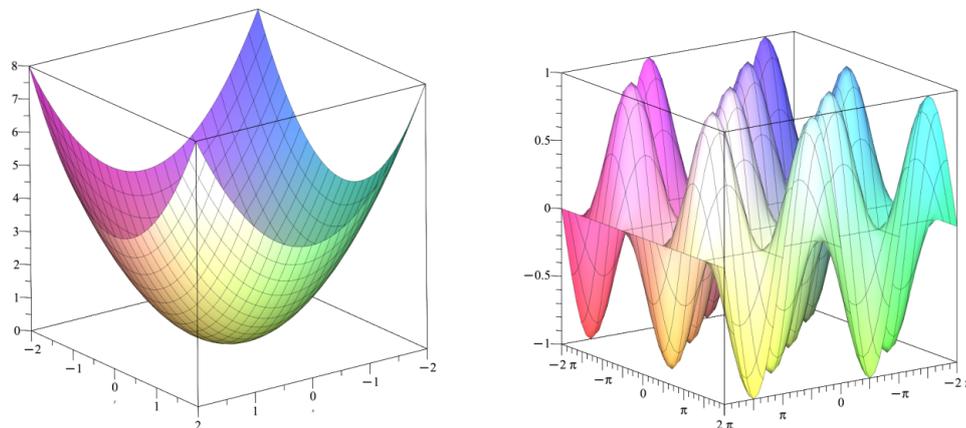


Abbildung 2.1: Die Graphen der Funktionen $f(x, y) = x^2 + y^2$ und $g(x, y) = \sin(x) \cos(y)$

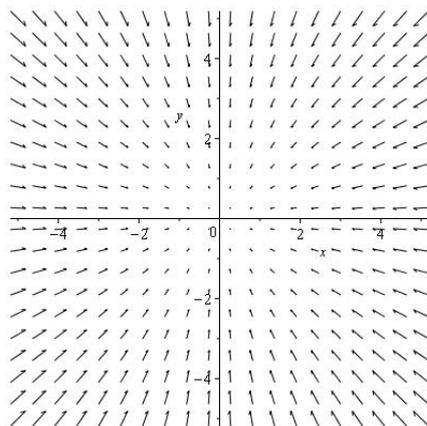


Abbildung 2.2: Visualisierung des Vektorfeldes $f(x, y) = (-x, -y)$

wobei

T = Temperatur in Kelvin,

$R \approx 8,314 \frac{\text{J}}{\text{mol K}}$, universelle Gaskonstante,

V = Volumen in m^3 ,

n = Stoffmenge, d.h. Anzahl Mol, 1 Mol entspricht $6,022 \cdot 10^{23}$ Gasteilchen,

a, b = Konstanten die vom Gas abhängen.

Definition 2.2. Sei $D \subseteq \mathbb{R}^m$. Eine Abbildung $f : D \rightarrow \mathbb{R}$ heißt Skalarfeld. Abbildungen $f : D \rightarrow \mathbb{R}^m$ (gleiche Dimension des Zielraums wie die Definitionsmenge!) heißen Vektorfelder.

Jede Funktion $f : D \rightarrow \mathbb{R}^n$ lässt sich als Vektor von Funktionen mit Werten in \mathbb{R} als

$$f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix}$$

schreiben. Die Funktion $f_j : D \rightarrow \mathbb{R}$, die die j -te Zeile von f angibt, heißt dann j -te Koordinatenfunktion von f .

2.2 Folgen und Stetigkeit

In diesem Abschnitt wiederholen und verallgemeinern wir wichtige Grundlagen der reellen Analysis aus der Höheren Mathematik I.

Erinnerung 2.3. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist stetig in $x_0 \in \mathbb{R}$, falls

$$\lim_{j \rightarrow \infty} f(x_j) = f(x_0)$$

gilt für alle Folgen $(x_j)_{j \in \mathbb{N}} = (x_1, x_2, x_3, \dots)$, die gegen x_0 konvergieren oder kürzer mit dem Funktionsgrenzwert ausgedrückt, wenn $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ ist.

Wie lässt sich die Stetigkeit für $f : \mathbb{R}^m \rightarrow \mathbb{R}$ definieren? Wir gehen ähnlich vor, benötigen aber Folgen von Vektoren, da die Argumente von f Vektoren sind.

Definition 2.4. Eine Folge in \mathbb{R}^m ist eine Abbildung von \mathbb{N} nach \mathbb{R}^m . Konkret schreiben wir $(v^{(j)})_{j \in \mathbb{N}}$ für eine Folge, wobei $v^{(j)} = (v_1^{(j)}, \dots, v_m^{(j)})^T \in \mathbb{R}^m$ für jedes $j \in \mathbb{N}$ Vektoren sind. Für $k = 1, \dots, m$ heißt die reellwertige Folge $(v_k^{(j)})_{j \in \mathbb{N}}$ die k -te Koordinatenfolge von $(v^{(j)})_{j \in \mathbb{N}}$.

Bemerkung 2.5. Beachten Sie die Änderung der Notation! Der Folgengliedindex steht nun *oben*, um unten Platz für den Koordinatenindex zu machen.

Beispiel 2.6. (a) Für $m = 2$ definiert

$$v^{(j)} = \begin{pmatrix} 1/j \\ 1 - 1/j \end{pmatrix}, \quad j \in \mathbb{N},$$

eine Folge in \mathbb{R}^2 . Beide Koordinatenfolgen $(v_1^{(j)})_{j \in \mathbb{N}} = (1/j)_{j \in \mathbb{N}}$ und $(v_2^{(j)})_{j \in \mathbb{N}} = (1 - 1/j)_{j \in \mathbb{N}}$ konvergieren in \mathbb{R} und ihre Grenzwerte sind 0 bzw. 1.

(b) Jetzt betrachten wir

$$v^{(j)} = \begin{pmatrix} (-1)^j \\ 1/j^2 \end{pmatrix}, \quad j \in \mathbb{N}.$$

2 Mehrdimensionale Differentialrechnung

Hier konvergiert die zweite Koordinatenfolge $(v_2^{(j)})_{j \in \mathbb{N}} = (1/j^2)_{j \in \mathbb{N}}$ gegen Null. Die erste Koordinatenfolge $(v_1^{(j)})_{j \in \mathbb{N}} = ((-1)^j)_{j \in \mathbb{N}}$ konvergiert nicht, sondern divergiert. Aber beide Koordinatenfolgen sind beschränkt.

(c) Schließlich betrachten wir die Folge

$$v^{(j)} = \begin{pmatrix} j \\ j^2 \end{pmatrix}, \quad j \in \mathbb{N}.$$

Beide Koordinatenfolgen $(v_1^{(j)})_{j \in \mathbb{N}} = (j)_{j \in \mathbb{N}}$ und $(v_2^{(j)})_{j \in \mathbb{N}} = (j^2)_{j \in \mathbb{N}}$ divergieren und sind unbeschränkt.

Definition 2.7. Sei $(v^{(j)})_{j \in \mathbb{N}}$ eine Folge im \mathbb{R}^m .

- (a) Die Folge $(v^{(j)})_{j \in \mathbb{N}}$ heißt beschränkt, falls alle m Koordinatenfolgen $(v_k^{(j)})_{j \in \mathbb{N}}$ für $k = 1, \dots, m$ beschränkt sind.
- (b) Die Folge $(v^{(j)})_{j \in \mathbb{N}}$ heißt konvergent, falls alle m Koordinatenfolgen $(v_k^{(j)})_{j \in \mathbb{N}}$ für $k = 1, \dots, m$ gegen reelle Zahlen v_1, \dots, v_m konvergieren. Dann nennen wir den Vektor $v = (v_1, \dots, v_m)^T \in \mathbb{R}^m$ Grenzwert der Folge $(v^{(j)})_{j \in \mathbb{N}}$ und schreiben $v = \lim_{j \rightarrow \infty} v^{(j)}$.

Schauen wir noch einmal die Folgen in Beispiel 2.6 an, so stellen wir fest, dass die Folge in (a) beschränkt und konvergent, die in (b) beschränkt aber nicht konvergent und die in (c) weder beschränkt noch konvergent ist. Die vierte theoretisch mögliche Situation, nicht beschränkt und konvergent, kann es nicht geben, weil genauso wie bei reellen Folgen jede konvergente Folge in \mathbb{R}^m automatisch beschränkt sein muss.

Wir können nun genauso wie in \mathbb{R} Stetigkeit mittels konvergenter Folgen definieren.

Definition 2.8. Sei $D \subseteq \mathbb{R}^m$.

- (a) Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt stetig in $v \in D$, falls für jede konvergente Folge $(v^{(j)})_{j \in \mathbb{N}}$ mit $v^{(j)} \in D$ für alle $j \in \mathbb{N}$ und $\lim_{j \rightarrow \infty} v^{(j)} = v$ gilt

$$\lim_{j \rightarrow \infty} f(v^{(j)}) = f(v).$$

Man nennt f stetig auf D oder auch schlicht stetig, wenn f in jedem $v \in D$ stetig ist.

- (b) Eine Funktion $f : D \rightarrow \mathbb{R}^n$ heißt stetig in v bzw. auf D , falls jede Koordinatenfunktion f_1, \dots, f_n stetig in v bzw. auf D ist.

Beim Lesen der Stetigkeitsdefinition sollte eine besondere Betonung auf dem Wort „jede“ in „jede Folge“ liegen. Es reicht hier nicht einzelne Beispielfolgen zu betrachten, sondern die Bedingung muss wirklich für jede denkbare Folge gelten. Umgekehrt lässt sich Unstetigkeit durch die Angabe einer einzigen ganz konkreten Spielverderberfolge

nachweisen. Denn wenn die Bedingung nur für eine einzige Folge verletzt ist, ist es mit Stetigkeit schon vorbei.

Beispiel 2.9. Wir betrachten das Skalarfeld $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, welches durch

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = \begin{cases} \frac{1}{x^2+y^2}, & \text{für } (x, y) \neq (0, 0), \\ 0, & \text{für } (x, y) = (0, 0), \end{cases}$$

definiert wird und weisen zum Einen nach, dass diesen in jedem Punkt $v = (x, y)^T \in \mathbb{R}^2$ mit $v \neq 0$ stetig ist und in $v = 0$ unstetig ist. Sei zuerst $v = (x, y)^T \in \mathbb{R}^2$ mit $(x, y)^T \neq 0$. Für den Nachweis, dass f in v stetig ist, sei $(v^{(j)})_{j \in \mathbb{N}} = ((x_j, y_j)^T)_{j \in \mathbb{N}}$ eine Folge, die gegen $v = (x, y)^T$ konvergiert. Es gilt

$$\lim_{j \rightarrow \infty} f(x_j, y_j) = \lim_{j \rightarrow \infty} \frac{1}{x_j^2 + y_j^2} = \frac{1}{\lim_{j \rightarrow \infty} (x_j^2 + y_j^2)} = \frac{1}{x^2 + y^2} = f(x, y).$$

Hieraus folgt die Stetigkeit in allen Punkten $(x, y) \neq (0, 0)$. Dabei haben wir nur Rechenregeln für konvergente reelle Folgen benutzt.

Schließlich zeigen wir, dass f im Ursprung *nicht* stetig ist. Dazu wählen wir die Folge $(v^{(j)})_{j \in \mathbb{N}} = ((1/j, 0)^T)_{j \in \mathbb{N}}$ in \mathbb{R}^2 , die gegen $(0, 0)^T$ konvergiert. Für jedes $j \in \mathbb{N}$ gilt

$$f(v^{(j)}) = \frac{1}{\frac{1}{j^2} + 0^2} = j^2$$

und diese Folge divergiert. Also geht $(f(v^{(j)}))$ insbesondere nicht gegen $f(0, 0) = 0$ und damit ist f nicht stetig in $(0, 0)^T$.

Wie im eindimensionalen können wir aus stetigen Funktionen durch Addition, Subtraktion und Multiplikation neue stetige Funktionen generieren. Bei der Division muss man wie im eindimensionalen Fall vorsichtig sein und Division durch 0 ausschließen.

Satz 2.10. Seien $D \subseteq \mathbb{R}^m$ und $f, g : D \rightarrow \mathbb{R}$ Funktionen, die stetig in $v \in D$ sind. Dann gelten:

- (a) Die Funktionen $f + g$, $f - g$ und $f g$ sind stetig in v .
- (b) Gilt $g(v) \neq 0$, so ist auch f/g stetig in v .
- (c) Sei $h : \mathbb{R} \rightarrow \mathbb{R}$ stetig in $f(v)$. Dann ist auch $h \circ f : D \rightarrow \mathbb{R}$ stetig in v .

Mit der Hilfe dieser Regeln können wir Stetigkeit leichter überprüfen, denn wir können nun kompliziertere Funktionen aus einfacheren Bausteinen zusammensetzen.

Beispiel 2.11. (a) Der Ausdruck in der van-der-Waals-Gleichung aus Beispiel 2.1 (f) für den Druck p eines Gases in Funktion von Temperatur T und Volumen V

$$p(T, V) = \frac{RT}{v/n - b} - \frac{an^2}{V^2} \tag{2.1}$$

2 Mehrdimensionale Differentialrechnung

ist stetig auf der Menge

$$D := (0, \infty) \times (bn, \infty) = \{(T, V)^T \in \mathbb{R}^2 : T > 0 \text{ und } V > bn\}.$$

Dabei schließen wir den Wert $V = bn$ aus, um sicherzustellen, dass in (2.1) nicht durch Null dividiert wird.

(b) Das Vektorfeld $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, welches durch

$$f(x, y) = \begin{pmatrix} \sin(x) \cos(y) \\ xy^2 \end{pmatrix}$$

gegeben ist, ist auf \mathbb{R}^2 stetig. Um dies nachzuweisen, müssen wir die Stetigkeit beider Koordinatenfunktionen von f einsehen. Diese sind $(x, y) \mapsto \sin(x) \cos(y)$ und $(x, y) \mapsto xy^2$. Da Polynome und die trigonometrischen Funktionen Sinus und Cosinus stetig sind, sind beide Koordinatenfunktionen wegen der Rechenregeln aus Satz 2.10 stetig.

(c) Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & \text{falls } (x, y) \neq (0, 0), \\ 0 & \text{falls } (x, y) = (0, 0), \end{cases}$$

ist Quotient zweier Funktionen $f_n(x, y) = xy$ und $f_z(x, y) = x^2 + y^2$. Beide Funktionen sind Polynome und daher stetig. Weiterhin gilt $f_z(x, y) \neq 0$, für alle $(x, y) \neq (0, 0)$. Also ist $f = f_n/f_z$ in jedem Punkt von $\mathbb{R}^2 \setminus \{0\}$ stetig.

Wie in Beispiel 2.9 gibt es höchstens für $v = (0, 0)^T$ Schwierigkeiten. Wir betrachten die Folge $(v^{(j)})_{j \in \mathbb{N}} = ((1/j, 1/j)^T)_{j \in \mathbb{N}}$. Diese konvergiert gegen $v = (0, 0)^T$ und es gilt für jedes $j \in \mathbb{N}$

$$f(v^{(j)}) = \frac{1/j \cdot 1/j}{(1/j)^2 + (1/j)^2} = \frac{1/j^2}{2 \cdot 1/j^2} = \frac{1}{2}.$$

Zwar konvergiert diese Bildfolge $(f(v^{(j)}))_{j \in \mathbb{N}}$ damit, aber nicht gegen den Funktionswert $0 = f(v) = f(0, 0)$. Also ist f nicht stetig in $(0, 0)^T$. Das sieht man auch am Graph der Funktion, der in Abbildung 2.3 links dargestellt ist.

(d) Die Funktion $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$g(x, y) = \begin{cases} \frac{x^2y}{x^2+y^2} & \text{falls } (x, y) \neq (0, 0), \\ 0 & \text{falls } (x, y) = (0, 0), \end{cases}$$

ähnelt auf den ersten Blick f aus Teil (c), aber sie verbirgt eine Überraschung. Wie oben begründet man, dass g in allen $v \in \mathbb{R}^2 \setminus \{0\}$ stetig ist.

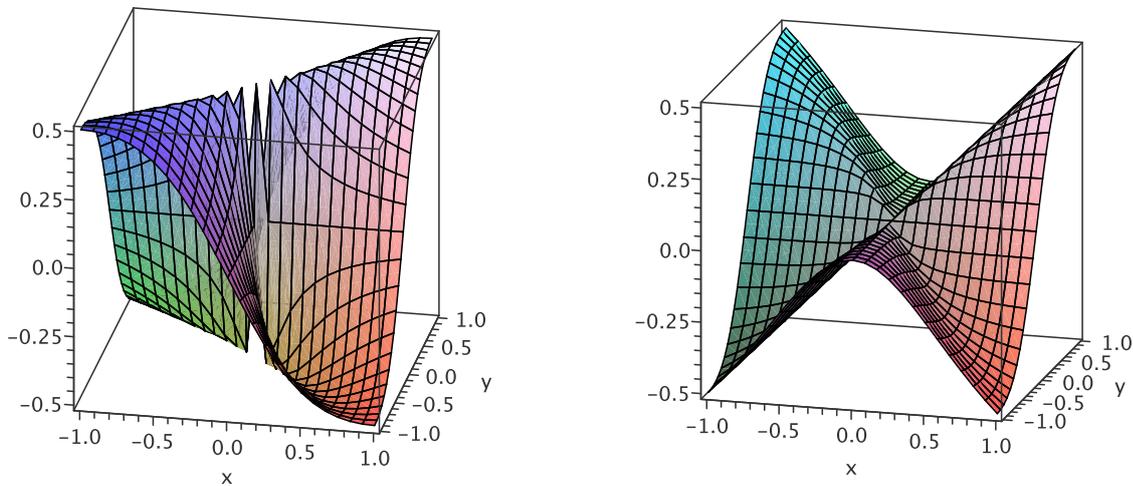


Abbildung 2.3: Graphen von $f(x, y) = \frac{xy}{x^2+y^2}$ links und $g(x, y) = \frac{x^2y}{x^2+y^2}$ rechts

Wir werden nun zeigen, dass g auch in $v = (0, 0)^T$ stetig ist. Dazu betrachten wir eine beliebige Folge $((x_1, y_1)^T, (x_2, y_2)^T, \dots)$ in \mathbb{R}^2 , die gegen $(0, 0)^T$ konvergiert. Dann gilt für alle $j \in \mathbb{N}$

$$0 \leq |g(x_j, y_j)| = \frac{|x_j^2 y_j|}{|x_j^2 + y_j^2|} = \frac{x_j^2 |y_j|}{x_j^2 + y_j^2} \leq \frac{x_j^2 |y_j|}{x_j^2} = |y_j|.$$

Da unsere Folge gegen den Ursprung $(0, 0)^T$ konvergiert, muss $\lim_{j \rightarrow \infty} y_j = 0$ gelten und mit dem „Sandwich-Theorem“ schließen wir $\lim_{j \rightarrow \infty} g(x_j, y_j) = 0 = g(0, 0)$. Also ist g stetig in $(0, 0)$. Vergleichen Sie hierzu auch die Visualisierung des Graphen in Abbildung 2.3.

Wir können nun ohne Änderungen die Definition des Funktionsgrenzwertes auf Funktionen in mehreren Variablen übertragen.

Definition 2.12. Sei $D \subseteq \mathbb{R}^m$, $f : D \rightarrow \mathbb{R}^n$ eine Funktion und $v_0 \in \mathbb{R}^m$. Wenn für jede Folge $((v^{(j)})_{j \in \mathbb{N}}$ mit

- $v^{(j)} \in D$ für alle $j \in \mathbb{N}$ und
- $v^{(j)} \neq v_0$ für alle $j \in \mathbb{N}$ und
- $(v^{(j)})_{j \in \mathbb{N}}$ konvergiert gegen v_0

der Grenzwert $\lim_{j \rightarrow \infty} f(v^{(j)})$ existiert, so bezeichnen wir diesen Grenzwert mit $\lim_{v \rightarrow v_0} f(v)$. (Der Grenzwert ist dann automatisch unabhängig von der Wahl der Folge $(v^{(j)})_{j \in \mathbb{N}}$.)

2 Mehrdimensionale Differentialrechnung

Bemerkung 2.13. (a) Wie beim Begriff der Stetigkeit in Definition 2.8, sollte auch beim Lesen dieser Definition eine besondere Betonung auf dem Lesen von „jede Folge“ liegen. Man beachte auch, dass obige Definition nur Sinn ergibt, wenn es auch Folgen im Definitionsbereich D gibt, die wie angegeben gegen v konvergieren. Wenn v mehrere Meilen vom Definitionsbereich wegliegt, wird man aber wohl auch keinen entsprechenden Grenzwert betrachten wollen.

(b) Mit diesem Funktionsgrenzwert können wir jetzt auch in der Definition von Stetigkeit, vgl. Definition 2.8, wieder kürzer formulieren: Eine Funktion $f : D \rightarrow \mathbb{R}^n$ ist stetig in $v_0 \in D$, wenn $\lim_{v \rightarrow v_0} f(v) = f(v_0)$ gilt.

In der reellen Analysis einer Dimension spielte das Suchen nach Maxima und Minima eine wichtige Rolle. Ähnliche Methoden existieren für Skalarfelder. Wir werden im nächsten Abschnitt die dafür wichtigen Begriff der partiellen Ableitung behandeln. Zuerst halten wir aber fest, dass Maxima und Minima in vielen Situationen sicher existieren. Dazu benötigen wir einige Begriffe, die besondere Arten von Teilmengen des Raumes \mathbb{R}^m auszeichnen.

Es sei an dieser Stelle an die Norm $\|v\| = \sqrt{v_1^2 + v_2^2 + \dots + v_m^2}$ erinnert, die für jeden Punkt $v \in \mathbb{R}^m$ seinen Abstand vom Ursprung angibt.

Definition 2.14. Sei $D \subseteq \mathbb{R}^m$.

(a) Wir nennen D beschränkt, falls ein $C \geq 0$ existiert mit

$$\|v\| \leq C \text{ für alle } v \in D.$$

(b) Wir nennen D abgeschlossen, falls für jede konvergente Folge $(v^{(j)})_{j \in \mathbb{N}}$, die in D liegt, auch der Grenzwert $\lim_{j \rightarrow \infty} v^{(j)}$ in D ist.

Beschränktheit ist genau das, was Sie sich anschaulich vorstellen: Die Menge hat nur eine begrenzte Ausdehnung. Abgeschlossenheit bedeutet anschaulich, dass jeder Punkt „am Rand“ auch zur Menge dazugehört. Wir definieren hier nicht was ein Rand ist und Sie dürfen sich bei den einfachen Mengen, die wir hier betrachten werden, auf Ihre Intuition verlassen. („Rand“ ist ein typisches Beispiel von etwas intuitiv sehr einfachem, das, je genauer man darüber nachdenkt, immer unklarer wird.)

Beispiel 2.15. (a) Das Rechteck mit Rand $[0, 1] \times [0, 2]$ ist abgeschlossen in \mathbb{R}^2 .

(b) Das Quadrat mit teilweise Rand $(0, 1) \times [0, 1]$ ist nicht abgeschlossen: Die Folge $((1/j, 1/j)^T)_{j \in \mathbb{N}}$ liegt in dieser Menge, konvergiert jedoch gegen den Ursprung $(0, 0)^T$ und der liegt nicht in $(0, 1) \times [0, 1]$.

(c) Beide Mengen $[0, 1] \times [0, 2]$ und $(0, 1) \times [0, 1]$ sind aber beschränkt, beispielsweise mit $C = 10$.

Die beiden Begriffe beschränkt und abgeschlossen, sind u. a. wegen des folgenden Resultats wichtig, dass auch unter der Bezeichnung *Satz von Weierstraß* firmiert.

Satz 2.16 (Weierstraß). Sei $D \subseteq \mathbb{R}^m$ eine beschränkte und abgeschlossene Menge und $f : D \rightarrow \mathbb{R}$ stetig. Dann nimmt f auf D sein Maximum und Minimum an, d. h. es existieren $v_{\max} \in D$ und $v_{\min} \in D$ mit

$$f(v_{\min}) \leq f(v) \leq f(v_{\max}) \quad \text{für alle } v \in D.$$

Dieser Satz ist ein oft unverzichtbares Hilfsmittel um die Existenz von Extremstellen abzusichern, auch wenn man sie nicht konkret ausrechnen kann.

2.3 Ableiten von reellwertigen Funktionen

Wir wollen nun den Begriff der Differenzierbarkeit und der Ableitung auf Funktionen von (Teilmengen von) \mathbb{R}^m nach \mathbb{R}^n erweitern und uns in einem ersten Schritt auf den Fall $n = 1$ konzentrieren. Dazu erinnern wir wieder zuerst an die eindimensionale Situation aus der Höheren Mathematik I.

Erinnerung 2.17. Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare Funktion, dann ist die Ableitung in $x_0 \in \mathbb{R}$ durch den Grenzwert des Differenzenquotienten

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \quad (2.2)$$

gegeben.

Für Funktionen mit m Variablen und $m \geq 2$ sind die Argumente x und x_0 Vektoren und dann ergibt leider der Differenzenquotient auf der rechten Seite keinen Sinn mehr, denn es gibt keine sinnvolle Möglichkeit, durch einen Vektor, hier $x - x_0$, zu dividieren. Dieses Problem werden wir später lösen.

Zunächst gilt es, ein eigentlich kleineres, eher technisches Problem zu lösen: Um überhaupt irgendeinen Grenzwert der (2.2) ersetzen soll, zu definieren, braucht es um den Basispunkt x_0 herum genügend Platz im Definitionsbereich von f . Schließlich muss man sich dem Punkt x_0 irgendwie innerhalb des Definitionsbereichs nähern können. Um diesen Platz sicherzustellen, führen wir die folgenden Konzepte ein.

Definition 2.18. (a) Sei $x_0 \in \mathbb{R}^m$ und $r > 0$, dann heißt

$$K(x_0, r) := \{x \in \mathbb{R}^m : \|x - x_0\| < r\}$$

Kugel vom Radius r um den Mittelpunkt x_0 .

(b) Eine Teilmenge $D \subseteq \mathbb{R}^m$ heißt offen, falls für jedes $x_0 \in D$ ein Radius $r > 0$ existiert, so dass $K(x_0, r) \subseteq D$ gilt.

2 Mehrdimensionale Differentialrechnung

Bemerkung 2.19. (a) Zum Verständnis der Definition der Kugel sei daran erinnert, dass der Term $\|a - b\|$ den Abstand von a und b angibt. Da steht also die Menge aller $x \in \mathbb{R}^m$ deren Abstand von x_0 kleiner als r ist. Im Dreidimensionalen ist das genau eine Kugel.

(b) Offenheit bedeutet, dass jeder Punkt der Menge noch eine ganze Kugel innerhalb der Menge um sich herum hat, die auch zur Menge gehört. Jeder Punkt der Menge liegt also kuschelig im Innern derselben. Damit ist die richtige Anschauung von Offenheit, dass kein Punkt des Randes zur Menge dazugehört.

(c) In Definition 2.14 haben wir abgeschlossene Teilmengen des \mathbb{R}^m als solche Mengen kennengelernt, bei denen der gesamte Rand zur Menge gehört. Damit sind offene und abgeschlossene Mengen in gewisser Weise zwei diametral entgegengesetzte Extremfälle, bei denen einmal nichts und einmal alles vom Rand dazugehört. Leider werden die beiden oft als ein entweder-oder-Pärchen im Kopf behalten, im Sinne von „Wenn die Menge nicht offen ist, dann muss sie ja abgeschlossen sein“. Das ist ungefähr so wie „Wenn sie nicht pleite ist, dann muss sie ja Milliardärin sein.“ Beides ist weit vorbei an der Realität. Beachten Sie bitte: Mengen sind keine Türen. Es gibt Mengen wie $[0, 1)$, die weder offen noch abgeschlossen sind und es gibt sogar solche, die gleichzeitig offen und abgeschlossen sind. Ein Beispiel hierfür ist \mathbb{R}^m selbst.

Beispiel 2.20. (a) Für $m = 1$, $x_0 \in \mathbb{R}$ und $r > 0$ ist $K(x_0, r) = (x_0 - r, x_0 + r)$. Eine Kugel im \mathbb{R}^1 ist also ein offenes Intervall. In Verallgemeinerung dessen ist jede Kugel $K(x_0, r)$ für beliebiges $x_0 \in \mathbb{R}^m$ und $r > 0$ offen. Der im \mathbb{R}^m neu eingeführte Begriff „offen“ verallgemeinert also das Konzept eines offenen Intervalls.

(b) Das Quadrat

$$(0, 1) \times (0, 1) = \{(x, y)^T \in \mathbb{R}^2 : 0 < x < 1 \text{ und } 0 < y < 1\}$$

ist ebenfalls offen.

(c) Allgemeiner seien $a = (a_1, \dots, a_m)^T$, $b = (b_1, \dots, b_m)^T \in \mathbb{R}^m$ mit $a_k < b_k$ für alle $k = 1, 2, \dots, m$. Dann ist der m -dimensionale Quader

$$\{(x_1, \dots, x_m)^T \in \mathbb{R}^m : a_k < x_k < b_k \text{ für alle } 1 \leq k \leq m\}$$

offen.

(d) Das Intervall $[0, 1) = \{x \in \mathbb{R} : 0 \leq x < 1\}$ ist *nicht* offen. Das Problem ist der linke Randpunkt, d. h. $0 \in [0, 1)$. Keine Kugel $K(0, r)$ mit $0 \in [0, 1)$ als Mittelpunkt liegt ganz in $[0, 1)$.

Nun kehren wir zurück zum eigentlichen Problem. Wie können wir Funktionen mit Vektoren als Argumenten ableiten? Dazu schauen wir zuerst ein Beispiel an.

2.3 Ableiten von reellwertigen Funktionen

Beispiel 2.21. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ die durch $f(x, y) = xe^y$ definierte Funktion.

- (a) Sei $y_0 \in \mathbb{R}$ beliebig. Wir betrachten $g(x) = f(x, y_0)$ als Funktion, indem wir die zweite Variable als festen Parameter verstehen. Es handelt sich dann um eine Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ und somit ist das Differenzieren für uns kein Problem. Wir leiten g an der Stelle x mit den üblichen Methoden ab und profitieren davon, dass e^{y_0} jetzt einfach eine Konstante ist:

$$g'(x) = e^{y_0} = \lim_{h \rightarrow 0} \frac{f(x+h, y_0) - f(x, y_0)}{h}.$$

- (b) Umgekehrt können wir auch die x -Koordinate festhalten. Wir fixieren also ein beliebiges $x_0 \in \mathbb{R}$ und betrachten $h(y) = f(x_0, y)$. Die Funktion h geht nun auch von \mathbb{R} nach \mathbb{R} und wir leiten wieder „wie gewohnt“ ab, indem wir uns daran erinnern, dass die Ableitung der Exponentialfunktion die Exponentialfunktion ist. Beachten Sie, dass wir nun nach dem Buchstaben y differenzieren müssen und x_0 eine Konstante ist. Wir bekommen

$$h'(y) = x_0 e^y = \lim_{h \rightarrow 0} \frac{f(x_0, y+h) - f(x_0, y)}{h}.$$

Fazit: für eine Funktion f in zwei Variablen gibt es mehrere Ableitungen. Wir nennen g' bzw. h' partielle Ableitungen von f nach x bzw. y .

Wir übertragen das allgemein auf Skalarfelder mit Variablen in beliebiger Dimension.

Definition 2.22. Sei $D \subseteq \mathbb{R}^m$ eine offene Menge, $f : D \rightarrow \mathbb{R}$ ein Skalarfeld und $x \in D$. Weiter verwenden wir für $k = 1, 2, \dots, m$ die Schreibweise

$$e_k = (0, \dots, 0, \underbrace{1}_{k\text{-te Komponente}}, 0, \dots, 0)^T$$

für den Nur-eine-Eins-an-der- k -ten-Stelle-Vektor.

- (a) Existiert der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x + he_k) - f(x)}{h},$$

so heißt f in x partiell nach x_k differenzierbar. Der Grenzwert heißt partielle Ableitung nach x_k von f in x und wird mit

$$\frac{\partial f}{\partial x_k}(x) \text{ oder } \partial_k f(x) \text{ oder } \partial_{x_k} f(x) \text{ oder } f_{x_k}(x)$$

bezeichnet.¹

¹An der Stelle von x_k werden wir stillschweigend andere übliche Koordinatenbezeichnungen benutzen. Ist z.B. $f(x, y) = xe^y$ so schreiben wir $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$.

2 Mehrdimensionale Differentialrechnung

- (b) Ist f in x nach allen Koordinaten x_1, x_2, \dots, x_m partiell differenzierbar, so heißt f partiell differenzierbar in x . Ist f weiter in jedem Punkt von D partiell differenzierbar, so nennen wir f partiell differenzierbar auf D oder schlicht partiell differenzierbar. Trifft dies zu und sind außerdem alle partiellen Ableitungsfunktionen

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_m}$$

stetige Funktionen auf D , so nennt man f stetig partiell differenzierbar.

Bemerkung 2.23. (a) Partiiell differenzieren nach x_k bedeutet, dass man alle anderen Variablen als feste Parameter betrachtet. In diesem Sinne kann man wie aus der Höheren Mathematik I gewohnt ableiten und wir müssen keine neuen Ableitungsregeln formulieren.

- (b) Anschaulich bezeichnet

$$\frac{\partial f}{\partial x_k}(x)$$

die Änderungsrate von f wenn die Variable x_k variiert wird und alle anderen Variablen festgehalten werden.

Beispiel 2.24. (a) Wir betrachten $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = x^3y - 2x^2y^2$. Es gilt

$$\frac{\partial f}{\partial x}(x, y) = 3x^2y - 4xy^2 \text{ und } \frac{\partial f}{\partial y}(x, y) = x^3 - 4x^2y.$$

- (b) Es gibt keinen Grund, sich auf den \mathbb{R}^2 zu beschränken. Sei also $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ mit $g(x, y, z) = xe^{xz+y^2}$ gegeben. Dann gilt

$$\begin{aligned} \frac{\partial g}{\partial x}(x, y, z) &= e^{xz+y^2} + xe^{xz+y^2} \cdot z = (1+xz)e^{xz+y^2}, \\ \frac{\partial g}{\partial y}(x, y, z) &= xe^{xz+y^2} \cdot 2y = 2xye^{xz+y^2}, \\ \frac{\partial g}{\partial z}(x, y, z) &= xe^{xz+y^2} \cdot x = x^2e^{xz+y^2}. \end{aligned}$$

- (c) Wir betrachten die Gleichung von van der Waals, vgl. Beispiel 2.1 (f),

$$p(T, V) = \frac{RT}{V/n - b} - \frac{an^2}{V^2}$$

auf dem Definitionsbereich $D = (0, \infty) \times (bn, \infty)$.

Die partielle Ableitung $\frac{\partial p}{\partial T}$ beschreibt die Änderung des Drucks, die durch Veränderung der Temperatur bei konstantem Volumen entsteht (isochore Zustandsänderung). Es ist

$$\frac{\partial p}{\partial T} = \frac{R}{V/n - b}.$$

2.3 Ableiten von reellwertigen Funktionen

Die partielle Ableitung $\frac{\partial p}{\partial V}$ beschreibt die Druckänderung, die durch Veränderung des Volumens bei konstanter Temperatur entsteht (isotherme Zustandsänderung). Es gilt

$$\frac{\partial p}{\partial V} = -\frac{RT}{(V/n - b)^2 n} + 2\frac{an^2}{V^3}.$$

Die reine Existenz der partiellen Ableitungen ermöglicht es leider nicht, daraus hilfreiche Schlussfolgerungen über die Funktion zu ziehen. Das wird das folgende Beispiel vor Augen führen.

Beispiel 2.25. Wir knüpfen an Beispiel 2.11 (c) an. Wie wir dort gesehen haben, ist die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & \text{falls } (x, y) \neq (0, 0), \\ 0, & \text{falls } (x, y) = (0, 0), \end{cases}$$

nicht stetig im Ursprung $(0, 0)^T$ aber stetig in allen anderen Punkten des \mathbb{R}^2 , vgl. Abbildung 2.3.

Wie sieht es mit den partiellen Ableitungen im Koordinatenursprung aus? Naiv betrachten sollte die Funktion f dort in keiner Weise differenzierbar sein, denn sie ist dort nicht einmal stetig und wir wissen aus der Höheren Mathematik I, dass Differenzierbarkeit ein stärkerer Begriff als Stetigkeit ist. Allerdings gilt

$$\lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0/h^2 - 0}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0.$$

Also existiert die partielle Ableitung $\frac{\partial f}{\partial x}$ bei $(0, 0)^T$ mit Wert 0. Da die Funktion in x und y symmetrisch ist, findet man genauso

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0$$

und auch die partielle Ableitung von f nach y existiert im Ursprung und hat den Wert 0.

Außerhalb des Ursprungs gilt nach der Quotientenregel

$$\frac{\partial f}{\partial x}(x, y) = \frac{y(x^2 + y^2) - xy \cdot 2x}{(x^2 + y^2)^2} = \frac{y^3 - x^2y}{(x^2 + y^2)^2}$$

für $(x, y) \neq (0, 0)$. Setzt man $x = 0$ ein, findet man wegen

$$\frac{\partial f}{\partial x}(0, y) = \frac{y^3 - 0 \cdot y}{(0 + y^2)^2} = \frac{y^3}{y^4} = \frac{1}{y},$$

dass die partielle Ableitungsfunktion $\frac{\partial f}{\partial x}$ im Ursprung nicht stetig ist. Also ist f nicht stetig partiell differenzierbar.

2 Mehrdimensionale Differentialrechnung

Was sagt uns das nun? Alleine die partielle Ableitbarkeit einer Funktion sagt nicht wirklich etwas aus. Insbesondere kann man in einem Fall wie hier betrachtet, die Ableitungen in keiner Weise als Steigungen o.ä. interpretieren. Glücklicherweise sind solche Pathologien nur möglich, wenn die partiellen Ableitungen wie hier im Beispiel nicht stetig sind. Wir werden uns im Weiteren daher auf die Behandlung stetig partiell differenzierbarer Funktionen beschränken.

Für stetig partiell differenzierbare Funktionen mit Werten in \mathbb{R} fassen wir zunächst mal alle partiellen Ableitungen in einem Vektor zusammen und geben diesem einen Namen.

Definition 2.26. Sei $D \subseteq \mathbb{R}^m$ offen und $f : D \rightarrow \mathbb{R}$ stetig partiell differenzierbar auf D . Für $x = (x_1, \dots, x_m)^T \in D$ heißt dann der Vektor

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_m}(x) \right)$$

Gradient von f in x .

Da wir nun über stetig partiell differenzierbare Funktionen reden, hat der Gradient durchaus eine anschauliche Bedeutung und sogar eine sehr wichtige, die, vielleicht nicht ganz unerwartet, wieder etwas mit Steigungen zu tun hat.

Bemerkung 2.27. Sei $f : D \rightarrow \mathbb{R}^m$ stetig partiell differenzierbar. Anschaulich ist der Gradient $\nabla f(x)$ von f in x dann ein Vektor in \mathbb{R}^m , der in die Richtung zeigt, in der die Werte von f am stärksten anwachsen. Seine Norm ist das Ausmaß dieser Steigung.

Da der Gradient immer in die Richtung des steilsten Anstiegs zeigt, steht er außerdem in jedem Punkt genau senkrecht auf der durch diesen Punkt gehenden Höhenlinie der Funktion.

Beispiel 2.28. Die durch

$$f(x, y) = \sqrt{1 - x^2 - y^2}$$

definierte Funktion $f : K(0, 1) \rightarrow \mathbb{R}$ beschreibt die Höhe einer Halbkugel, denn ihr Graph ist genau die obere Halbsphäre, vgl. Abbildung 2.4. Wir berechnen

$$\begin{aligned} \nabla f(x, y) &= \left(\frac{1}{2\sqrt{1-x^2-y^2}} \cdot (-2x), \frac{1}{2\sqrt{1-x^2-y^2}} \cdot (-2y) \right) \\ &= \frac{1}{\sqrt{1-x^2-y^2}} (-x, -y). \end{aligned}$$

Dieser Gradient ist direkt neben der Funktion ebenfalls in Abbildung 2.4 illustriert. Man sieht hier schön, dass er immer in die Richtung des steilsten Anstiegs zum Ursprung hin zeigt und seine Größe zum Ursprung hin zusammen mit der Steigung des Graphen abnimmt. Die Höhenlinie dieser Funktion sind konzentrische Kreise um den Ursprung. Auf diesen steht der Gradient jeweils senkrecht.

2.4 Ableiten von Vektor-wertigen Funktionen

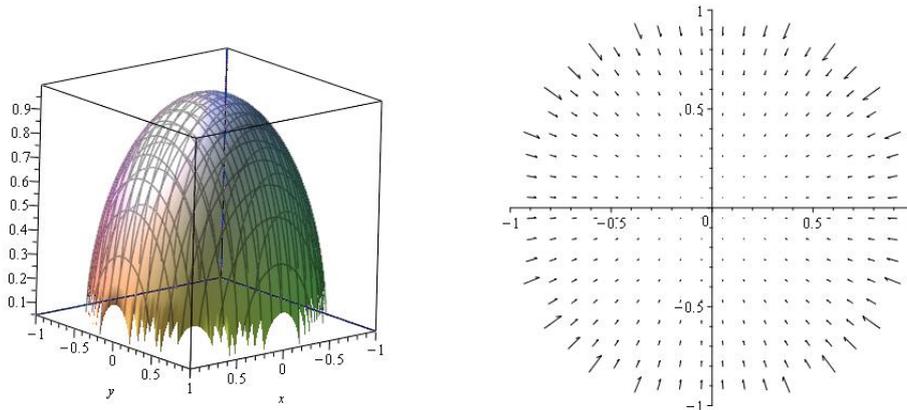


Abbildung 2.4: Der Graph der Funktion $f(x, y) = \sqrt{1 - x^2 - y^2}$ und eine Visualisierung ihres Gradienten

2.4 Ableiten von Vektor-wertigen Funktionen

Wir wenden uns nun dem allgemeinen Fall zu, dass die betrachtete Funktion von (einer Teilmenge des) \mathbb{R}^m in den \mathbb{R}^n abbildet. Wie schon zuvor spielen wir den Begriff der Ableitungen dann auf die einzelnen Koordinatenfunktionen f_1, f_2, \dots, f_n zurück und sind dann wieder im Fall von reellwertigen Funktionen, vgl. Definition 2.22, so dass wir große Teile der folgenden Definition schlicht von dort abschreiben können:

Definition 2.29. Sei $D \subseteq \mathbb{R}^m$ und $f : D \rightarrow \mathbb{R}^n$ eine Funktion mit Koordinatenfunktionen $f_1, f_2, \dots, f_n : D \rightarrow \mathbb{R}$.

- (a) Sei $x \in D$ und $k \in \{1, 2, \dots, m\}$. Dann ist f in x partiell nach x_k differenzierbar, wenn jede Koordinatenfunktion f_1, f_2, \dots, f_m dies ist. Wir schreiben dann mit der gleichen Vielfalt der Notationen wie schon in Definition 2.22

$$\frac{\partial f}{\partial x_k}(x) = \partial_k f(x) = \partial_{x_k} f(x) = f_{x_k}(x) = \begin{pmatrix} \partial_k f_1(x) \\ \partial_k f_2(x) \\ \vdots \\ \partial_k f_m(x) \end{pmatrix}$$

für den (Spalten-)Vektor der entsprechenden partiellen Ableitungen der Koordinatenfunktionen.

- (b) Ist f in x nach allen Koordinaten x_1, x_2, \dots, x_m partiell differenzierbar, so heißt f partiell differenzierbar in x . Ist f weiter in jedem $x \in D$ partiell differenzierbar, so nennen wir f partiell differenzierbar auf D oder schlicht partiell differenzierbar. Trifft dies zu und sind außerdem alle partiellen Ableitungsfunktionen

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_m}$$

2 Mehrdimensionale Differentialrechnung

stetige Funktionen auf D , so nennt man f stetig partiell differenzierbar auf D .

Auch für vektorwertige Funktionen beschränken wir uns in den weiteren Betrachtungen zur Differenzialrechnung darauf, ausschließlich stetig partiell differenzierbare Funktionen zu betrachten. Die Berechnung der partiellen Ableitungen einer solchen Funktion ist nicht schwierig, es können nur sehr viele werden. Deshalb ist es sehr sinnvoll, ein wenig Ordnung zu schaffen und alle Ableitungen zu einer Matrix zusammenzufassen.

Definition 2.30. Sei $D \subseteq \mathbb{R}^m$ offen und $f : D \rightarrow \mathbb{R}^n$ eine stetig partiell differenzierbare Funktion. Die Matrix

$$J_f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{pmatrix} = \begin{pmatrix} \text{---} & \nabla f_1(x) & \text{---} \\ \vdots & \vdots & \vdots \\ \text{---} & \nabla f_n(x) & \text{---} \end{pmatrix} \in M_{n,m}(\mathbb{R})$$

heißt Jacobi-Matrix von f in x .

Beispiel 2.31. Für das Vektorfeld $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ mit

$$f(x, y, z) = \begin{pmatrix} z \sin(xy) \\ x^2 y^3 z \\ x - y + z \end{pmatrix}$$

gilt $f_1(x, y, z) = z \sin(xy)$, $f_2(x, y, z) = x^2 y^3 z$ und $f_3(x, y, z) = x - y + z$ und damit ist

$$\begin{aligned} \frac{\partial f_1}{\partial x}(x, y, z) &= yz \cos(xy), & \frac{\partial f_1}{\partial y}(x, y, z) &= xz \cos(xy), & \frac{\partial f_1}{\partial z}(x, y, z) &= \sin(xy), \\ \frac{\partial f_2}{\partial x}(x, y, z) &= 2xy^3 z, & \frac{\partial f_2}{\partial y}(x, y, z) &= 3x^2 y^2 z, & \frac{\partial f_2}{\partial z}(x, y, z) &= x^2 y^3, \\ \frac{\partial f_3}{\partial x}(x, y, z) &= 1, & \frac{\partial f_3}{\partial y}(x, y, z) &= -1, & \frac{\partial f_3}{\partial z}(x, y, z) &= 1, \end{aligned}$$

bzw. als Vektoren geschrieben

$$\frac{\partial f}{\partial x} = \begin{pmatrix} yz \cos(xy) \\ 2xy^3 z \\ 1 \end{pmatrix}, \quad \frac{\partial f}{\partial y} = \begin{pmatrix} xz \cos(xy) \\ 3x^2 y^2 z \\ -1 \end{pmatrix}, \quad \frac{\partial f}{\partial z} = \begin{pmatrix} \sin(xy) \\ x^2 y^3 \\ 1 \end{pmatrix}$$

oder noch kompakter als Jacobi-Matrix

$$J_f(x) = \begin{pmatrix} yz \cos(xy) & xz \cos(xy) & \sin(xy) \\ 2xy^3 z & 3x^2 y^2 z & x^2 y^3 \\ 1 & -1 & 1 \end{pmatrix}.$$

Uns hindert nichts daran, höhere partielle Ableitungen zu betrachten.

2.4 Ableiten von Vektor-wertigen Funktionen

Definition 2.32. Sei $D \subseteq \mathbb{R}^m$ offen und $f : D \rightarrow \mathbb{R}^n$ stetig partiell differenzierbar auf D . Ist $\partial f / \partial x_k : D \rightarrow \mathbb{R}^n$ auf D erneut stetig partiell differenzierbar, so heißt

$$\frac{\partial^2 f}{\partial x_l \partial x_k} = \frac{\partial}{\partial x_l} \frac{\partial f}{\partial x_k}$$

partielle Ableitung 2. Ordnung von f nach x_l und x_k . Analog definiert man Ableitungen 3, 4, 5, ...ter Ordnung.

Beispiel 2.33. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert durch $f(x, y) = x^3 y + x e^y$. Dann gilt

$$\begin{array}{ll} \frac{\partial f}{\partial x}(x, y) = 3x^2 y + e^y, & \frac{\partial f}{\partial y}(x, y) = x^3 + x e^y, \\ \frac{\partial^2 f}{\partial x^2}(x, y) = 6xy, & \frac{\partial^2 f}{\partial x \partial y}(x, y) = 3x^2 + e^y, \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) = 3x^2 + e^y, & \frac{\partial^2 f}{\partial y^2}(x, y) = x e^y, \\ \frac{\partial^3 f}{\partial x^3}(x, y) = 6y, & \frac{\partial^3 f}{\partial x^2 \partial y}(x, y) = 6x, \\ \frac{\partial^3 f}{\partial y \partial x^2}(x, y) = 6x, & \frac{\partial^3 f}{\partial y \partial x \partial y}(x, y) = e^y, \\ \frac{\partial^3 f}{\partial x \partial y \partial x}(x, y) = 6x, & \frac{\partial^3 f}{\partial x \partial y^2}(x, y) = e^y, \\ \frac{\partial^3 f}{\partial y^2 \partial x}(x, y) = e^y, & \frac{\partial^3 f}{\partial y^3}(x, y) = x e^y. \end{array}$$

Betrachtet man die Ergebnisse in obigem Beispiel genauer, dann fällt auf, dass es nicht auf die Reihenfolge der Differentiation ankommt. So gilt beispielsweise

$$\frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial x^2 \partial y} = \frac{\partial^3 f}{\partial y \partial x^2} = 6x.$$

Dies ist kein Zufall. Tatsächlich ist die Differentiationsreihenfolge für stetig partiell differenzierbare Funktionen egal. Diese Aussage wird als *Satz von Schwarz* bezeichnet.

Satz 2.34 (Schwarz). Sei $D \subseteq \mathbb{R}^m$ eine offene Teilmenge und $f : D \rightarrow \mathbb{R}^n$ eine p -mal stetig partiell differenzierbare Funktion. Dann kommt es bei der Bildung der partiellen Ableitungen bis zur p -ten Ordnung nicht auf die Reihenfolge der Differentiation an.

Für Vektorfelder, also Funktionen von (Teilmengen von) \mathbb{R}^m in den Raum \mathbb{R}^m mit der gleichen Dimension, ist die Jacobi-Matrix eine quadratische Matrix. In diesem Fall gibt es weitere aus den partiellen Ableitungen, also den Einträgen der Jacobi-Matrix, abgeleitete Ableitungsgrößen, die eine Bedeutung haben und daher relevant sind. Wir wollen hier exemplarisch die sogenannte Divergenz behandeln.

2 Mehrdimensionale Differentialrechnung

Definition 2.35. Sei $D \subseteq \mathbb{R}^m$ offen und $f = (f_1, \dots, f_m)^T : D \rightarrow \mathbb{R}^m$ ein stetig partiell differenzierbares Vektorfeld. Dann heißt $\operatorname{div} f : D \rightarrow \mathbb{R}$ mit

$$\operatorname{div} f := \sum_{k=1}^m \frac{\partial f_k}{\partial x_k}$$

Divergenz von f .

Bemerkung 2.36. Bei der Berechnung der Divergenz werden jeweils die partiellen Ableitungen der Koordinatenfunktionen von f nach genau der Variablen, die den gleichen Index hat, addiert. Damit ist das Ergebnis der Rechnung eine Zahl, so dass die aus dem Vektorfeld f bestimmte Funktion $\operatorname{div} f$ ein Skalarfeld ist.

Für die Veranschaulichung der Divergenz von f stelle man sich das Vektorfeld f am besten als ein Magnetfeld oder ein Geschwindigkeitsfeld einer fließenden Flüssigkeit vor. In jedem Punkt x ist $f(x)$ also die Feldstärke bzw. der Geschwindigkeitsvektor der Strömung an dieser Stelle. Dann beschreibt die Divergenz $\operatorname{div} f$ die sogenannte *Quelldichte* des Feldes, d. h. ein positiver Wert von $\operatorname{div} f(x)$ deutet auf eine Quelle des Feldes, ein negativer auf eine Senke.

Beispiel 2.37. (a) Die Divergenz des Felds $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit $f(x, y) = (x, y)^T$ ist

$$\operatorname{div} f = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} = 1 + 1 = 2.$$

Anschaulich entsteht Feldstärke in jedem Punkt der Ebene.

(b) Wir betrachten $g := -f$ mit dem Vektorfeld f aus Teil (a), also $g(x, y) = (-x, -y)^T$, vgl. Abbildung 2.2. Dann zeigen die Feldpfeile zum Koordinatenursprung und es ist $\operatorname{div} g = -\operatorname{div} f = -2$ immer negativ. Dieses Feld hat in jedem Punkt eine Senke.

(c) Das Feld $h(x, y) = (y, -x)$ hat als Divergenz

$$\operatorname{div} h = \frac{\partial h_1}{\partial x} + \frac{\partial h_2}{\partial y} = 0 + 0 = 0.$$

Es ist sogenannt *divergenzfrei* und hat keine Quellen oder Senken.

2.5 Rechnen mit Ableitungen

Um Ableitungen wie die Jacobi-Matrix bzw. den Gradienten einer Funktion auszurechnen, berechnet man ihre verschiedenen partiellen Ableitungen, die am Ende immer Ableitungen einer skalaren Komponentenfunktion nach einer einzigen Variablen sind. Über die Techniken hinaus, die sie zur Berechnung solcher „eindimensionalen“ Ableitungen schon aus der Höheren Mathematik I kennen, brauchen wir also keine neuen

Methoden oder Rechenregeln. Trotzdem wollen wir die wichtigsten Ableitungsregeln noch mal im Matrixkontext aufschreiben, da insbesondere die Kettenregel gut aufzeigt, welche Bedeutung die Matrixmultiplikation für die mehrdimensionale Differentialrechnung hat.

Rechenregel 2.38. Sei $D \subseteq \mathbb{R}^m$ offen.

- (a) (**Linearität**) Seien $a, b \in \mathbb{R}$ und $f, g : D \rightarrow \mathbb{R}^n$ stetig partiell differenzierbar. Dann ist auch $af + bg : D \rightarrow \mathbb{R}^n$ stetig partiell differenzierbar und es gilt für jedes $x \in D$

$$J_{af+bg}(x) = aJ_f(x) + bJ_g(x).$$

Ist $n = 1$, d. h. die Funktionen f und g sind reellwertig, so schreibt man statt der Jacobi-Matrix üblicherweise den Gradienten. Mit diesem sieht die Formel dann so aus:

$$\nabla(af + bg)(x) = a\nabla f(x) + b\nabla g(x).$$

- (b) (**Produktregel**) Wir schreiben diese nur für den reellwertigen Fall hin. Sind $f, g : D \rightarrow \mathbb{R}$ stetig partiell differenzierbar, dann ist das Produkt $fg : D \rightarrow \mathbb{R}$ auch stetig partiell differenzierbar und es gilt für jedes $x \in D$

$$\nabla(fg)(x) = f(x)\nabla g(x) + g(x)\nabla f(x).$$

- (c) (**Kettenregel**) Seien $f : D \rightarrow \mathbb{R}^n$ stetig partiell differenzierbar, $E \subseteq \mathbb{R}^n$ offen mit $f(D) \subseteq E$ und $g : E \rightarrow \mathbb{R}^p$ stetig partiell differenzierbar. Dann ist die Verkettung $g \circ f : D \rightarrow \mathbb{R}^p$ ebenfalls stetig partiell differenzierbar und es gilt

$$J_{g \circ f}(x) = J_g(f(x)) \cdot J_f(x). \quad (2.3)$$

für alle $x \in D$.

Bemerkung 2.39. Die Gleichung (2.3) sieht vielleicht erst einmal ungewohnt aus, deshalb ist es wichtig zu sehen, dass sie genau die gleiche Form hat wie die Kettenregel in einer Dimension: Ableitung (=Jacobi-Matrix) der äußeren Funktion mit der inneren Funktion eingesetzt und dann mal die Ableitung der inneren Funktion. Eigentlich passiert hier also nichts neues.

Der Unterschied ist, dass die mehrdimensionale Kettenregel nun eine Gleichung von *Matrizen* ist. Links vom Gleichheitszeichen steht die $p \times m$ -Matrix $J_{g \circ f}(x)$. Die Matrix $J_g(f(x))$ ist eine $p \times n$ -Matrix und $J_f(x)$ ist vom Format $n \times m$. Das Matrixprodukt $J_g(f(x)) \cdot J_f(x)$ ist also ebenfalls eine $p \times m$ -Matrix und alles passt zusammen.

Beispiel 2.40. Wir betrachten die Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}^2$ und $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(t) = (t^2, t^3)$ und $g(x, y) = x^3y + xe^y$. Beide Funktionen sind stetig partiell differenzierbar und die Verkettung $g \circ f : \mathbb{R} \rightarrow \mathbb{R}$ ist

$$(g \circ f)(t) = g(f(t)) = g(t^2, t^3) = (t^2)^3 t^3 + t^2 e^{t^3} = t^9 + t^2 e^{t^3}.$$

2 Mehrdimensionale Differentialrechnung

Durch direktes Ausrechnen erhalten wir

$$\frac{d(g \circ f)}{dt} = (g \circ f)'(t) = 9t^8 + 2te^{t^3} + t^2 e^{t^3} \cdot 3t^2 = 9t^8 + 2te^{t^3} + 3t^4 e^{t^3}.$$

Diese Ableitung können wir auch mit der mehrdimensionalen Kettenregel bestimmen. Dazu halten wir

$$J_f(t) = \begin{pmatrix} 2t \\ 3t^2 \end{pmatrix} \quad \text{und} \quad J_g(x, y) = \nabla g(x, y) = (3x^2y + e^y, x^3 + xe^y)$$

fest. Um die Kettenregel anzuwenden setzen wir $f(t)$ in J_g ein und multiplizieren die entsprechende Matrix mit J_f . Das liefert

$$\begin{aligned} J_g(f(t)) \cdot J_f(t) &= (3(t^2)^2 t^3 + e^{t^3}, (t^2)^3 + t^2 e^{t^3}) \cdot \begin{pmatrix} 2t \\ 3t^2 \end{pmatrix} = (3t^7 + e^{t^3}, t^6 + t^2 e^{t^3}) \begin{pmatrix} 2t \\ 3t^2 \end{pmatrix} \\ &= 6t^8 + 2te^{t^3} + 3t^8 + 3t^4 e^{t^3} = 9t^8 + 2te^{t^3} + 3t^4 e^{t^3} \end{aligned}$$

genau wie oben.

Erinnern wir uns an die Höhere Mathematik I: Geometrisch beschreibt die Ableitung $f'(x_0)$ einer differenzierbaren Funktion die Steigung der Tangente an den Graphen von f im Punkt mit der x -Koordinate x_0 . In Formeln ausgedrückt gilt

$$f(x) = \underbrace{f(x_0) + f'(x_0)(x - x_0)}_{\text{lineare Näherung durch Tangente}} + \underbrace{r(x)}_{\text{Fehler der Näherung}},$$

wobei $r(x)$ den Näherungsfehler (also den Abstand zwischen dem Graphen von f und der approximierenden Tangente) bezeichnet. Für diesen Fehler gilt

$$\lim_{x \rightarrow x_0} \frac{|r(x)|}{|x - x_0|} = 0,$$

d. h. er geht schneller gegen Null als der Abstand von x zu x_0 schrumpft. Gleiches gilt auch in mehreren Dimensionen.

Satz 2.41 (Ableitung als beste lineare Näherung). *Es sei $D \subseteq \mathbb{R}^m$ offen und $f : D \rightarrow \mathbb{R}^n$ stetig partiell differenzierbar, sowie $x_0 \in D$. Dann ist $J_f(x_0)$ die einzige Matrix in $M_{n,m}(\mathbb{R})$, für die gilt*

$$f(x) = \underbrace{f(x_0) + J_f(x_0) \cdot (x - x_0)}_{\text{lineare Näherung}} + \underbrace{r(x)}_{\text{Fehler der Näherung}}$$

mit einem Fehler $r : D \rightarrow \mathbb{R}^n$, der

$$\lim_{x \rightarrow x_0} \frac{\|r(x)\|}{\|x - x_0\|}$$

erfüllt.

2.6 Bestimmen von Extrema in mehreren Variablen

Wieder besinnen wir uns auf ein schon bekanntes eindimensionales Resultat:

Erinnerung 2.42. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ zweimal differenzierbar. Dann gilt

- (a) Besitzt f in $x_0 \in \mathbb{R}$ ein lokales Extremum (d. h. ein lokales Minimum oder ein lokales Maximum), so gilt $f'(x_0) = 0$.
- (b) Gilt $f'(x_0) = 0$ und $f''(x_0) \neq 0$, so besitzt f ein lokales Extremum in x_0 . Für $f''(x_0) < 0$ ist x_0 ein lokales Maximum und für $f''(x_0) > 0$ ist x_0 ein lokales Minimum von f .

Lokale Extrema von Funktionen lassen sich also durch die Ableitungen charakterisieren. Wie übertragen wir dies nun auf mehrere Dimensionen?

Zunächst ist festzuhalten, dass die Frage nach besonders großen oder kleinen Werten einer Funktion nur Sinn ergibt, wenn die Werte der Funktion Zahlen in \mathbb{R} sind. Da man die Größe von Vektoren nicht unbedingt vergleichen kann (ist $(1, 0)^T$ oder $(0, 1)^T$ größer?), ist für eine Funktion $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ mit $n \geq 2$ schon die Frage nach einem Extremwert sinnfrei. Wir beschränken uns in diesem Abschnitt also auf Funktionen $f : D \rightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}^m$. Im gesamten Kapitel sei $D \subseteq \mathbb{R}^m$ eine offene Menge.

Definition 2.43. Sei $E \subseteq \mathbb{R}^m$ und $f : E \rightarrow \mathbb{R}$ eine Funktion. Dann hat f in $x_0 \in E$ ein

- (a) globales Minimum, falls $f(x) \geq f(x_0)$ für alle $x \in E$ gilt.
- (b) globales Maximum, falls $f(x) \leq f(x_0)$ für alle $x \in E$ gilt.
- (c) lokales Minimum, falls es ein $r > 0$ gibt mit $f(x) \geq f(x_0)$ für alle $x \in E \cap K(x_0, r)$.
- (d) lokales Maximum, falls es ein $r > 0$ gibt mit $f(x) \leq f(x_0)$ für alle $x \in E \cap K(x_0, r)$.
- (e) lokales bzw. globales Extremum, falls f in x_0 ein lokales bzw. globales Minimum oder Maximum besitzt.

Wie kann man nun die Extremstellen und Extremwerte eines Skalarfelds, beispielsweise

$$f(x, y) = (x^2 + 2y^2)e^{-x^2 - y^2},$$

vgl. Abbildung 2.5, bestimmen? Die Antwort ist erfreulich: die Situation ist im Prinzip genauso wie in Erinnerung 2.42 dargestellt, man muss die beteiligten Ableitungen nur durch die richtigen mehrdimensionalen Dinge ersetzen. Das wollen wir nun herausarbeiten. Zuerst vermerken wir wieder, dass in jeder Extremstelle im Inneren des Definitionsbereichs die Ableitung verschwindet.

2 Mehrdimensionale Differentialrechnung

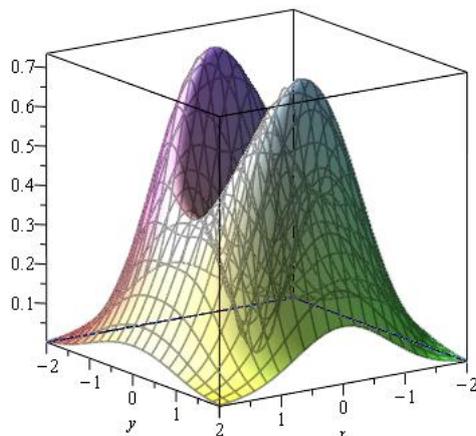


Abbildung 2.5: Der Graph der Funktion $f(x, y) = (x^2 + 2y^2)e^{-x^2 - y^2}$

Satz 2.44. Sei $f : D \rightarrow \mathbb{R}$ stetig partiell differenzierbar. Ist $x_0 \in D$ eine lokale Extremstelle von f , so gilt $\nabla f(x_0) = 0$.

Warnung 2.45. Genau wie in einer Dimension ist die Umkehrung dieser Aussage falsch! Gilt $\nabla f(x_0) = 0$, so ist x_0 nicht notwendigerweise eine lokale Extremstelle von f . Ein Beispiel sehen Sie anhand der Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = x^2 - y^2$ in Abbildung 2.6. Für diese Funktion gilt im Ursprung $\nabla f(0, 0) = 0$, aber der Graph hat kein lokales Extremum, sondern einen Sattelpunkt. (Sollte Ihnen bisher unklar gewesen sein, warum man solche Punkte Sattelpunkte nennt, wird es vielleicht bei der Betrachtung dieses Bildes klarer.)

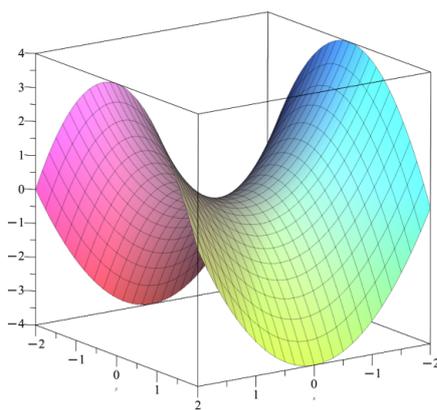


Abbildung 2.6: Der Graph der Funktion $f(x, y) = x^2 - y^2$

Trotzdem ist obiger Satz hilfreich, denn damit ist sichergestellt, dass an allen Punkten,

2.6 Bestimmen von Extrema in mehreren Variablen

an denen der Gradient nicht Null ist, auf keinen Fall ein Extremum vorliegen kann. Durch dieses Kriterium kann man die Untersuchung also meistens auf einige wenige Kandidaten reduzieren. Diese bekommen nun den üblichen Namen.

Definition 2.46. Sei $f : D \rightarrow \mathbb{R}$ ein stetig partiell differenzierbares Skalarfeld. Jedes $x \in D$ mit $\nabla f(x) = 0$ heißt kritischer Punkt von f .

Beispiel 2.47. Sei $f(x, y) = x^2 + 2y^2 - x^2y$. Der Gradient ist

$$\nabla f(x, y) = (2x - 2xy, 4y - x^2) = (2x(1 - y), 4y - x^2).$$

Die kritischen Punkte von f ergeben sich als die Lösungen des Gleichungssystems $\nabla f(x, y) = (0, 0)$, d. h.

$$\begin{cases} 2x(1 - y) = 0 \\ 4y - x^2 = 0 \end{cases}.$$

Zur Lösung betrachtet man zuerst den Fall $x = 0$. Dann ist die erste Gleichung erfüllt und die zweite liefert $4y = 0$, also auch $y = 0$. Das liefert schon mal den kritischen Punkt $(0, 0)^T$. Ist $x \neq 0$, so können wir in der ersten Gleichung durch x dividieren und erhalten $1 - y = 0$, d. h. $y = 1$. Die zweite Gleichung liefert dann $x^2 = 4$ und wir haben zusammen drei kritische Punkte gefunden:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Wie sieht man nun, ob und welche Extremstellen hier vorliegen? Um in der Analogie zum eindimensionalen Resultat zu bleiben, müssten wir wie in Teil (b) von Erinnerung 2.42 etwas über das positive oder negative Vorzeichen der zweiten Ableitungen aussagen. Diese schauen wir uns jetzt genauer an.

Wie wir im letzten Abschnitt festgestellt haben, gibt es viele Möglichkeiten, partielle Ableitungen zweiter Ordnung zu bilden. Welche zweiten Ableitung muss man also anschauen?

Ist $D \subseteq \mathbb{R}^m$ offen und $f : D \rightarrow \mathbb{R}$ eine Funktion, so ist die erste Ableitung von f der Zeilenvektor $\nabla f : D \rightarrow \mathbb{R}^m$ und damit ein Vektorfeld mit m Komponenten. Es gibt damit m^2 partielle Ableitungen zweiter Ordnung, denn jede der m Komponenten des Gradienten kann wieder nach jeder der m Variablen abgeleitet werden. All diese Ableitungen fassen wir erst einmal zusammen.

Definition 2.48. Sei $f : D \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar. Die Matrix

$$H_f(x) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_m}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1}(x) & \frac{\partial^2 f}{\partial x_m \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_m \partial x_m}(x) \end{pmatrix} \in M_{m,m}(\mathbb{R})$$

heißt Hesse-Matrix von f an der Stelle x .

2 Mehrdimensionale Differentialrechnung

Man beachte, dass die Hesse-Matrix auch insofern eine sinnvolle Betrachtung der zweiten Ableitung von f ist, weil sie die Jacobimatrix des Gradienten von f ist: $H_f = J_{\nabla f}$. Sie ist damit die Ableitung der Ableitung von f .

Beispiel 2.49. Wir betrachten weiter die Funktion $f(x, y) = x^2 + 2y^2 - x^2y$ aus Beispiel 2.47. Deren Hesse-Matrix H_f ergibt sich aus dem dort bestimmten Gradienten $\nabla f(x, y) = (2x - 2xy, 4y - x^2)$ zu

$$H_f(x, y) = \begin{pmatrix} 2 - 2y & -2x \\ -2x & 4 \end{pmatrix}.$$

Es gilt also an den kritischen Stellen:

$$H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}, \quad H_f(2, 1) = \begin{pmatrix} 0 & -4 \\ -4 & 4 \end{pmatrix} \quad \text{und} \quad H_f(-2, 1) = \begin{pmatrix} 0 & 4 \\ 4 & 4 \end{pmatrix}.$$

Nun ist aber in keiner Weise offensichtlich, was hier „positiv“ oder „negativ“ bedeuten soll. Dafür müssen wir noch ein wenig weiter ausholen.

Nach dem Satz von Schwarz, vgl. Satz 2.34, hat die Hesse-Matrix die für das Weitere entscheidende Eigenschaft, dass jeweils der Eintrag an der (j, k) -ten und an der (k, j) -ten Stelle übereinstimmen: $\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_j}$. Sie ist also immer eine symmetrische Matrix und damit diagonalisierbar, vgl. Satz 1.40. Damit beschreiben die Eigenwerte und Eigenvektoren die Matrix vollständig und wir können anhand der Eigenwerte ein Konzept für diese Matrizen einführen, das als Ersatz für positiv/negativ taugt.

Definition 2.50. Eine symmetrische Matrix $A = (a_{ij}) \in \mathbf{M}_{m,m}(\mathbb{R})$, d.h. $a_{ij} = a_{ji}$ für $i, j = 1, \dots, m$, heißt

- (a) positiv definit, falls alle ihre Eigenwerte positiv (d. h. > 0) sind.
- (b) negativ definit, falls alle ihre Eigenwerte negativ (d. h. < 0) sind.
- (c) indefinit, falls A positive und negative Eigenwerte besitzt.

Warnung 2.51. Hat die Hesse-Matrix den Eigenwert Null und alle anderen Eigenwerte ein identisches Vorzeichen, passt sie in keines dieser Schemata, es gibt also symmetrische Matrizen, die weder positiv definit, noch negativ definit noch indefinit sind!

Nun können wir den Satz formulieren, der Teil (b) von Erinnerung 2.42 ins mehrdimensionale hebt. Das Schöne ist, dass er nun (bis auf das Wörtchen „definit“) im Prinzip genauso aussieht wie im eindimensionalen.

Satz 2.52. Sei $f : D \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar und $x_0 \in D$ ein kritischer Punkt von f , d. h. es ist $\nabla f(x_0) = 0$. Dann gelten die folgenden Kriterien:

- (a) Ist $H_f(x_0)$ negativ definit, so liegt in x_0 ein lokales Maximum von f vor.

2.6 Bestimmen von Extrema in mehreren Variablen

- (b) Ist $H_f(x_0)$ positiv definit, so liegt in x_0 ein lokales Minimum von f vor.
- (c) Ist $H_f(x_0)$ indefinit, so liegt in x_0 kein lokales Extremum (sondern ein Sattelpunkt) vor.

Schaut man den Satz noch mal genauer an, stellt man sogar fest, dass er ein bisschen mehr kann als sein eindimensionales Pendant. Dort gab es nur die Fälle zweite Ableitung positiv, also Minimum, zweite Ableitung negativ, also Maximum und zweite Ableitung Null, also keine Ahnung. Der keine-Ahnung-Fall splittet sich hier weiter auf: Ist die Matrix indefinit, hat also sowohl einen positiven als auch einen negativen Eigenwert, so kann man sicher sagen, dass es ein Sattelpunkt ist. Wirklich keine Aussage trifft der Satz nur noch in dem Fall, dass die Hesse-Matrix Null als Eigenwert hat und alle anderen Eigenwerte entweder alle positiv oder alle negativ sind.

Um diesen Satz gut anwenden zu können, sind natürlich Kriterien für die Definitheit von Matrizen, für die man nicht mühsam alle Eigenwerte ausrechnen muss, praktisch. Tatsächlich kann man, ob eine Matrix positiv oder negativ definit ist, an bestimmten Determinanten ablesen. Das ist ein Teil der Aussage des folgenden Satzes.

Satz 2.53. Sei $A = (a_{ij})_{i,j=1,\dots,m} \in M_{m,m}(\mathbb{R})$ symmetrisch. Dann sind die folgenden Aussagen äquivalent:

- (a) Die Matrix A ist positiv (resp. negativ) definit.
- (b) Es gilt $v^T \cdot Av > 0$ (resp. < 0) für alle $v \in \mathbb{R}^m \setminus \{0\}$.
- (c) Es gilt

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0 \quad \text{für alle } 1 \leq k \leq m$$

(resp.

$$(-1)^k \det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0 \quad \text{für alle } 1 \leq k \leq m$$

im negativ definiten Fall.)

Beispiel 2.54. Wir untersuchen die Funktion $f(x, y) = (x^2 + 2y^2)e^{-x^2-y^2}$, deren Graph in Abbildung 2.5 abgebildet ist, auf lokale Extrema.

Zuerst suchen wir die kritischen Punkte. Dazu berechnen wir den Gradienten

$$\begin{aligned} \nabla f(x) &= (2xe^{-x^2-y^2} + (x^2 + 2y^2)e^{-x^2-y^2} \cdot (-2x), \\ &\quad 4ye^{-x^2-y^2} + (x^2 + 2y^2)e^{-x^2-y^2} \cdot (-2y)) \\ &= e^{-x^2-y^2} (2x - 2x^3 - 4xy^2, 4y - 2x^2y - 4y^3) \\ &= e^{-x^2-y^2} (2x(1 - x^2 - 2y^2), 2y(2 - x^2 - 2y^2)). \end{aligned}$$

2 Mehrdimensionale Differentialrechnung

Der Faktor $e^{-x^2-y^2}$ ist niemals Null. Es gilt also, alle $(x_0, y_0) \in \mathbb{R}^2$ zu finden, die

$$2x_0(1 - x_0^2 - 2y_0^2) = 0 \quad \text{und} \quad 2y_0(2 - x_0^2 - 2y_0^2) = 0 \quad (2.4)$$

erfüllen. Wir untersuchen dazu zwei Fälle.

Fall 1: $x_0 = 0$. Dann ist die erste Gleichung in (2.4) erfüllt. Die Zweite vereinfacht sich zu $2y_0(2 - 2y_0^2) = 0$, oder $y_0(1 - y_0^2) = 0$. Nun gilt entweder $y_0 = 0$ oder $y_0^2 = 1$. In anderen Worten, es gilt

$$y_0 \in \{-1, 0, 1\}.$$

Fall 2: $x_0 \neq 0$. Die erste Gleichung oben impliziert $x_0^2 + 2y_0^2 = 1$, also $x_0^2 = 1 - 2y_0^2$. Wir ersetzen den Term x_0^2 in der zweiten Gleichung (2.4) durch diesen Ausdruck und erhalten

$$0 = y_0(2 - (1 - 2y_0^2) - 2y_0^2) = y_0(1 + 2y_0^2 - 2y_0^2) = y_0.$$

Kurzum, es muss $y_0 = 0$ gelten. Wir nutzen diese Erkenntnis, und erhalten aus der ersten Gleichung in (2.4), dass $x_0(1 - x_0^2) = 0$ gelten muss. Wie im ersten Fall folgt

$$x_0 \in \{-1, 1\}.$$

Abschließend halten wir fest, dass die kritischen Punkt von f genau

$$\left\{ \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} \quad (2.5)$$

sind.

Um zu entscheiden, ob es sich bei einem dieser Punkte um ein lokales Extremum handelt, müssen wir die Hesse-Matrix berechnen. Zunächst berechnen wir die zweiten Ableitungen und erhalten

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x} \left((2x - 2x^3 - 4xy^2)e^{-x^2-y^2} \right) \\ &= (2 - 6x^2 - 4y^2)e^{-x^2-y^2} + (2x - 2x^3 - 4xy^2)e^{-x^2-y^2} \cdot (-2x) \\ &= (2 - 6x^2 - 4y^2 - 4x^2 + 4x^4 + 8x^2y^2)e^{-x^2-y^2} \\ &= (2 - 10x^2 - 4y^2 + 4x^4 + 8x^2y^2)e^{-x^2-y^2}, \end{aligned}$$

sowie

$$\begin{aligned} \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y} \left((2x - 2x^3 - 4xy^2)e^{-x^2-y^2} \right) \\ &= (-8xy)e^{-x^2-y^2} + (2x - 2x^3 - 4xy^2)e^{-x^2-y^2} \cdot (-2y) \\ &= (-8xy)e^{-x^2-y^2} + (-4xy + 4x^3y + 8xy^3)e^{-x^2-y^2} \\ &= (-12xy + 4x^3y + 8xy^3)e^{-x^2-y^2} \end{aligned} \quad (2.6)$$

2.6 Bestimmen von Extrema in mehreren Variablen

und

$$\begin{aligned} \frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y} \left((4y - 2x^2y - 4y^3)e^{-x^2-y^2} \right) \\ &= (4 - 2x^2 - 12y^2)e^{-x^2-y^2} + (4y - 2x^2y - 4y^3)e^{-x^2-y^2} \cdot (-2y) \\ &= (4 - 2x^2 - 12y^2 - 8y^2 + 4x^2y^2 + 8y^4)e^{-x^2-y^2} \\ &= (4 - 2x^2 - 20y^2 + 4x^2y^2 + 8y^4)e^{-x^2-y^2}. \end{aligned}$$

Die Berechnung von $\frac{\partial^2 f}{\partial x \partial y}$ können wir uns wegen des Satzes von Schwarz ersparen. Diese partielle Ableitung ist gleich (2.6).

Schließlich werten wir die Hesse-Matrix

$$\begin{aligned} H_f(x, y) &= \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x, y) & \frac{\partial^2 f}{\partial y \partial x}(x, y) \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) & \frac{\partial^2 f}{\partial y^2}(x, y) \end{pmatrix} \\ &= e^{-x^2-y^2} \begin{pmatrix} 2 - 10x^2 - 4y^2 + 4x^4 + 8x^2y^2 & -12xy + 4x^3y + 8xy^3 \\ -12xy + 4x^3y + 8xy^3 & 4 - 2x^2 - 20y^2 + 4x^2y^2 + 8y^4 \end{pmatrix} \end{aligned}$$

an den fünf Punkten aus (2.5) aus. Wir erhalten nach einer kurzer Rechnung

$$H_f(0, -1) = \begin{pmatrix} -2/e & 0 \\ 0 & -8/e \end{pmatrix},$$

$$H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix},$$

$$H_f(0, 1) = \begin{pmatrix} -2/e & 0 \\ 0 & -8/e \end{pmatrix},$$

$$H_f(-1, 0) = \begin{pmatrix} -4/e & 0 \\ 0 & 2/e \end{pmatrix},$$

$$H_f(1, 0) = \begin{pmatrix} -4/e & 0 \\ 0 & 2/e \end{pmatrix}.$$

Diese fünf Matrizen sind in Diagonalf orm, deshalb sind die Diagonaleinträge genau die Eigenwerte. An den Vorzeichen der Eigenwerte kann man die Definitheit der Matrix ablesen. Wir fassen das Resultat in einer Tabelle zusammen.

(x_0, y_0)	Eigenwerte	$H_f(x_0, y_0)$ ist ...	Der Punkt (x_0, y_0) ist eine ...
$(0, -1)$	$-2/e, -8/e$	negativ definit	lokale Maximalstelle
$(0, 0)$	$2, 4$	positiv definit	lokale Minimalstelle
$(0, 1)$	$-2/e, -8/e$	negativ definit	lokale Maximalstelle
$(-1, 0)$	$-4/e, 2/e$	indefinit	keine lokale Extremalstelle
$(1, 0)$	$-4/e, 2/e$	indefinit	keine lokale Extremalstelle

2 Mehrdimensionale Differentialrechnung

Wir fassen alles noch einmal in einem Kochrezept zusammen.

Methode 2.55. Gegeben ist $D \subseteq \mathbb{R}^m$ offen und eine zweimal stetig differenzierbare Funktion $f : D \rightarrow \mathbb{R}$. Gesucht sind die alle lokalen Extremstellen von f in D .

- (a) Bestimme den Gradienten ∇f und die Hesse-Matrix H_f von f .
- (b) Löse das Gleichungssystem $\nabla f(x) = 0$. Alle Lösungen sind kritische Punkte von f .
- (c) Für jeden kritischen Punkt $x_0 \in D$ von f :
 - (i) Bilde $H_f(x_0)$.
 - (ii) Stelle fest, ob $H_f(x_0)$ positiv definit, negativ definit, indefinit oder gar nichts davon ist, beispielsweise durch Bestimmung aller Eigenwerte (vgl. Definition 2.50) oder Berechnung der Teildeterminanten aus Satz 2.53. Dann gilt:
 - $H_f(x_0)$ positiv definit: f hat in x_0 lokale Minimalstelle,
 - $H_f(x_0)$ negativ definit: f hat in x_0 lokale Maximalstelle,
 - $H_f(x_0)$ indefinit: f hat in x_0 keine Extremstelle (Sattelpunkt),
 - $H_f(x_0)$ ist gar nichts davon: Methode liefert keine Erkenntnis.

2.7 Extremwertprobleme mit Nebenbedingung

Oft werden Extremwerte einer Funktion gesucht, aber die möglichen Belegungen der Variablen sind nicht unabhängig voneinander, sondern genügen weiteren Bedingungen, den sogenannten Nebenbedingungen. Besonders häufig tritt das in wirtschaftstheoretischen Fragestellungen auf, wo es meist solche zusätzlichen Restriktionen für die Lösungen eines Optimierungsproblems gibt. Wir betrachten ein einfaches, zweidimensionales Beispiel.

Beispiel 2.56. Wir suchen die Extremstellen und -werte der Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

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

deren Graph schon in Abbildung 2.6 dargestellt war, eingeschränkt auf den Einheitskreis

$$K := \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 = 1\},$$

d. h. mit der Nebenbedingung $x^2 + y^2 = 1$. Anschaulich suchen wir damit den höchsten und tiefsten Punkt auf der roten Linie in Abbildung 2.7.

2.7 Extremwertprobleme mit Nebenbedingung

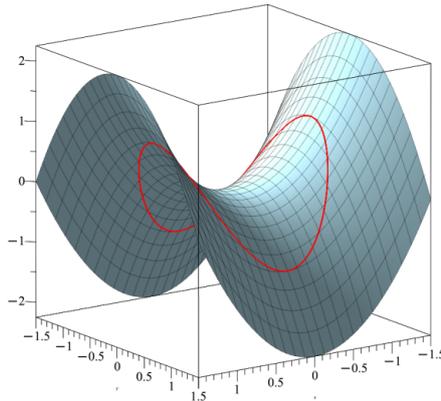


Abbildung 2.7: Maximiere/Minimiere $f(x, y) = x^2 - y^2$ auf dem Einheitskreis

In diesem Spezialfall können wir wie folgt vorgehen: Der Einheitskreis kann durch

$$K = \left\{ \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} : t \in [0, 2\pi] \right\}$$

beschrieben werden. Für jedes t hat der Vektor $(\cos(t), \sin(t))^T$ immer Norm Eins, liegt also auf dem Einheitskreis und lässt man t durch das Intervall $[0, 2\pi]$ gehen, so kommt man einmal um den Kreis herum. Die Abbildung

$$w : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad w(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix},$$

beschreibt also den Einheitskreis. Man nennt das eine *Parametrisierung* des Einheitskreises.

Man kann sich w als Weg vorstellen, der den Kreis im Gegenuhrzeigersinn durchläuft, wenn t von 0 nach 2π geht. Die Maxima/Minima von $f(x, y) = x^2 - y^2$ auf dem Einheitskreis entsprechen damit den Maxima/Minima der Verkettung

$$h(t) = (f \circ w)(t) = \cos(t)^2 - \sin(t)^2.$$

Es handelt sich um eine zweimal stetig differenzierbare Funktion $h : [0, 2\pi] \rightarrow \mathbb{R}$, so dass wir im eindimensionalen Fall sind und lediglich ableiten müssen, um die lokalen Maxima oder Minima zu finden.

Es gilt

$$h'(t) = -2 \cos(t) \sin(t) - 2 \sin(t) \cos(t) = -4 \cos(t) \sin(t).$$

Die Nullstellen der Ableitung im Intervall $[0, 2\pi]$ liegen bei den Nullstellen von Cosinus und Sinus also in $0, \pi/2, \pi, 3\pi/2, 2\pi$. Diese Werte von t entsprechen den Punkten

$$w(0) = w(2\pi) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad w(\pi/2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad w(\pi) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \text{und } w(3\pi/2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad (2.7)$$

2 Mehrdimensionale Differentialrechnung

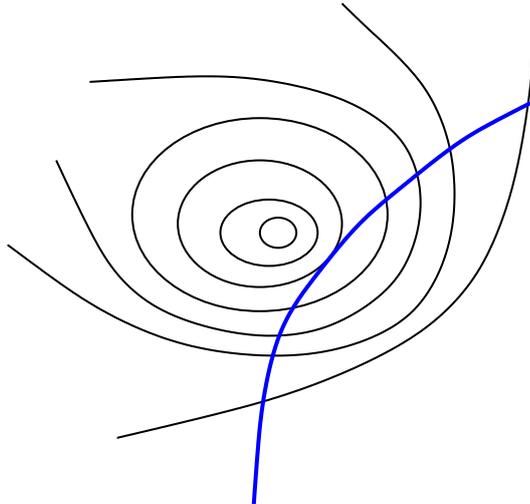


Abbildung 2.8: Das Höhenlinienprofil einer Funktion f (schwarz) und die Linie (dick, blau), die durch die Nebenbedingung $g(x, y) = 0$ beschrieben wird.

mit Werten

$$f(1, 0) = 1, \quad f(0, 1) = -1, \quad f(-1, 0) = 1, \quad \text{und} \quad f(0, -1) = -1.$$

Man kann nun noch die zweite Ableitung hinzuziehen oder folgende Überlegung machen: Die Kreislinie K ist eine beschränkte und abgeschlossene Teilmenge von \mathbb{R}^2 und die Funktion f ist stetig. Also muss diese nach unserem theoretischen Extremstellen-Existenzresultat im Satz von Weierstraß 2.16 mindestens eine Minimal- und eine Maximalstelle auf K haben. Diese können nach der obigen Überlegung nur an den vier kritischen Punkten liegen, also sind $(1, 0)^T$ und $(-1, 0)^T$ globale (und damit auch lokale) Maximalstellen sowie $(0, 1)^T$ und $(0, -1)^T$ globale (und damit auch lokale) Minimalstellen von f .

Unser ad hoc-Ansatz hat so gut funktioniert, weil wir den Einheitskreis explizit parametrisieren konnten. Das geht nicht für jede Nebenbedingung so direkt und daher suchen wir ein anderes Kriterium, um die kritischen Stellen für Extrema mit Nebenbedingung zu bestimmen. Wir gehen dabei davon aus, dass die Nebenbedingung in Form einer Gleichung gegeben ist, sich also mit einer geeigneten Funktion g schreiben lässt als $g(x, y) = 0$. Da man so ziemlich jede Bedingung in eine Gleichung umschreiben kann, ist das eine sehr schwache Annahme, insbesondere viel weniger, als die Anforderung, dass man die Bedingung parametrisieren kann. In obigem Beispiel wäre diese Funktion g als $g(x, y) = x^2 + y^2 - 1$ wählbar gewesen.

Wir betrachten das Problem anschaulich anhand von Abbildung 2.8. Dort ist in Schwarz das Höhenlinienprofil einer Funktion eingezeichnet und die dickere blaue Linie markiert die Punkte, die die Nebenbedingung erfüllen, also alle Punkte mit $g(x, y) = 0$. Gesucht ist der höchste Punkt auf dem blauen Wanderweg durch das Gebirge, das durch die

2.7 Extremwertprobleme mit Nebenbedingung

Höhenlinien der Funktion gegeben ist. Der Weg geht von unten kommend erst bergauf und schneidet dabei verschiedene Höhenlinien, bis er eine Höhenlinie erreicht ohne sie zu schneiden, ab dann geht es wieder abwärts und der Weg schneidet dieselben Höhenlinien in der entgegengesetzten Richtung. Der höchste Punkt auf dem Weg ist offensichtlich genau an der Stelle erreicht, an dem dieser die Höhenlinie erreicht ohne sie zu schneiden.

An dieser Stelle muss der durch die Nebenbedingung gegebene Weg sich genau tangential an die Höhenlinie anschmiegen. Unser gesuchter Punkt ist also dadurch gekennzeichnet, dass die Höhenlinie von f und der Wanderweg, der eine Höhenlinie von g darstellt, genau in dieselbe Richtung zeigen. Nun müssen wir uns nur noch daran erinnern, dass der Gradient einer Funktion in jedem Punkt immer genau senkrecht auf der Höhenlinie der Funktion steht, vgl. Bemerkung 2.27. An unserem gesuchten Punkt, müssen also die Gradienten von f und von g entweder in dieselbe oder genau entgegengesetzte Richtungen schauen. Egal wie, sie müssen auf einer Geraden liegen, d. h. einer muss ein (positives oder negatives) Vielfaches des anderen sein!

Diese Überlegung ist schon die ganze Magie hinter dem folgenden Satz, der die übliche Methode zur Bestimmung von Extremstellen unter Nebenbedingungen begründet.

Satz 2.57 (Methode von Lagrange). *Seien $D \subseteq \mathbb{R}^m$ eine offene Menge und $f, g : D \rightarrow \mathbb{R}$ stetig differenzierbar. Hat f unter der Nebenbedingung $g(x) = 0$ ein Extremum im Punkt $x_0 \in D$ und gilt $\nabla g(x_0) \neq 0$, so existiert ein $\lambda \in \mathbb{R}$ mit*

$$\nabla f(x_0) = \lambda \nabla g(x_0).$$

Der Wert λ in obigem Satz heißt dann *Lagrange-Multiplikator*.

Methode 2.58 (Extrema unter Nebenbedingungen mit Lagrange). Gegeben ist eine Funktion $f : \mathbb{R}^m \rightarrow \mathbb{R}$ und eine Funktion $g : \mathbb{R}^m \rightarrow \mathbb{R}$ und gesucht sind die Extremstellen von f unter der Nebenbedingung $g(x) = 0$. Dazu geht man wie folgt vor:

- (a) Bestimme die Gradienten ∇f und ∇g .
- (b) Überprüfe die Voraussetzung des Satzes 2.57 von Lagrange: Für alle $x \in \mathbb{R}^m$, die die Nebenbedingung erfüllen, d. h. für die $g(x) = 0$ gilt, ist $\nabla g(x) \neq 0$.
- (c) Löse das Gleichungssystem mit $m + 1$ Gleichungen und $m + 1$ Variablen, dass sich aus den Gleichungen $\nabla f(x) = \lambda \nabla g(x)$ und der Nebenbedingung $g(x) = 0$ ergibt.
- (d) Dann kann man sicher sagen, dass die Funktion höchstens an den Stellen, die sich in (c) ergeben, eine Extremstelle haben kann.

Das tatsächliche Vorliegen einer Extremstelle ist im Allgemeinen schwerer zu begründen. Ein einfacher Weg eröffnet sich, wenn die von der Nebenbedingung beschriebene

2 Mehrdimensionale Differentialrechnung

Teilmenge des \mathbb{R}^m abgeschlossen und beschränkt ist. Dann weiß man nach Satz 2.16, dass auf jeden Fall ein Maximum und Minimum existieren muss. Da es jeweils nur in einer der Lösungen aus (c) liegen kann, ist dann die Lösung mit dem kleinsten Funktionswert Minimalstelle und die mit dem größten Funktionswert Maximalstelle.

Beispiel 2.59. (a) Sei $x, y, z \geq 0$. Wir betrachten einen Quader Q im Raum \mathbb{R}^3 , der auf der x - y -Ebene liegt und Eckpunkte bei $(\pm x, \pm y, z)$ und $(\pm x, \pm y, 0)$ besitzt. Das Volumen von Q ist dann $4xyz$. Gesucht ist ein solcher Quader mit maximalem Volumen, welcher in der Einheitskugel liegt. Klar ist, dass für ein maximales Volumen, die Ecken auf der Einheitskugel, also dem Rand der Einheitskugel liegen müssen, d. h. wir wollen die Funktion

$$f(x, y, z) = 4xyz$$

unter allen Argumenten x, y, z maximieren, für die $x^2 + y^2 + z^2 = 1$ gilt. Die Nebenbedingung ist also $g(x, y, z) = 0$ mit

$$g(x, y, z) = x^2 + y^2 + z^2 - 1.$$

Die Funktion f ist stetig und die Einheitskugel, die durch die Nebenbedingung beschrieben wird, ist abgeschlossen und beschränkt. Wegen Satz 2.16 muss also ein solches Maximum existieren, wir müssen es nur noch finden. Dazu verwenden wir die Lagrange-Methode.

Zunächst ist

$$\nabla f(x, y, z) = (4yz, 4xz, 4xy) \quad \text{und} \quad \nabla g(x, y, z) = (2x, 2y, 2z).$$

Weiter gilt $\nabla g(x, y, z) = (0, 0, 0)$ nur wenn $(x, y, z) = (0, 0, 0)$ ist, und dieser Punkt liegt nicht auf der Einheitskugel. Wir dürfen die Lagrange-Methode also anwenden. Dazu müssen wir das Gleichungssystem aus

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) \quad \text{und} \quad g(x, y, z) = 0$$

lösen. Konkret haben wir also die Lösungen von

$$\begin{cases} 4yz = 2\lambda x \\ 4xz = 2\lambda y \\ 4xy = 2\lambda z \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

zu bestimmen. Wir dürfen dafür voraussetzen, dass alle drei Werte x, y, z ungleich Null sind, denn sonst hat unser Quader Volumen Null und das ist sicher nicht maximal. Dann liefern die ersten drei Gleichungen

$$\lambda = 2 \frac{yz}{x} = 2 \frac{xz}{y} = 2 \frac{xy}{z}.$$

2.7 Extremwertprobleme mit Nebenbedingung

Durch Multiplikation mit xy bekommen wir hier aus dem zweiten Gleichheitszeichen $y^2z = x^2z$ und damit $y^2 = x^2$ und wegen $x, y \geq 0$ auch $x = y$. Damit liest sich das dritte Gleichheitszeichen als

$$z = \frac{xz}{x} = \frac{xz}{y} = \frac{xy}{z} = \frac{x^2}{z},$$

also folgt auch $x^2 = z^2$ und wieder wegen $x, z \geq 0$ auch $x = y = z$.

Aus der Nebenbedingung $x^2 + y^2 + z^2 = 1$ folgt dann $3x^2 = 1$ also $x = 1/\sqrt{3}$ und wir haben den gesuchten Eckpunkt

$$x = y = z = \frac{1}{\sqrt{3}} \quad \text{mit} \quad f\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) = 4 \cdot \left(\frac{1}{\sqrt{3}}\right)^3 = \frac{4}{\sqrt{3}^3} = \frac{4}{3\sqrt{3}}$$

gefunden.

- (b) Schließlich behandeln wir das Beispiel vom Beginn dieses Abschnitts auch noch einmal mit der Methode von Lagrange. D. h. wir suchen die Extremstellen von $f(x, y) = x^2 - y^2$ unter der Nebenbedingung $g(x, y) = x^2 + y^2 - 1 = 0$.

Wieder müssen nach Satz 2.16 beide Extremwerte existieren, denn die Nebenbedingung beschreibt die Kreislinie des Einheitskreises, die beschränkt und abgeschlossen ist, und die Funktion f ist stetig. Hier ist

$$\nabla f(x, y) = (2x, -2y) \quad \text{und} \quad \nabla g(x, y) = (2x, 2y)$$

und die Methode von Lagrange ist anwendbar, denn die einzige Nullstelle von ∇g liegt in $(0, 0)$ und erfüllt damit nicht die Nebenbedingung. Das Gleichungssystem aus $\nabla f(x) = \lambda \nabla g(x)$ und $g(x) = 0$ lautet hier

$$\begin{cases} 2x = \lambda 2x \\ -2y = \lambda 2y \\ x^2 + y^2 = 1. \end{cases}$$

Zur Lösung unterscheiden wir zwei Fälle:

1. Fall: $x = 0$. Dann ist die erste Gleichung automatisch erfüllt, die dritte liefert $y = \pm 1$ und die zweite lautet dann $\mp 2 = \pm 2\lambda$, also $\lambda = -1$. Damit haben wir zwei kritische Punkte in $(0, 1)^T$ und $(0, -1)^T$.

2. Fall: $x \neq 0$. Dann können wir die erste Gleichung durch x teilen und bekommen $\lambda = 1$. Damit liest sich die zweite Gleichung als $2y = -2y$, was sofort zu $y = 0$ führt. Die dritte Gleichung sagt dann $x = \pm 1$ und wir haben mit $(1, 0)^T$ und $(-1, 0)^T$ zwei weitere kritische Punkte.

Da Maximum und Minimum nach der Vorüberlegung existieren müssen, ergibt ein Vergleich der Funktionswerte

$$f(1, 0) = f(-1, 0) = 1 \quad \text{und} \quad f(0, 1) = f(0, -1) = -1,$$

2 Mehrdimensionale Differentialrechnung

dass $(0, 1)^T$ und $(0, -1)^T$ die Minimalstellen und $(1, 0)^T$ und $(-1, 0)^T$ die Maximalstellen von f unter der Nebenbedingung $g = 0$ sind. Dies stimmt mit unserer ersten Rechnung überein.

3 Mehrdimensionale Integration

In der Höheren Mathematik I wurde gezeigt, wie man eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ integriert. Das bestimmte Integral

$$\int_a^b f(x) \, dx$$

lässt sich als Fläche unter dem Graphen von f interpretieren. In diesem Kapitel möchten wir den Integrationsbegriff auf Skalarfelder ausweiten. Dabei beschränken wir uns auf Funktionen, die auf Teilmengen des \mathbb{R}^2 und \mathbb{R}^3 definiert sind.

3.1 Zweidimensionale Integration

Sei $D \subseteq \mathbb{R}^2$ und $f : D \rightarrow \mathbb{R}$ eine stetige Funktion. Wir möchten das Integral

$$\int_D f \, d(x, y)$$

so definieren, dass es sich für Funktionen f mit positiven Werten als Volumen unter dem Graphen, also dem Volumen der Menge

$$\{(x, y, z)^T \in D \times \mathbb{R} : 0 \leq z \leq f(x, y)\}$$

interpretieren lässt. Der Hauptunterschied zum eindimensionalen Fall zeigt sich schnell: Die Menge D , über die wir integrieren möchten, kann viel komplizierter sein. Wir beschränken uns daher im Folgenden auf gutartige Mengen, die es erlauben das zweidimensionale Integral in zwei eindimensionale Integrale zu zerlegen.

Definition 3.1. Sei $D \subseteq \mathbb{R}^2$.

- (a) Wir nennen D y -projizierbar, falls ein Intervall $[a, b]$ in \mathbb{R} und Funktionen $\underline{y}, \bar{y} : [a, b] \rightarrow \mathbb{R}$ existieren mit

$$D = \{(x, y)^T \in \mathbb{R}^2 : x \in [a, b] \text{ und } \underline{y}(x) \leq y \leq \bar{y}(x)\}. \quad (3.1)$$

- (b) Umgekehrt nennen wir D x -projizierbar, falls ein Intervall $[c, d]$ in \mathbb{R} und Funktionen $\underline{x}, \bar{x} : [c, d] \rightarrow \mathbb{R}$ existieren mit

$$D = \{(x, y)^T \in \mathbb{R}^2 : y \in [c, d] \text{ und } \underline{x}(y) \leq x \leq \bar{x}(y)\}. \quad (3.2)$$

3 Mehrdimensionale Integration

(c) Wir nennen D eine Standardmenge, falls D sowohl x - als auch y -projizierbar ist.

Beispiel 3.2. (a) Die Menge

$$D := \{(x, y)^T \in \mathbb{R}^2 : x \in [2, 4] \text{ und } x^2 \leq y \leq x^3\}$$

ist y -projizierbar mit $[a, b] = [2, 4]$, $\underline{y}(x) = x^2$ und $\bar{y}(x) = x^3$. Diese Menge ist in Abbildung 3.1 abgebildet.

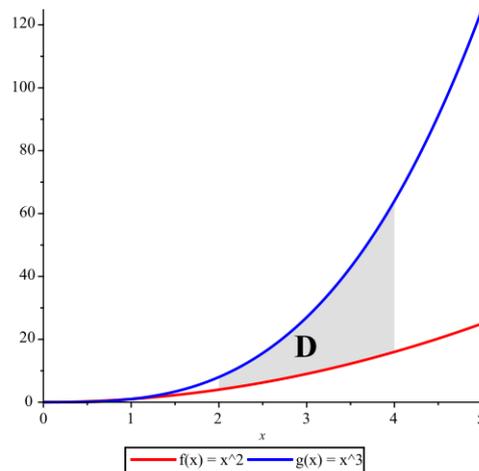


Abbildung 3.1: Die y -projizierbare Menge D aus Beispiel 3.2 (a)

(b) Nicht jede Menge ist irgendwie projizierbar. Beispielsweise ist die Menge

$$E := \{(x, y)^T \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4\},$$

vgl. Abbildung 3.2, weder x - noch y -projizierbar. In diesem Fall kann man das Problem allerdings leicht umgehen. Beispielsweise kann man E in den Teil oberhalb und den Teil unterhalb der x -Achse zerlegen. Die beiden Hälften von E sind dann y -projizierbar und man bekommt das Integral über E als Summe dieser beiden Integrale.

(c) Die Kreisscheibe

$$F := \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$$

ist eine Standardmenge. Sie lässt sich durch

$$F = \left\{ (x, y)^T \in \mathbb{R}^2 : x \in [-1, 1] \text{ und } \underbrace{-\sqrt{1-x^2}}_{=\underline{y}(x)} \leq y \leq \underbrace{\sqrt{1-x^2}}_{=\bar{y}(x)} \right\}$$

und

$$F = \left\{ (x, y)^T \in \mathbb{R}^2 : y \in [-1, 1] \text{ und } \underbrace{-\sqrt{1-y^2}}_{=\underline{x}(y)} \leq x \leq \underbrace{\sqrt{1-y^2}}_{=\bar{x}(y)} \right\}$$

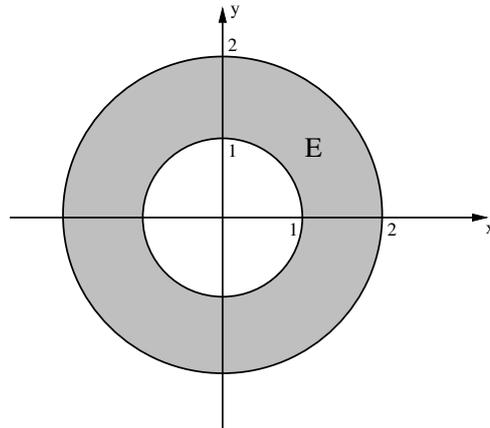


Abbildung 3.2: Eine Menge, die weder x - noch y -projizierbar ist.

beschreiben.

Definition 3.3. Ist $D \subseteq \mathbb{R}^2$ wie in (3.1) eine y -projizierbare Menge und $f : D \rightarrow \mathbb{R}$ eine stetige Funktion, dann integrieren wir „scheibchenweise“:

$$\int_D f \, d(x, y) = \int_a^b \left(\int_{\underline{y}(x)}^{\bar{y}(x)} f(x, y) \, dy \right) dx.$$

Ist D eine x -projizierbare Menge wie in (3.2), so integrieren wir stattdessen zuerst über die x -Koordinate und dann über die y -Koordinate:

$$\int_D f \, d(x, y) = \int_c^d \left(\int_{\underline{x}(y)}^{\bar{x}(y)} f(x, y) \, dx \right) dy.$$

Der folgende Satz von Fubini besagt, dass für Standardmengen die Integrationsreihenfolge nicht relevant ist.

Satz 3.4 (Satz von Fubini). Sei $D \subseteq \mathbb{R}^2$ eine Standardmenge und $f : D \rightarrow \mathbb{R}$ ein stetiges Skalarfeld. Dann hängt das Integral nicht von der Reihenfolge der Integration ab. Konkret: Für

$$\begin{aligned} D &= \{(x, y)^T \in \mathbb{R}^2 : x \in [a, b] \text{ und } \underline{y}(x) \leq y \leq \bar{y}(x)\} \\ &= \{(x, y)^T \in \mathbb{R}^2 : y \in [c, d] \text{ und } \underline{x}(y) \leq x \leq \bar{x}(y)\} \end{aligned}$$

gilt

$$\int_D f(x, y) \, d(x, y) = \int_a^b \int_{\underline{y}(x)}^{\bar{y}(x)} f(x, y) \, dy \, dx = \int_c^d \int_{\underline{x}(y)}^{\bar{x}(y)} f(x, y) \, dx \, dy.$$

3 Mehrdimensionale Integration

Beispiel 3.5. (a) Das Quadrat $D = [0, 1] \times [0, 1]$ ist eine Standardmenge. Wir möchten $f(x, y) = x^3 - y^3 - x^2y + 2$ auf D integrieren. Es gilt

$$\begin{aligned} \int_D f(x, y) \, d(x, y) &= \int_0^1 \int_0^1 (x^3 - y^3 - x^2y + 2) \, dy \, dx \\ &= \int_0^1 \left(x^3y - \frac{1}{4}y^4 - \frac{1}{2}x^2y^2 + 2y \right) \Big|_{y=0}^{y=1} \, dx \\ &= \int_0^1 \left(x^3 - \frac{1}{4} - \frac{1}{2}x^2 + 2 \right) \, dx \\ &= \left(\frac{x^4}{4} - \frac{x}{4} - \frac{x^3}{6} + 2x \right) \Big|_{x=0}^{x=1} \\ &= \frac{1}{4} - \frac{1}{4} - \frac{1}{6} + 2 = \frac{11}{6}. \end{aligned}$$

(b) Die Menge

$$D = \{(x, y)^T \in \mathbb{R}^2 : x \in [0, 2] \text{ und } x^2 \leq y \leq 4\}$$

steht schon in y -projizierter Form da. Wir integrieren die Funktion $f(x, y) = x^2 + y^2$ über D :

$$\begin{aligned} \int_D f(x, y) \, d(x, y) &= \int_0^2 \int_{x^2}^4 (x^2 + y^2) \, dy \, dx \\ &= \int_0^2 \left(x^2y + \frac{y^3}{3} \right) \Big|_{y=x^2}^{y=4} \, dx \\ &= \int_0^2 \left(4x^2 + \frac{64}{3} - x^4 - \frac{x^6}{3} \right) \, dx \\ &= \left(\frac{4}{3}x^3 + \frac{64}{3}x - \frac{1}{5}x^5 - \frac{x^7}{21} \right) \Big|_{x=0}^{x=2} \\ &= \frac{4288}{105}. \end{aligned}$$

Als nächstes wollen wir die von Integralen in einer Dimension bekannte Substitutionsregel übertragen und erinnern uns: Ist $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion und $g : [a, b] \rightarrow [c, d]$ bijektiv, so gilt

$$\int_a^b f(x) \, dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f(g(u))g'(u) \, du.$$

Eine analoge Regel gilt in zwei Dimensionen. Dazu betrachten wir zwei Mengen $D, E \subseteq \mathbb{R}^2$, die x - oder y -projizierbar sind und eine bijektive Abbildung $\Phi : E \rightarrow D$.

Wie vorher die Funktion g rechnet also Φ die Koordinaten von Punkten $(u, v) \in E$, eins-zu-eins in Koordinaten von Punkten

$$\Phi(u, v) = (\Phi_1(u, v), \Phi_2(u, v)) = (x, y)$$

in D um. Ist Φ hinreichend gutartig,¹ so können wir wieder eine Substitutionsregel formulieren:

Rechenregel 3.6 (Substitutionsregel). Seien D, E, Φ wie oben und $f : D \rightarrow \mathbb{R}$ eine stetige Funktion, dann ist

$$\int_D f(x, y) \, d(x, y) = \int_E f(\Phi(u, v)) |\det J_\Phi(u, v)| \, d(u, v),$$

wobei J_Φ die Jacobi-Matrix von Φ ist.

Dabei muss man es am Rand des Integrationsgebiets mit der Bijektivität von Φ nicht ganz genau nehmen, denn Integrale über einzelne Punkte oder Linien sind Null. Wir werden das gleich im folgenden Beispiel sehen.

Beispiel 3.7. (a) Wir möchten die Funktion $f(x, y) = e^{-x^2-y^2}$ auf der Kreisscheibe

$$D := \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$$

integrieren. Diese Aufgabe weist eine Kreissymmetrie auf und in solchen Fällen ist es oft zweckmäßig die folgende Substitution in sogenannte *Polarkoordinaten* zu machen. Dabei werden die Punkte der Ebene nicht durch x - und y -Koordinaten beschrieben, sondern durch ihren Abstand $r > 0$ vom Ursprung und den Winkel $\varphi \in [0, 2\pi]$, den der Vektor $(x, y)^T$ mit der positiven x -Achse bildet, vgl. Abbildung 3.3. Die Umrechnung erfolgt mit

$$\begin{aligned} x &= r \cos \varphi \\ y &= r \sin \varphi. \end{aligned}$$

Damit ist die Substitutionsfunktion Φ gegeben durch

$$\begin{pmatrix} x \\ y \end{pmatrix} = \Phi(r, \varphi) = \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \end{pmatrix}.$$

Mit

$$E = \{(r, \varphi) \in [0, \infty) \times [0, 2\pi] : r \in [0, 1] \text{ und } \varphi \in [0, 2\pi]\} = [0, 1] \times [0, 2\pi].$$

gilt $\Phi(E) = D$, denn E beschreibt alle Punkte mit beliebigem Winkel und Abstand zum Ursprung zwischen Null und Eins und das entspricht genau dem Einheitskreis.

¹Das Vektorfeld Φ muss stetig differenzierbar auf einer offenen Umgebung von E sein

3 Mehrdimensionale Integration

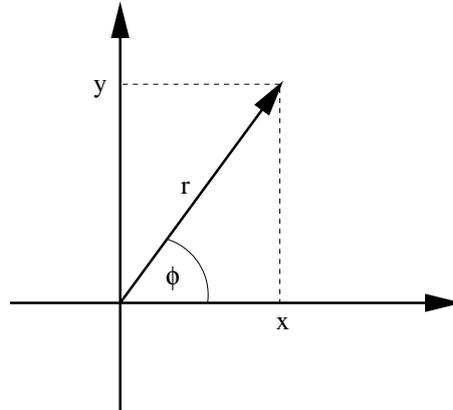


Abbildung 3.3: Die Polarkoordinaten (r, ϕ) eines Punktes $(x, y)^T$ in der Ebene

Für die Anwendung der Substitutionsregel benötigen wir noch die Jacobi-Matrix von Φ . Diese ist

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

und damit gilt

$$\det J_{\Phi}(r, \varphi) = r \cos^2(\varphi) + r \sin^2(\varphi) = r(\cos^2(\varphi) + \sin^2(\varphi)) = r. \quad (3.3)$$

Wegen $x^2 + y^2 = r^2 \cos^2(\varphi) + r^2 \sin^2(\varphi) = r^2$ liefert die Substitutionsregel

$$\int_D e^{-x^2-y^2} d(x, y) = \int_E e^{-r^2} |\det J_{\Phi}(r, \varphi)| d(r, \varphi).$$

Die Menge E ist nun schlicht ein Rechteck, was zu einfachen Integrationsgrenzen führt:

$$\begin{aligned} &= \int_0^{2\pi} \int_0^1 e^{-r^2} r dr d\varphi = \int_0^{2\pi} 1 d\varphi \int_0^1 e^{-r^2} r dr \\ &= 2\pi \int_0^1 e^{-r^2} r dr. \end{aligned}$$

Jetzt wenden wir die eindimensionale Substitutionsregel mit $s = r^2$ an:

$$\begin{aligned} &= 2\pi \int_0^1 \frac{1}{2} e^{-s} ds = 2\pi \left(-\frac{1}{2} e^{-s} \Big|_{s=0}^{s=1} \right) \\ &= 2\pi \left(-\frac{1}{2} e^{-1} + \frac{1}{2} \right) = \pi \left(1 - \frac{1}{e} \right). \end{aligned}$$

Eine Bemerkung zur formalen Korrektheit: Die Voraussetzung für die Substitutionsregel, wie wir sie oben formuliert haben, ist in diesem Beispiel nicht

3.1 Zweidimensionale Integration

erfüllt. Die Funktion Φ ist nicht injektiv auf $[0, 1] \times [0, 2\pi]$, denn die Punkte mit $\varphi = 0$ und $\varphi = 2\pi$ werden jeweils auf denselben Punkt auf der x -Achse abgebildet. Trotzdem kann man die Argumentation mit leichtem Aufwand korrekt begründen und das Resultat ist richtig.

(b) Hier berechnen wir das Integral von $f(x, y) = xy$ auf dem Viertel-Kreisring

$$D = \left\{ (x, y)^T \in \mathbb{R}^2 : x, y \geq 0 \text{ und } \frac{1}{2} \leq \sqrt{x^2 + y^2} \leq 1 \right\},$$

vgl. Abbildung 3.4.

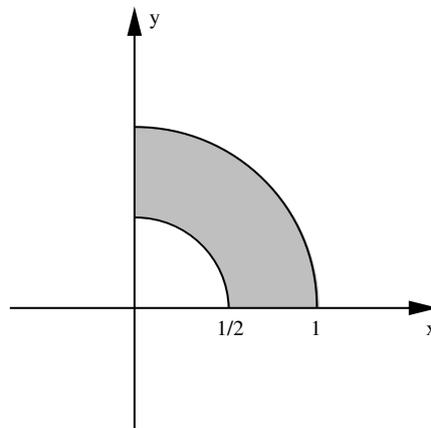


Abbildung 3.4: Die Menge D aus Beispiel 3.7 (b)

Wie im ersten Beispiel benutzen wir Polarkoordinaten (r, φ) . In diesen Koordinaten schreibt sich D als

$$E = \left\{ (r, \varphi) \in [0, \infty) \times [0, 2\pi] : \frac{1}{2} \leq r \leq 1 \text{ und } 0 \leq \varphi \leq \frac{\pi}{2} \right\} = \left[\frac{1}{2}, 1 \right] \times \left[0, \frac{\pi}{2} \right].$$

Wir erhalten

$$\begin{aligned} \int_D xy \, d(x, y) &= \int_E r \cos(\varphi) \cdot r \sin(\varphi) \underbrace{|\det J_\Phi(r, \varphi)|}_{=r} \, dr \, d\varphi \\ &= \int_0^{\pi/2} \int_{1/2}^1 r^3 \cos(\varphi) \sin(\varphi) \, dr \, d\varphi \\ &= \int_0^{\pi/2} \cos(\varphi) \sin(\varphi) \, d\varphi \int_{1/2}^1 r^3 \, dr \end{aligned}$$

3 Mehrdimensionale Integration

Eine Stammfunktion von $\cos(\varphi) \sin(\varphi)$ ist $\frac{1}{2} \cdot \sin^2(\varphi)$. Damit können wir weiterrechnen:

$$= \frac{1}{2} \sin^2(\varphi) \Big|_{\varphi=0}^{\varphi=\pi/2} \cdot \frac{r^4}{4} \Big|_{r=1/2}^{r=1} = \frac{1}{2} (1 - 0) \cdot \left(\frac{1}{4} - \frac{1}{64} \right) = \frac{15}{128}.$$

3.2 Dreidimensionale Integration

Mit drei Dimensionen haben wir im Alltag am häufigsten zu tun, insofern ist es kein Wunder, dass auch das Integrieren in drei Dimensionen eine besondere Relevanz hat. Bezeichnet beispielsweise für eine vernünftige Teilmenge K des \mathbb{R}^3 , wie einen Quader, eine Kugel, einen Zylinder o.ä., $\rho(x, y, z)$ in jedem Punkt $(x, y, z)^T \in K$ die Dichte von K , so berechnet sich die *Masse* von K zu

$$m = \int_K \rho(x, y, z) \, d(x, y, z),$$

das *Volumen* ist

$$V = \int_K 1 \, d(x, y, z)$$

und ist die Masse $m > 0$, so bekommt man den *Schwerpunkt* $(x_s, y_s, z_s)^T$ von K durch die Berechnung der Integrale

$$\begin{aligned} x_s &= \frac{1}{m} \int_K x \rho(x, y, z) \, d(x, y, z), \\ y_s &= \frac{1}{m} \int_K y \rho(x, y, z) \, d(x, y, z), \\ z_s &= \frac{1}{m} \int_K z \rho(x, y, z) \, d(x, y, z). \end{aligned}$$

Analog zum zweidimensionalen Fall integriert man über projizierbare Mengen, um die Größen oben präzise zu berechnen, es wird alles nur um eine weitere Variable aufgepumpt. Projizierbare Mengen lassen sich nun also beispielsweise so beschreiben

$$K = \{(x, y, z)^T \in \mathbb{R}^3 : a \leq z \leq b, \quad \underline{y}(z) \leq y \leq \bar{y}(z), \quad \underline{x}(y, z) \leq x \leq \bar{x}(y, z)\},$$

wobei $\underline{y}, \bar{y}, \underline{x}, \bar{x}$ geeignete Funktionen sind. Alternativ kann man die Reihenfolge der Variablen natürlich auch wieder ändern, also etwa

$$K = \{(x, y, z)^T \in \mathbb{R}^3 : a \leq x \leq b, \quad \underline{y}(x) \leq y \leq \bar{y}(x), \quad \underline{z}(x, y) \leq z \leq \bar{z}(x, y)\}.$$

mit geeigneten Funktionen $\underline{y}, \bar{y}, \underline{z}, \bar{z}$ betrachten. Essentiell ist dabei nur, dass die Randvorgaben jeder einzelnen Variable immer nur von den schon vorher erwähnten Variablen abhängig sein darf und nicht von denen, die noch später kommen. Anstatt eine komplizierte allgemeine Definition hinzuschreiben, verdeutlichen wir das an zwei Beispielen.

Beispiel 3.8. (a) Der Quader

$$Q := \{(x, y, z)^T \in \mathbb{R}^3 : 0 \leq z \leq 1, \quad 0 \leq y \leq 2, \quad 0 \leq x \leq 3\}.$$

ist einfach zu handhaben, da alle Integrationsgrenzen konstant sind. Nehmen wir als Dichte

$$\rho(x, y, z) = xyz,$$

dann ist beispielsweise die Masse gegeben durch

$$\begin{aligned} \int_Q \rho(x, y, z) \, d(x, y, z) &= \int_0^1 \int_0^2 \int_0^3 xyz \, dx \, dy \, dz \\ &= \int_0^1 z \, dz \int_0^2 y \, dy \int_0^3 x \, dx \\ &= \frac{1}{2} z^2 \Big|_{z=0}^{z=1} \cdot \frac{1}{2} y^2 \Big|_{y=0}^{y=2} \cdot \frac{1}{2} x^2 \Big|_{x=0}^{x=3} \\ &= \frac{1}{8} (1-0)(4-0)(9-0) = \frac{36}{8} = \frac{9}{2}. \end{aligned}$$

(b) Der Tetraeder T mit den Eckpunkten $(0, 0, 0)^T$, $(1, 0, 0)^T$, $(0, 1, 0)^T$ und $(0, 0, 1)^T$ lässt sich beschreiben durch

$$T = \{(x, y, z)^T \in \mathbb{R}^3 : 0 \leq x \leq 1, \quad 0 \leq y \leq 1-x, \quad 0 \leq z \leq 1-x-y\}.$$

Sein Volumen ergibt sich also zu

$$\begin{aligned} \int_T 1 \, d(x, y, z) &= \int_0^1 \int_0^{1-x} \int_0^{1-x-y} 1 \, dz \, dy \, dx = \int_0^1 \int_0^{1-x} (1-x-y) \, dy \, dx \\ &= \int_0^1 \left(y - xy - \frac{1}{2} y^2 \right) \Big|_{y=0}^{y=1-x} \, dx \\ &= \int_0^1 \left(1-x - x(1-x) - \frac{1}{2} (1-x)^2 \right) \, dx \\ &= \int_0^1 \left(1-x - x + x^2 - \frac{1}{2} + x - \frac{1}{2} x^2 \right) \, dx \\ &= \int_0^1 \left(\frac{1}{2} - x + \frac{1}{2} x^2 \right) \, dx = \left(\frac{1}{2} x - \frac{1}{2} x^2 + \frac{1}{6} x^3 \right) \Big|_{x=0}^{x=1} = \frac{1}{6}. \end{aligned}$$

Auch die Substitutionsregel (Rechenregel 3.6) gilt genauso in drei Dimensionen. Wir behandeln das im folgenden Beispiel anhand der häufig verwendeten Koordinatenwechsel in Zylinder- und Kugelkoordinaten.

Beispiel 3.9. (a) Wir möchten die Dichtefunktion

$$\rho(x, y, z) = x^2 z^2$$

3 Mehrdimensionale Integration

über den Zylinder

$$Z = \{(x, y, z)^T \in \mathbb{R}^3 : -4 \leq z \leq 4, \quad x^2 + y^2 \leq 1\}$$

integrieren. Diese Menge lässt sich mit einiger Mühe als projizierbar erkennen, denn mit dem Ergebnis aus Beispiel 3.2 (c) gilt:

$$Z = \{(x, y, z)^T \in \mathbb{R}^3 : -4 \leq z \leq 4, \quad 0 \leq y \leq 1, \quad -\sqrt{1-y^2} \leq x \leq \sqrt{1-y^2}\},$$

die Berechnung der Masse von Z wird in dieser Darstellung wegen der Wurzeln aber sehr beschwerlich. Wir verwenden deshalb *Zylinderkoordinaten*, was bedeutet, dass wir analog zu Beispiel 3.7 die Variablen x und y durch die zugehörigen Polarkoordinaten r und φ substituieren und die Variable z unverändert lassen. Insgesamt ersetzen wir also die Koordinaten (x, y, z) durch (r, φ, z) . Das ergibt einen Koordinatenwechsel durch die Abbildung

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \Phi(r, \varphi, z) = \begin{pmatrix} r \cos(\varphi) \\ r \sin(\varphi) \\ z \end{pmatrix}.$$

Auch hier muss man die Determinante der Jacobi-Matrix J_Φ bestimmen. Wegen

$$J_\Phi(r, \varphi, z) = \begin{pmatrix} \cos(\varphi) & -r \sin(\varphi) & 0 \\ \sin(\varphi) & r \cos(\varphi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

bekommt man mit Entwicklung nach der letzten Spalte und derselben Rechnung wie in (3.3) diese Determinante als $\det J_\Phi(r, \varphi, z) = r$.

In Zylinderkoordinaten bekommt die Menge Z die Form

$$S = \{(r, \varphi, z) \in [0, \infty) \times [0, 2\pi] \times \mathbb{R} : -4 \leq z \leq 4, \quad 0 \leq r \leq 1, \quad 0 \leq \varphi \leq 2\pi\}.$$

Also gilt

$$\begin{aligned} \int_Z \rho(x, y, z) \, d(x, y, z) &= \int_Z x^2 z^2 \, d(x, y, z) \\ &= \int_S (r \cos(\varphi))^2 z^2 \cdot |\det(J_\Phi(r, \varphi, z))| \, d(r, \varphi, z) \\ &= \int_S r^2 \cos^2(\varphi) z^2 r \, d(r, \varphi, z) \\ &= \int_{-4}^4 \int_0^1 \int_0^{2\pi} z^2 r^3 \cos^2(\varphi) \, d\varphi \, dr \, dz \\ &= \int_{-4}^4 z^2 \, dz \int_0^1 r^3 \, dr \int_0^{2\pi} \cos^2(\varphi) \, d\varphi \end{aligned}$$

3.2 Dreidimensionale Integration

Hier brauchen wir eine Stammfunktion von $\cos^2(\varphi)$. Diese ist gegeben durch $\frac{1}{2} \cdot (\varphi + \cos(\varphi) \sin(\varphi))$ und wir erhalten

$$\begin{aligned} &= \frac{1}{3} z^3 \Big|_{z=-4}^{z=4} \cdot \frac{1}{4} r^4 \Big|_{r=0}^{r=1} \cdot \frac{1}{2} (\varphi + \cos(\varphi) \sin(\varphi)) \Big|_{\varphi=0}^{\varphi=2\pi} \\ &= \frac{1}{24} (64 - (-64)) \cdot (1 - 0) \cdot (2\pi + 0 - 0) = \frac{32}{3} \pi. \end{aligned}$$

(b) Nun wenden wir uns der Integration über Kugeln zu. Sei also

$$K = \{(x, y, z)^T \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq 1\} \quad (3.4)$$

die Kugel um den Nullpunkt mit Radius 1. Wir möchten

$$\int_K \rho(x, y, z) \, d(x, y, z)$$

mit der Dichtefunktion

$$\rho(x, y, z) = e^{(x^2+y^2+z^2)^{3/2}}$$

bestimmen. Wie wir aus der Geographie wissen, beschreibt man Punkte auf einer Kugeloberfläche am besten durch zwei Winkel, die dann Längen- und Breitengrad genannt werden. Um alle Punkte in der Kugel zu bekommen, müssen wir neben den beiden Winkeln noch den Abstand vom Ursprung angeben. Wir bekommen also die *Kugelkoordinaten* (r, φ, θ) mit

$r \in [0, \infty)$	Radius,
$\varphi \in [-\pi, \pi]$	geographische Länge,
$\theta \in [0, \pi]$	geographische Breite.

Die Transformation in Kugelkoordinaten lautet dann

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \Phi(r, \varphi, \theta) = \begin{pmatrix} r \sin(\theta) \cos(\varphi) \\ r \sin(\theta) \sin(\varphi) \\ r \cos(\theta) \end{pmatrix}.$$

Die Jacobi-Matrix J_Φ und ihre Determinante lassen sich wie üblich berechnen.² Man findet

$$\det J_\Phi = r^2 \sin(\theta).$$

Die Kugel K transformiert in Kugelkoordinaten zu

$$\begin{aligned} L &= \{(r, \varphi, \theta) \in [0, \infty) \times [-\pi, \pi] \times [0, \pi] : r \in [0, 1], \varphi \in [-\pi, \pi], \theta \in [0, \pi]\} \\ &= [0, 1] \times [-\pi, \pi] \times [0, \pi]. \end{aligned}$$

²Das ist eine feige Formulierung des Skript-Autors: Die Rechnung ist nicht besonders tiefsinnig, aber sie ist aufwändig, lang und nervig. Das ist der wahre Grund, warum sie hier nicht steht. Es ist sinnvoll sich diese Determinante als Blackbox zu merken oder irgendwo aufzuschreiben.

3 Mehrdimensionale Integration

Weiter ist $r = \sqrt{x^2 + y^2 + z^2}$, also können wir

$$(x^2 + y^2 + z^2)^{3/2} = \sqrt{x^2 + y^2 + z^2}^3 = r^3$$

schreiben. Damit gilt

$$\rho(x, y, z) = e^{(x^2+y^2+z^2)^{3/2}} = e^{r^3}.$$

Also ist die Masse hier gegeben durch

$$\begin{aligned} \int_K \rho(x, y, z) \, d(x, y, z) &= \int_L e^{r^3} r^2 \sin(\theta) \, d(r, \varphi, \theta) \\ &= \int_0^\pi \int_{-\pi}^\pi \int_0^1 r^2 e^{r^3} \sin(\theta) \, dr \, d\varphi \, d\theta \\ &= \int_0^\pi \sin(\theta) \, d\theta \int_{-\pi}^\pi 1 \, d\varphi \int_0^1 r^2 e^{r^3} \, dr. \end{aligned}$$

Wir beobachten $(e^{r^3})' = 3r^2 e^{r^3}$ wegen der Kettenregel, also ist $\frac{1}{3} \cdot e^{r^3}$ eine Stammfunktion von $r^2 e^{r^3}$. Es folgt

$$\begin{aligned} &= -\cos(\theta) \Big|_{\theta=0}^{\theta=\pi} \cdot 2\pi \cdot \frac{1}{3} e^{r^3} \Big|_{r=0}^{r=1} \\ &= (1 + 1) \cdot 2\pi \cdot \frac{1}{3} (e - 1) = \frac{4\pi}{3} (e - 1). \end{aligned}$$

(c) Das Volumen der Kugel mit Radius R

$$K_R = \{(x, y, z)^T \in \mathbb{R}^3 : \sqrt{x^2 + y^2 + z^2} \leq R\}$$

lässt sich ebenfalls mit Kugelkoordinaten berechnen. Hier ist die transformierte Version $L_R = [0, R] \times [-\pi, \pi] \times [0, \pi]$ und es gilt

$$\begin{aligned} V &= \int_{K_R} 1 \, d(x, y, z) = \int_{L_R} r^2 \sin(\theta) \, d(r, \varphi, \theta) \\ &= \int_0^\pi \int_{-\pi}^\pi \int_0^R r^2 \sin(\theta) \, dr \, d\varphi \, d\theta \\ &= \int_0^R r^2 \, dr \int_{-\pi}^\pi 1 \, d\varphi \int_0^\pi \sin(\theta) \, d\theta \\ &= \frac{1}{3} r^3 \Big|_{r=0}^{r=R} \cdot 2\pi \cdot (-\cos(\theta)) \Big|_{\theta=0}^{\theta=\pi} = \frac{2\pi}{3} R^3 (1 + 1) = \frac{4\pi}{3} R^3. \end{aligned}$$

Das ist genau die bekannte Formel für das Volumen einer Kugel.

3.3 Kurvenintegrale

Im diesem Abschnitt untersuchen wir die zwei folgenden Fragestellungen:

- (i) Gegeben sei eine Kurve, z. B. ein Kreis, wie können wir ihre Länge bestimmen?
- (ii) Liegt diese Kurve in einem Kraftfeld, z. B. dem Gravitationsfeld eines Planeten, welche Arbeit wird beim Transport einer Masse entlang des Weges verrichtet?

Definition 3.10. Sei $[a, b] \subseteq \mathbb{R}$ ein Intervall. Ein Weg ist eine stetig differenzierbare Funktion $\gamma : [a, b] \rightarrow \mathbb{R}^m$. Die Bildmenge von γ heißt Kurve.

Beispiel 3.11. Der Einheitskreis wird durch den Weg

$$\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

beschrieben.

Beim Integrieren einer Funktion entlang eines Weges ist zu unterscheiden, ob die Funktion ein Skalarfeld oder ein Vektorfeld ist. Wir beginnen mit dem Fall von Skalarfeldern.

Definition 3.12 (Wegintegral für Skalarfelder). Sei $\gamma : [a, b] \rightarrow \mathbb{R}^m$ ein Weg und $f : \mathbb{R}^m \rightarrow \mathbb{R}$ eine stetige Funktion. Das Wegintegral von f entlang γ ist

$$\int_{\gamma} f(x) \, ds := \int_a^b f(\gamma(t)) \|\gamma'(t)\| \, dt.$$

Das Wegintegral wird auch Kurvenintegral genannt.

Beispiel 3.13. Wir berechnen das Wegintegral der konstanten Funktion $f(x) = 1$ entlang des Weges $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$ mit

$$\gamma(t) = R \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}.$$

Dieser Weg beschreibt einen Kreis mit Radius $R > 0$ um den Ursprung, vgl. mit Beispiel 3.11. Dann gilt

$$\gamma'(t) = R \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \quad \text{sowie} \quad \|\gamma'(t)\| = R\sqrt{(-\sin(t))^2 + \cos(t)^2} = R.$$

Also ist

$$\int_{\gamma} 1 \, ds = \int_0^{2\pi} 1 \|\gamma'(t)\| \, dt = \int_0^{2\pi} R \, dt = 2\pi R.$$

Es handelt sich um den Umfang des von γ beschriebenen Kreises mit Radius R .

3 Mehrdimensionale Integration

Dieses Ergebnis ist kein Zufall, sondern ein Spezialfall des folgenden allgemeinen Zusammenhangs.

Satz 3.14 (Weg-/Kurvenlänge). *Die Länge eines Weges $\gamma : [a, b] \rightarrow \mathbb{R}^m$ ist gegeben durch*

$$\int_{\gamma} 1 \, ds = \int_a^b \|\gamma'(t)\| \, dt.$$

Beispiel 3.15. Ist $f : [a, b] \rightarrow \mathbb{R}$ eine differenzierbare Funktion, so können wir daraus eine Formel für die Länge des zugehörigen Funktionsgraphen gewinnen. Dazu parametrisieren wir die Kurve des Graphen durch

$$\gamma : [a, b] \rightarrow \mathbb{R}^2, \quad \text{mit} \quad \gamma(t) = \begin{pmatrix} t \\ f(t) \end{pmatrix}$$

und berechnen die Weglänge durch

$$\int_{\gamma} 1 \, ds = \int_a^b \|\gamma'(t)\| \, dt = \int_a^b \sqrt{1 + (f'(t))^2} \, dt.$$

Für die Funktion $f(t) = t^{3/2}$ auf dem Intervall $[0, 1]$ ergibt das mit $f'(t) = 3/2 \cdot \sqrt{t}$ beispielsweise

$$\begin{aligned} \int_0^1 \sqrt{1 + \frac{9}{4}t} \, dt &= \int_1^{13/4} \sqrt{u} \cdot \frac{4}{9} \, du = \frac{4}{9} \cdot \frac{2}{3} u^{3/2} \Big|_{u=1}^{u=13/4} = \frac{8}{27} \left[\left(\frac{13}{4}\right)^{3/2} - 1 \right] \\ &= \frac{8}{27} \left[\frac{13\sqrt{13}}{8} - 1 \right] = \frac{13\sqrt{13} - 8}{27}. \end{aligned}$$

wobei die Substitution $u = 1 + 9/4 \cdot t$ verwendet wurde.

Wir betrachten noch ein weiteres Beispiel eines Kurvenintegrals über ein anderes Skalarfeld als die konstante-Eins-Funktion.

Beispiel 3.16. Gegeben sei der Weg

$$\gamma : [0, 1] \rightarrow \mathbb{R}, \quad \gamma(t) = \begin{pmatrix} t \\ t^2 \end{pmatrix}$$

und die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $f(x, y) = 6x + 2\sqrt{y}$. Dann gilt $\gamma'(t) = (1, 2t)^T$ und mit der Substitution $u = 1 + 4t^2$ ergibt sich

$$\begin{aligned} \int_{\gamma} f(x, y) \, ds &= \int_0^1 f(\gamma(t)) \|\gamma'(t)\| \, dt = \int_0^1 f(t, t^2) \sqrt{1 + 4t^2} \, dt \\ &= \int_0^1 8t \sqrt{1 + 4t^2} \, dt = \int_1^5 \sqrt{u} \, du = \frac{2}{3} u^{3/2} \Big|_{u=1}^{u=5} = \frac{2}{3} (5\sqrt{5} - 1). \end{aligned}$$

Nun schauen wir uns Kurvenintegrale über Vektorfelder an.

Definition 3.17. Sei $\gamma : [a, b] \rightarrow \mathbb{R}^m$ ein Weg und $F : \mathbb{R}^m \rightarrow \mathbb{R}^m$ ein stetiges Vektorfeld. Dann setzt man

$$\int_{\gamma} F(x) dx := \int_a^b \langle F(\gamma(t)), \gamma'(t) \rangle dt = \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt$$

mit dem Skalarprodukt $\langle \cdot, \cdot \rangle$ bzw. \cdot . Dieses Integral heißt Wegintegral oder Kurvenintegral des Vektorfelds F entlang von γ .

Bemerkung 3.18. Interpretiert man F als ein Kraftfeld, so beschreibt obiges Wegintegral die Arbeit, die vom Feld verrichtet wird, wenn eine Einheitsmasse entlang γ von $\gamma(a)$ nach $\gamma(b)$ transportiert wird. Die Ableitung $\gamma'(t)$ entspricht der *Geschwindigkeit* des Weges γ . Die Arbeit ist betragsmäßig am größten, wenn das Kraftfeld parallel zur Geschwindigkeit des Weges wirkt.

Beispiel 3.19. (a) Sei $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gegeben durch $F(x, y) = (y, -x)^T$. Wir berechnen das Kurvenintegral $\int_{\gamma} F(x) dx$ für den Weg

$$\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2, \quad \gamma(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix},$$

der einmal den Einheitskreis umläuft. Es gilt

$$\begin{aligned} \int_{\gamma} F(x) dx &= \int_0^{2\pi} \langle F(\gamma(t)), \gamma'(t) \rangle dt = \int_0^{2\pi} \left\langle \begin{pmatrix} \sin(t) \\ -\cos(t) \end{pmatrix}, \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \right\rangle dt \\ &= \int_0^{2\pi} (-\sin^2(t) - \cos^2(t)) dt = -\int_0^{2\pi} 1 dt = -2\pi. \end{aligned}$$

(b) Sei $M > 0$ die Masse eines Planeten im Koordinatenursprung und $m > 0$ die Masse eines Punktes bei $x \in \mathbb{R}^3 \setminus \{0\}$. Dann wirkt das Gravitationsfeld von M auf m durch das Kraftfeld

$$F(x) = -GMm \frac{x}{\|x\|^3},$$

wobei G die Gravitationskonstante bezeichnet.

Ist $\gamma : [0, 1] \rightarrow \mathbb{R}^3 \setminus \{0\}$ ein Weg, so ist

$$\int_{\gamma} F(x) dx = -GMm \int_0^1 \|\gamma(t)\|^{-3} \langle \gamma(t), \gamma'(t) \rangle dt$$

die Arbeit, die das Gravitationsfeld an m verrichtet, wenn es den Weg γ entlang transportiert wird.

3 Mehrdimensionale Integration

Definition 3.20 (Stammfunktion). Sei $D \subseteq \mathbb{R}^m$ offen und $F : D \rightarrow \mathbb{R}^m$ ein stetiges Vektorfeld. Eine stetig differenzierbare Funktion $\Phi : D \rightarrow \mathbb{R}$ heißt Stammfunktion oder Potenzial von F , falls $\nabla\Phi = F$.

Stammfunktionen sind in höheren Dimensionen leider ziemlich selten, aber unglaublich wertvoll, denn mit ihrer Hilfe lassen sich Kurvenintegrale über Vektorfelder sehr leicht auswerten. Ist nämlich Φ eine Stammfunktion von F und $\gamma : [a, b] \rightarrow D$ ein Weg, so gilt nach der Kettenregel „rückwärts“

$$\int_{\gamma} F(x) \, dx = \int_a^b F(\gamma(t)) \cdot \gamma'(t) \, dt = \int_a^b \nabla\Phi(\gamma(t)) \cdot \gamma'(t) \, dt = \int_a^b \frac{d}{dt}(\Phi(\gamma(t))) \, dt$$

Nun können wir den Hauptsatz der Differential- und Integralrechnung anwenden und finden

$$= \Phi(\gamma(b)) - \Phi(\gamma(a)).$$

Dieses Ergebnis ist in mehrfacher Hinsicht bemerkenswert. Zum Einen ist das Integral verschwunden und man bekommt seinen Wert schlicht daraus das Potenzial an zwei Punkten auszuwerten. Zum Anderen bedeutet das, dass der Wert des Kurvenintegrals nur vom Anfangs- und Endpunkt der Kurve abhängt und nicht vom Verlauf des Weges dazwischen. Ist das Vektorfeld F also durch ein Potenzial gegeben, so ist die Arbeit, die beim Durchqueren des zugehörigen Kraftfeldes entsteht, bzw. benötigt wird, nur davon abhängig wo man startet und wo man endet und nicht vom Weg dazwischen. Man nennt ein solches Integral *wegunabhängig* und in der Physik heißt ein solches Kraftfeld *konservativ*. Wir halten dieses Ergebnis fest.

Satz 3.21. Sei $D \subseteq \mathbb{R}^m$ offen. Besitzt ein Vektorfeld $F : D \rightarrow \mathbb{R}^m$ eine Stammfunktion $\Phi : D \rightarrow \mathbb{R}$, so gilt für jeden Weg $\gamma : [a, b] \rightarrow D$

$$\int_{\gamma} F(x) \, dx = \Phi(\gamma(b)) - \Phi(\gamma(a)).$$

Das Wegintegral hängt in diesem Fall also nur vom Anfangs- und Endpunkt des Weges ab. Ist der Weg geschlossen (Anfangspunkt=Endpunkt), so ist das Wegintegral 0.

Beispiel 3.22. (a) Das Gravitationsfeld

$$F(x) = -GMm \frac{x}{\|x\|^3}$$

aus Beispiel 3.19 (b) besitzt ein Potenzial, denn für die Funktion

$$\Phi(x) = \frac{GMm}{\|x\|} = \frac{GMm}{\sqrt{x_1^2 + x_2^2 + x_3^2}} = GMm(x_1^2 + x_2^2 + x_3^2)^{-1/2}$$

gilt

$$\begin{aligned}\nabla\Phi(x) &= GMm\left(-\frac{1}{2}(x_1^2+x_2^2+x_3^2)^{-3/2}\cdot 2x_1, -\frac{1}{2}(x_1^2+x_2^2+x_3^2)^{-3/2}\cdot 2x_2, \right. \\ &\quad \left. -\frac{1}{2}(x_1^2+x_2^2+x_3^2)^{-3/2}\cdot 2x_3\right) \\ &= -GMm\frac{1}{(\sqrt{x_1^2+x_2^2+x_3^2})^3}(x_1, x_2, x_3) = -GMm\frac{x}{\|x\|^3} = F(x).\end{aligned}$$

Φ ist das Gravitations- oder Coulomb-Potenzial, durch das wir uns alltaglich bewegen und das Gravitations-Kraftfeld ist konservativ. Wenn Sie auf einen Berg steigen, verrichten Sie eine Arbeit, die nur von Ihrer Masse und dem Hohenunterschied abhangt, auf welchem Weg Sie ans Ziel gelangen, ist dabei vollig irrelevant. (Auch wenn sich das meist nicht so anfuhlt. . .).

- (b) Nachdem im vorigen Beispiel das Potenzial einfach angegeben war und nur nachgerechnet wurde, dass es eines ist, schauen wir uns noch in einem Beispiel an, wie eine solche Stammfunktion wirklich berechnet werden kann. Dazu betrachten wir das Vektorfeld $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$f(x, y) = \begin{pmatrix} \sin(y) + 2xy \\ x \cos(y) + x^2 + 1 \end{pmatrix}.$$

Fur eine Stammfunktion $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ von f muss gelten $\partial_x\Phi(x, y) = f_1(x, y) = \sin(y) + xy$. Integrieren wir nach x , liefert das

$$\Phi(x, y) = \int (\sin(y) + 2xy) dx = x \sin(y) + x^2y + c(y)$$

mit einer Integrationskonstante, die noch von y abhangen kann. Leiten wir diesen Ansatz wiederum nach y ab, so muss die zweite Komponente f_2 von f herauskommen. Es muss also gelten

$$\partial_y\Phi(x, y) = x \cos(y) + x^2 + c'(y) = x \cos(y) + x^2 + 1.$$

Damit ist $c'(y) = 1$, wir konnen also beispielsweise $c(y) = y$ nehmen. Das ergibt die Stammfunktion

$$\Phi(x, y) = x \sin(y) + x^2y + y.$$

4 Gewöhnliche Differentialgleichungen

4.1 Erste Beispiele und Begriffe

Beispiel 4.1 (Natürliches Wachstum). Eine Population bestehe zur Zeit t aus $N(t)$ Individuen und wir nehmen an, dass die Population eine über die Zeit konstante Geburtenrate und Sterberate hat. Das heißt es gibt die zwei Konstanten

β = Anzahl Geburten pro Individuum und Zeiteinheit

δ = Anzahl Todesfälle pro Individuum und Zeiteinheit.

Damit ist in einem Zeitintervall $[t, t + \Delta t]$

die Anzahl der Geburten = $\beta N(t) \Delta t$,

die Anzahl der Todesfälle = $\delta N(t) \Delta t$.

Für die Größe der Population zur Zeit $t + \Delta t$ folgt

$$N(t + \Delta t) = N(t) + \beta N(t) \Delta t - \delta N(t) \Delta t,$$

bzw. nach umstellen

$$\frac{N(t + \Delta t) - N(t)}{\Delta t} = (\beta - \delta) N(t)$$

Diese Gleichung ist nur näherungsweise gültig, weil bei der Berechnung der Anzahl von Geburten bzw. Todesfällen im Zeitintervall $[t, t + \Delta t]$ für die Anzahl N der Individuen der konstante Wert $N(t)$ verwendet wurde. Tatsächlich wird N im Intervall $[t, t + \Delta t]$ variieren. Die Näherung wird umso genauer, je kleiner Δt ist. Wir lassen deshalb Δt gegen Null gehen und erhalten

$$N'(t) = \lim_{\Delta t \rightarrow 0} \frac{N(t + \Delta t) - N(t)}{\Delta t} = (\beta - \delta) N(t).$$

Um den Grenzprozess $\Delta t \rightarrow 0$ durchführen zu können, muss die Funktion $N(t)$ differenzierbar, also insbesondere stetig sein. Dies erscheint zunächst unrealistisch, da $N(t)$ nur diskrete Werte aus \mathbb{N}_0 annimmt. Dennoch ist diese Idealisierung sinnvoll, wenn die Population aus sehr vielen Individuen besteht. In diesem Fall entspricht eine Änderung der Anzahl um Eins einer sehr kleinen relativen Änderung von $N(t)$. Betrachtet

4 Gewöhnliche Differentialgleichungen

man z.B. die Population einer gewissen chemischen Spezies (Molekülsorte), so wird N typischerweise im Bereich 10^{20} bis 10^{24} liegen.

Für große Populationen kann die Anzahl der Individuen also mit kleinem (relativem) Fehler durch eine stetige Funktion $N(t)$ beschrieben werden. Unter der zusätzlichen Annahme, dass $N(t)$ auch differenzierbar ist, ergibt sich das *mathematische Modell*

$$N'(t) = (\beta - \delta) N(t) \quad (4.1)$$

zur Beschreibung des Wachstums dieser Population.

Gleichungen dieser Form, deren unbekannte Größe eine Funktion ist, die zusammen mit ihren Ableitungen auftritt, heißen *Differentialgleichungen* (kurz DGL). In diesem einfachen Beispiel ist das nur die Funktion und die erste Ableitung, im Allgemeinen dürfen natürlich auch höhere Ableitungen der gesuchten Funktion auftreten. Den Grad der höchsten auftretenden Ableitung nennt man die *Ordnung* der Differentialgleichung. Die Differentialgleichung in (4.1) ist also eine von erster Ordnung. Ein Beispiel einer Differentialgleichung von 3. Ordnung wäre etwa

$$y'''(t) - 2y'(t) + y(t)^4 = e^t.$$

Die naturwissenschaftliche Beschreibung der Natur ist voll von Differentialgleichungen, denn wann immer ein Prozess oder eine Größe beschrieben werden soll, bei der die Änderungsrate der Größe, also die Ableitung, vom Zustand der Größe selbst abhängt, entsteht ganz natürlich eine Differentialgleichung. Die kurze Einführung in das Thema hier soll aufzeigen, wie typischerweise Lösungen von Differentialgleichungen aussehen und wie diese in einigen wichtigen Spezialfällen auch berechnet werden können.

Bemerkung 4.2 (Diskussion der natürlichen Wachstumsgleichung). Die *natürliche Wachstumsgleichung* aus (4.1) ist von der Form

$$y'(t) = k y(t) \quad (4.2)$$

mit $k := \beta - \delta \in \mathbb{R}$ konstant. Man rechnet leicht nach, dass jede Funktion der Form $y(t) = Ce^{kt}$ eine Lösung ist, wobei $C \in \mathbb{R}$ eine beliebige Konstante ist und die folgende Überlegung zeigt, dass das auch alle Lösungen sind: Für jede Lösung u von (4.2) ist

$$\frac{d}{dt}(u(t)e^{-kt}) = u'(t)e^{-kt} - ku(t)e^{-kt} = ku(t)e^{-kt} - ku(t)e^{-kt} = 0.$$

Also ist die Funktion $u(t)e^{-kt}$ konstant, und es gibt ein $C \in \mathbb{R}$ mit $u(t)e^{-kt} = C$. Das bedeutet aber gerade, dass $u(t) = Ce^{kt}$ ist.

Die Lösung der natürlichen Wachstumsgleichung ist also nur bis auf eine Konstante bestimmt und das ist auch ein ganz typisches Verhalten von Differentialgleichungen. Man spricht deshalb von der *allgemeinen Lösung* der Differentialgleichung. Denkt man darüber nach, ist es auch nicht verblüffend, denn nur durch die Wachstumsbedingung

4.1 Erste Beispiele und Begriffe

(„Der Zuwachs der Population ist proportional zu ihrer Größe“) ist das System nicht eindeutig festgelegt. Damit festliegt wie groß die Population zu einem Zeitpunkt ist, muss man auch noch wissen, wie groß sie zum Beginn der Beobachtung war. Tatsächlich ist genau das die Bedeutung der obigen Konstante C , denn nach der Rechnung

$$u(0) = C e^{k \cdot 0} = C \cdot 1 = C$$

gibt der Wert C genau die Population zum Zeitpunkt $t = 0$ an.

Schreibt man also zusätzlich den Wert der Lösung zu einem bestimmten (Zeit-)Punkt t_0 durch Vorgabe eines *Anfangswertes* y_0 vor, so ist die Lösung eindeutig bestimmt und $y(t) = y_0 e^{k(t-t_0)}$ ist die einzige Lösung des sogenannten *Anfangswertproblems* (kurz AWP)

$$\begin{cases} y'(t) = k y(t), \\ y(t_0) = y_0. \end{cases}$$

Hier sind ein paar weitere wichtige Beispiele von elementaren Differentialgleichungen.

Beispiel 4.3. (a) Eine Kräftebilanzrechnung am *Federpendel* (wobei x die Position, t die Zeit bezeichne), vgl. Abbildung 4.1, liefert:

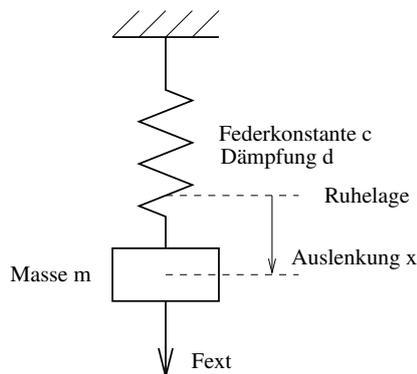


Abbildung 4.1: Das Federpendel aus Beispiel 4.3 (a)

$$\begin{array}{ll} \text{mit zur Geschw. prop. Dämpfung} & m x''(t) = -c x(t) \\ \text{mit zusätzl. äußerer Kraft} & m x''(t) = -c x(t) - d x'(t) \\ & m x''(t) = -c x(t) - d x'(t) + F_{ext}(t) \end{array}$$

Man erhält hier damit eine Differentialgleichung vom Typ

$$y''(t) + a y'(t) + b y(t) = f(t),$$

wobei $a, b \in \mathbb{R}$ und die Funktion f gegeben ist, während y die gesuchte Funktion bezeichnet. Diese Differentialgleichung ist von 2. Ordnung und *linear*, d. h. y

4 Gewöhnliche Differentialgleichungen

und seine Ableitungen, hier y' und y'' , treten ausschließlich linear auf (also keine Terme wie $y(t)^2$, $y(t)y'(t)$, $\sin(y(t))$, etc.).

Analog zu linearen Gleichungssystemen nennt man eine solche lineare Differentialgleichung *homogen*, falls $f = 0$ gilt, und sonst *inhomogen*.

- (b) Für das *Fadenpendel* in Abbildung 4.2 bezeichnen wir mit ϕ die Auslenkung und t die Zeit.

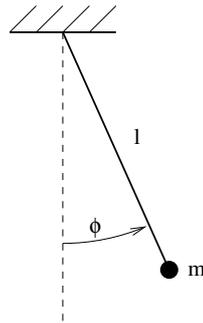


Abbildung 4.2: Das Fadenpendel aus Beispiel 4.3 (b)

Hier liefert eine physikalische Betrachtung der Kräfte die Differentialgleichung

$$\phi''(t) + \frac{g}{l} \sin(\phi(t)) = 0,$$

wobei l die Länge des Pendels und g die Gravitationskraft ist. Dies ist eine *nichtlineare* Differentialgleichung 2. Ordnung. Man kann zeigen, dass die zugehörigen Anfangswertprobleme (mit Vorgaben $\phi(0) = \phi_0$, $\phi'(0) = \phi_1$) eine eindeutige Lösung besitzen – eine formelmäßige Berechnung der Lösung ist hier aber i. A. nicht möglich.

4.2 Differentialgleichungen mit getrennten Variablen

In diesem Abschnitt betrachten wir eine erste spezielle Klasse von Differentialgleichungen erster Ordnung, für die es ein allgemeines Lösungsverfahren gibt. Dazu muss sich die Gleichung in einer speziellen Form schreiben lassen, der wir zunächst ihren Namen geben.

Definition 4.4. Eine Differentialgleichung erster Ordnung der Form

$$y'(t) = g(t) \cdot h(y(t)) \quad (\text{kurz } y' = g(t)h(y))$$

mit zwei Funktionen g und h heißt Differentialgleichung mit getrennten Variablen.

4.2 Differentialgleichungen mit getrennten Variablen

Man lässt in der Schreibung auch oft die t -Abhängigkeit von y weg und schreibt dann kurz $y' = g(t)h(y)$. Das sollte man aber nur machen, wenn man sicher ist, dass das nicht später in der Rechnung zu Verwirrungen führt.

Beispiel 4.5. (a) Sei $k \in \mathbb{R}$. Die Differentialgleichung

$$y'(t) = -ky(t)^2$$

ist von getrennten Variablen mit konstantem $g(t) = -k$ und $h(y) = y^2$.

(b) Die Differentialgleichung

$$y'(t) = e^{y(t)} \sin(t)$$

ist ebenfalls von getrennten Variablen mit $g(t) = \sin(t)$ und $h(y) = e^y$.

(c) Auch die Differentialgleichung

$$y'(t) = e^{y(t)+\sin(t)}$$

ist von getrennten Variablen, selbst wenn man es ihr nicht sofort ansieht. Wegen $e^{y(t)+\sin(t)} = e^{y(t)} \cdot e^{\sin(t)}$ lässt sie sich nämlich in eine Form mit getrennten Variablen umschreiben. Es ist hier $g(t) = e^{\sin(t)}$ und $h(y) = e^y$.

(d) Schließlich ist $y'(t) = y(t)^2 + t^2$ nicht von getrennten Variablen.

Zur Lösung solcher Gleichungen kann man folgende Rechnung machen: Man trennt die Variablen auf die rechte und linke Seite der Gleichung:

$$\frac{y'}{h(y)} = g(t) \quad (\text{falls } h(y) \neq 0)$$

und integriert dann über die Variable t . Eine Anwendung der Substitutionsregel später findet man:

$$\int g(t) dt = \int \frac{y'(t)}{h(y(t))} dt = \int \frac{dy}{h(y)}.$$

Eine Berechnung dieser unbestimmten Integrale (mit Integrationskonstante!) und Auflösen nach y ergibt alle Lösungen $y(t)$, für die $h(y(t)) \neq 0$ gilt. Der Spezialfall $h(y(t)) = 0$ muss gesondert behandelt werden.

Beispiel 4.6. (a) Wir beobachten zunächst, dass unsere Gleichung des natürlichen Wachstums $y'(t) = k y(t)$ aus (4.1) mit $g(t) = k$ und $h(y) = y$ von getrennten Variablen ist und berechnen noch mal nach dem oben beschriebenen Verfahren die Lösung. Für $y(t) \neq 0$ ist

$$\int \frac{dy}{y} = \int k dt, \quad \text{also} \quad \ln(|y|) = kt + c$$

mit einer Integrationskonstanten c (Eigentlich entsteht natürlich sowohl beim Integral über y als auch beim Integral über t eine Konstante, aber die beiden lassen

4 Gewöhnliche Differentialgleichungen

sich immer zu einer zusammenfassen.) Nun müssen wir nach y auflösen. Dazu wenden wir die Exponentialfunktion an und machen dann eine Fallunterscheidung für den Betrag. Das liefert

$$|y| = e^{kt+c} = e^c e^{kt}, \quad \text{also} \quad y(t) = e^c e^{kt} \text{ oder } y(t) = -e^c e^{kt}.$$

Dann lassen sich beide Fälle zusammenfassen zu $y(t) = \tilde{c}e^{kt}$ mit $\tilde{c} \neq 0$. Schließlich ist auch die konstante Nullfunktion $y(t) = 0$ eine Lösung, die dem Fall $c = 0$ entspricht. Damit haben alle Lösungen die Form $y(t) = Ce^{kt}$ mit $C \in \mathbb{R}$ beliebig.

(b) Für $k \in \mathbb{R}$ gegeben, betrachten wir das Anfangswertproblem

$$\begin{cases} y'(t) = -k y(t)^2 \\ y(0) = y_0 \end{cases}$$

mit einem Anfangswert $y_0 > 0$. Für $y(t) \neq 0$ trennen wir die Variablen und integrieren. Das liefert

$$\int \frac{dy}{y^2} = - \int k dt, \quad \text{also} \quad -\frac{1}{y} = -kt + c$$

mit einem $c \in \mathbb{R}$. Auflösen nach y liefert

$$\frac{1}{y(t)} = kt - c, \quad \text{d.h.} \quad y(t) = \frac{1}{kt - c}.$$

Schließlich können wir noch die Konstante c anhand der Anfangsbedingung bestimmen. Es muss gelten

$$y_0 = y(0) = \frac{1}{-c}, \quad \text{also ist} \quad c = -\frac{1}{y_0}$$

Zusammen ist

$$y(t) = \frac{1}{kt + \frac{1}{y_0}} = \frac{y_0}{y_0 kt + 1}$$

die Lösung des AWP.

Ist nicht nur eine Differentialgleichung mit getrennten Veränderlichen, sondern auch ein Anfangswert gegeben, hat man also ein Anfangswertproblem

$$\begin{cases} y'(t) = g(t)h(y(t)) \\ y(t_0) = y_0 \end{cases}$$

zu lösen, so kann bei obiger Methodik der Anfangswert sofort mit eingerechnet werden. Dazu integriert man bestimmt von t_0 bis t :

$$\int_{y(t_0)}^{y(t)} \frac{dx}{h(x)} = \int_{y_0}^{y(t)} \frac{dx}{h(x)} = \int_{t_0}^t g(s) ds \quad (4.3)$$

Auch hier liefert das Ausrechnen der Integrale mit nachfolgendem Auflösen nach $y(t)$ die Lösung.

4.3 Lineare Differentialgleichungen 1. Ordnung

Beispiel 4.7. Wir lösen das Anfangswertproblem

$$\begin{cases} y'(t) = t \cdot y(t) \\ y(0) = 1. \end{cases}$$

Dann ist $g(t) = t$ und $h(y) = y$. Wir rechnen also entsprechend (4.3) für $y(t) > 0$

$$\begin{aligned} \int_1^{y(t)} \frac{dx}{x} &= \int_0^t s \, ds \quad \Rightarrow \quad \ln(x) \Big|_{x=1}^{y(t)} = \frac{1}{2} s^2 \Big|_{s=0}^t \quad \Rightarrow \quad \ln(y(t)) - \ln(1) = \frac{1}{2} t^2 \\ &\Rightarrow \quad y(t) = e^{t^2/2}. \end{aligned}$$

Das dass tatsächlich eine Lösung ist, verifiziert man leicht durch eine Probe: Tatsächlich ist $y(0) = e^0 = 1$ und

$$y'(t) = e^{t^2/2} \frac{1}{2} \cdot 2t = t e^{t^2/2} = t \cdot y(t).$$

4.3 Lineare Differentialgleichungen 1. Ordnung

Eine *lineare Differentialgleichung 1. Ordnung* hat die allgemeine Form

$$y'(t) + a(t)y(t) = f(t) \tag{4.4}$$

mit gegebenen Funktionen a und f und gesuchter Funktion y . Damit die Lösung eindeutig wird, muss auch eine solche Gleichung wieder mit einer Anfangsbedingung kombiniert werden.

Für lineare Differentialgleichungen gibt es eine bemerkenswerte Parallele zur Lösung von linearen Gleichungssystemen. Wir hatten dort gesehen, dass man alle Lösungen eines inhomogenen linearen Gleichungssystems $Ax = b$ bekommt, indem man *eine* Lösung des Systems, genannt spezielle Lösung oder Partikulärlösung, zu *allen* Lösungen des zugehörigen homogenen LGS $Ax = 0$ addiert.

Dasselbe passiert hier auch. Die *zugehörige homogene* Differentialgleichung ist gegeben durch

$$y'(t) + a(t)y(t) = 0. \tag{4.5}$$

Ist nun y_s eine Lösung von (4.4) und y_h eine Lösung der zugehörigen homogenen Gleichung (4.5), so ist deren Summe $y_s + y_h$ wegen

$$\begin{aligned} (y_s + y_h)'(t) + a(t)(y_s(t) + y_h(t)) &= y_s'(t) + a(t)y_s(t) + y_h'(t) + a(t)y_h(t) \\ &= f(t) + 0 = f(t) \end{aligned}$$

ebenfalls ein Lösung von (4.4) und man kann auch zeigen, dass alle Lösungen auf diese Weise erzeugt werden. Auch in diesem Zusammenhang nennt man daher y_s *spezielle Lösung* oder *Partikulärlösung*

4 Gewöhnliche Differentialgleichungen

Die Aufgabe zur Bestimmung der allgemeinen Lösung von (4.4) lässt sich also in zwei Teilschritte zerlegen. Wir brauchen erstens alle Lösungen der zugehörigen homogenen Gleichung (4.5) und zweitens irgendeine Lösung von (4.4).

Der erste Schritt ist im Wesentlichen schon erledigt, denn die homogene Differentialgleichung $y'(t) = -a(t)y(t)$ ist eine mit getrennten Variablen. Bezeichnen wir mit $A(t)$ eine Stammfunktion von $a(t)$, so erhalten wir mit dem Verfahren aus dem vorhergehenden Abschnitt die allgemeine Lösung der homogenen Gleichung als

$$y_h(t) = ce^{-A(t)} \quad \text{mit } c \in \mathbb{R} \text{ beliebig.} \quad (4.6)$$

Wir brauchen also nur noch eine Lösung von (4.4). Dazu verwenden wir den Ansatz der *Variation der Konstanten*:

$$y_s(t) = c(t)e^{-A(t)},$$

wir machen also aus der Konstanten c in der allgemeinen homogenen Lösung eine Funktion $c(t)$. Setzen wir diesen Ansatz in (4.4) ein, so erhalten wir

$$\begin{aligned} y'_s(t) + a(t)y_s(t) &= c'(t)e^{-A(t)} + c(t)e^{-A(t)}(-A'(t)) + a(t)c(t)e^{-A(t)} \\ &= c'(t)e^{-A(t)} - y_s(t)a(t) + a(t)y_s(t) = c'(t)e^{-A(t)} \stackrel{!}{=} f(t). \end{aligned}$$

Das liefert $c'(t) = e^{A(t)}f(t)$ und eine Integration später bekommen wir

$$c(t) = \int e^{A(t)}f(t) dt, \quad \text{also} \quad y_s(t) = e^{-A(t)} \int e^{A(r)}f(r) dr \Big|_{r=t}.$$

Dabei bedeutet die Notation mit em senkrechten Strich hinter der Formel, dass erst das unbestimmte Integral zu bestimmen ist, und dann im entstehenden Ausdruck die Variable r durch t zu ersetzen ist.

Methode 4.8 (Lineare DGL 1. Ordnung). Gesucht sind alle Lösungen der linearen Differentialgleichung erster Ordnung

$$y'(t) + a(t)y(t) = f(t)$$

oder die Lösung eines zugehörigen Anfangswertproblems

$$\begin{cases} y'(t) + a(t)y(t) = f(t), \\ y(t_0) = y_0. \end{cases}$$

Diese findet man mit den folgenden Schritten:

- (a) Bestimme eine Stammfunktion A von a . Ist die Lösung des Anfangswertproblems gesucht, vereinfacht sich später die Rechnung, wenn man hier schon vorausschauend die Stammfunktion nimmt, die $A(t_0) = 0$ erfüllt. In diesem Fall geht es direkt weiter im zweiten Satz von Schritt (e).

4.3 Lineare Differentialgleichungen 1. Ordnung

(b) Alle Lösungen der homogenen Gleichung sind dann schon gegeben als

$$y_h(t) = ce^{-A(t)} \quad \text{mit } c \in \mathbb{R} \text{ beliebig.}$$

(c) Bestimme eine spezielle Lösung durch

$$y_s(t) = e^{-A(t)} \int e^{A(r)} f(r) dr \Big|_{r=t}.$$

Die auftretende Integrationskonstante kann hier ignoriert, bzw. irgendwie gesetzt werden, es wird ja nur irgendeine spezielle Lösung benötigt.

(d) Alle Lösungen der Differentialgleichung sind dann gegeben durch die *Variation-der-Konstanten-Formel*:

$$y(t) = y_h(t) + y_s(t) = ce^{-A(t)} + e^{-A(t)} \int e^{A(r)} f(r) dr \Big|_{r=t} \quad \text{mit } c \in \mathbb{R} \text{ beliebig.}$$

(e) Ist auch ein Anfangswert gegeben, kann man durch Einsetzen von t_0 und Einstellen der Konstanten c , die Lösung des AWP erhalten. Hat man in ersten Schritt A so gewählt, dass $A(t_0) = 0$ ist, so liefert das direkt die Lösungsformel

$$y(t) = y_0 e^{-A(t)} + e^{-A(t)} \int_{t_0}^t e^{A(r)} f(r) dr.$$

Beispiel 4.9. (a) Wir betrachten das Anfangswertproblem

$$\begin{cases} y'(t) + y(t) = \sin(t), \\ y(0) = 1. \end{cases}$$

Hier ist also $a(t) = 1$, $f(t) = \sin(t)$, $t_0 = 0$ und $y_0 = 1$. Eine Stammfunktion A von a mit $A(0) = 0$ ist $A(t) = t$. Als Lösung ergibt sich also

$$y(t) = 1 \cdot e^{-A(t)} + e^{-A(t)} \int_0^t e^{A(r)} \sin(r) dr = e^{-t} + e^{-t} \int_0^t e^r \sin(r) dr.$$

Mit der Stammfunktion $\int e^r \sin(r) dr = 1/2 \cdot e^r (\sin(r) - \cos(r))$ folgt

$$y(t) = e^{-t} + e^{-t} \left(\frac{1}{2} e^r (\sin(r) - \cos(r)) \right) \Big|_{r=0}^t = \frac{3}{2} e^{-t} + \frac{1}{2} (\sin t - \cos t).$$

(b) Wir bestimmen alle Lösungen der linearen Differentialgleichung

$$y'(t) + \frac{1}{t} y(t) = t + 1$$

4 Gewöhnliche Differentialgleichungen

für $t > 0$. Hier ist also $a(t) = 1/t$ und für $t > 0$ ist eine Stammfunktion davon $A(t) = \ln(t)$. Das ergibt für die allgemeine Lösung des zugehörigen homogenen Systems

$$y_h(t) = ce^{-A(t)} = ce^{-\ln(t)} = c \cdot \frac{1}{t} \quad \text{mit } c \in \mathbb{R} \text{ beliebig.}$$

Für die spezielle Lösung setzen wir den Ansatz der Variation der Konstanten an: $y_s(t) = c(t)/t$. Dann soll gelten

$$t + 1 \stackrel{!}{=} y'_s(t) - \frac{1}{t}y_s(t) = \frac{c'(t)}{t} - \frac{c(t)}{t^2} + \frac{1}{t} \frac{c(t)}{t} = \frac{c'(t)}{t},$$

also ist

$$c'(t) = t^2 + t.$$

Eine mögliche Wahl von c ist damit $c(t) = t^3/3 + t^2/2$ und wir bekommen alle Lösungen als

$$y(t) = y_h(t) + \frac{c(t)}{t} = \frac{c}{t} + \frac{1}{3}t^2 + \frac{1}{2}t \quad \text{mit } c \in \mathbb{R} \text{ beliebig.}$$

4.4 Lineare Differentialgleichungen 2. Ordnung

In Anwendungen treten besonders häufig Differentialgleichungen von 2. Ordnung auf. Wir wollen uns auch davon die linearen Gleichungen anschauen. Lineare Differentialgleichungen 2. Ordnung haben die allgemeine Form

$$y''(t) + a(t)y'(t) + b(t)y(t) = f(t)$$

mit gegebenen Funktionen a , b und f und gesuchtem y . Auch hier gilt

allg. Lsg. der inhom. DGL = spez. Lsg. der inh. DGL + allg. Lsg. der hom. DGL.

Da die Gleichung 2. Ordnung ist, benötigt man nun zur Angabe der Menge aller Lösungen zwei Konstanten und dementsprechend zur eindeutigen Festlegung einer Lösung auch zwei Anfangsbedingungen, meistens in der Form $y(t_0) = y_0$ und $y'(t_0) = y_1$.

Allerdings gibt es kein allgemeines Verfahren zur Berechnung der Lösungen mehr. Ein solches lässt sich nur für den wichtigen Spezialfall *konstanter Koeffizienten* angeben, d. h. wenn $a(t) \equiv a$ und $b(t) \equiv b$ konstant sind. Daher werden wir uns auf diesen Fall beschränken. Es seien also im Folgenden $a, b \in \mathbb{R}$.

Auch hier löst man zunächst das zugehörige homogene Problem mit $f = 0$:

$$y''(t) + ay'(t) + by(t) = 0.$$

4.4 Lineare Differentialgleichungen 2. Ordnung

Damit bei der Summe auf der linken Seite Null herauskommen kann, müssen die Funktionen y'' , y' und y alle „vom selben Typ sein“. Das ist die Motivation den Ansatz $y(t) = e^{\lambda t}$ mit $\lambda \in \mathbb{C}$ zu versuchen. Setzt man diesen ein, so erhält man

$$\lambda^2 e^{\lambda t} + a\lambda e^{\lambda t} + b e^{\lambda t} = 0 \quad \Leftrightarrow \quad \boxed{\lambda^2 + a\lambda + b = 0} \quad (4.7)$$

Diese quadratische Gleichung hat die beiden Lösungen

$$\lambda_1 = -\frac{a}{2} + \sqrt{D}, \quad \lambda_2 = -\frac{a}{2} - \sqrt{D} \quad \text{mit} \quad D = \frac{a^2}{4} - b.$$

Das Vorzeichen der Diskriminante D entscheidet über den Typ der allgemeinen Lösung.

1. Fall $D > 0$: Dann sind beide Nullstellen λ_1 und λ_2 reell und verschieden. Man bekommt die allgemeine Lösung

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} \quad \text{mit } c_1, c_2 \in \mathbb{R} \text{ beliebig.}$$

2. Fall $D = 0$: Dann gibt es eine doppelte reelle Nullstelle $\lambda := \lambda_1 = \lambda_2$. Die allgemeine Lösung ist in diesem Fall

$$y(t) = c_1 e^{\lambda t} + c_2 t e^{\lambda t} \quad \text{mit } c_1, c_2 \in \mathbb{R} \text{ beliebig.}$$

3. Fall $D < 0$: Dann gibt es zwei konjugiert komplexe Nullstellen

$$\lambda_1 = -\frac{a}{2} + i\omega, \quad \lambda_2 = -\frac{a}{2} - i\omega \quad \text{mit } \omega = \sqrt{-D}.$$

Die entsprechenden komplexen Lösungen $e^{(-a/2+i\omega)t}$ und $e^{(-a/2-i\omega)t}$ der Differentialgleichung lassen sich mit der Eulerschen Formel umschreiben und man bekommt dann für die allgemeine Lösung heraus:

$$y(t) = e^{-a/2 \cdot t} (c_1 \cos(\omega t) + c_2 \sin(\omega t)) \quad \text{mit } c_1, c_2 \in \mathbb{R} \text{ beliebig.}$$

Bemerkung 4.10. Die Funktionen y_1 und y_2 spannen jeweils alle Lösungen der Differentialgleichung durch die Bildung $y(t) = c_1 y_1(t) + c_2 y_2(t)$ mit $c_1, c_2 \in \mathbb{R}$ auf. Sie heißen *Fundamentallösungen* der linearen Differentialgleichung 2. Ordnung.

Beispiel 4.11. Die Differentialgleichung zum in Beispiel 4.3 (a) beschriebenen ungedämpften Federpendel ist linear von 2. Ordnung mit konstanten Koeffizienten:

$$m y''(t) = -c y(t), \quad \text{also} \quad y''(t) + 0 \cdot y'(t) + \frac{c}{m} y(t) = 0.$$

4 Gewöhnliche Differentialgleichungen

Hier ist also $a = 0$ und $b = c/m$. Das Polynom aus (4.7) lautet entsprechend $\lambda^2 + \frac{c}{m} = 0$ mit den konjugiert komplexen Nullstellen $\pm i\sqrt{c/m}$. Wir sind also in Fall 3 und die allgemeine Lösung lautet

$$y(t) = e^{-a/2 \cdot t} (c_1 \cos(\omega t) + c_2 \sin(\omega t)) = c_1 \cos\left(\sqrt{\frac{c}{m}} t\right) + c_2 \sin\left(\sqrt{\frac{c}{m}} t\right).$$

Wie zu erwarten entsteht eine periodische Schwingung. Die Größe $\omega = \sqrt{c/m}$ ist deren Kreisfrequenz.

Methode 4.12. Für gegebene $a, b \in \mathbb{R}$ und $f : \mathbb{R} \rightarrow \mathbb{R}$ ist die allgemeine Lösung y der linearen Differentialgleichung 2. Ordnung

$$y''(t) + ay'(t) + by(t) = f(t)$$

gesucht. Dazu geht man folgendermaßen vor:

- (a) Bestimme die Nullstellen $\lambda_1 = -a/2 + \sqrt{D}$ und $\lambda_2 = -a/2 - \sqrt{D}$ mit $D = a^2/4 - b$ des Polynoms $\lambda^2 + a\lambda + b = 0$.
- (b) Je nach Vorzeichen von D ist die allgemeine Lösung der homogenen(!) Gleichung $y''(t) + ay'(t) + by(t) = 0$ dann gegeben durch

$$y(t) = c_1 y_1(t) + c_2 y_2(t) \text{ mit } c_1, c_2 \in \mathbb{R} \text{ beliebig.}$$

Dabei ist

$$\begin{array}{lll} y_1(t) = e^{\lambda_1 t}, & y_2(t) = e^{\lambda_2 t} & \text{für } D > 0 \\ y_1(t) = e^{\lambda_1 t}, & y_2(t) = t e^{\lambda_1 t} & \text{für } D = 0 \\ y_1(t) = e^{-\frac{a}{2} t} \cos \omega t, & y_2(t) = e^{-\frac{a}{2} t} \sin \omega t & \text{für } D < 0, \end{array}$$

mit der Abkürzung $\omega = \sqrt{-D}$.

- (c) Bestimme irgendwie eine spezielle Lösung der inhomogenen Gleichung, wenn es nicht anders geht mit Variation der Konstanten. d. h. dem Ansatz

$$y(t) = c_1(t)y_1(t) + c_2(t)y_2(t).$$

Eine oft einfachere Methode ist der sogenannte „Ansatz vom Typ der rechten Seite“: Versuche für

eine rechte Seite f der Form	den Ansatz $y_s(t) =$
$e^{at}(a_1 \cos kt + a_2 \sin kt)$	$e^{at}(b_1 \cos kt + b_2 \sin kt)$
$a_n t^n + \dots + a_1 t + a_0$	$b_n t^n + \dots + b_1 t + b_0$

Beispiel 4.13. Wir setzen Beispiel 4.11 fort und betrachten wieder das ungedämpfte Federpendel mit Kreisfrequenz $\omega = \sqrt{m/c}$, das mit einer periodischen externen Kraft mit einer u. U. anderen Frequenz ω_0 angeregt wird:

$$y''(t) + \frac{c}{m} y(t) = \sin(\omega_0 t). \quad (4.8)$$

4.4 Lineare Differentialgleichungen 2. Ordnung

Wir kennen schon alle Lösungen des zugehörigen homogenen Systems als

$$y_h(t) = c_1 \cos(\omega t) + c_2 \sin(\omega t)$$

und brauchen noch irgendeine spezielle Lösung, um alle Lösungen zu kennen.

Die Inhomogenität f ist hier gegeben durch $f(t) = \sin(\omega_0 t)$ und der Ansatz vom Typ der rechten Seite ist demnach

$$y_s(t) = b_1 \cos(\omega_0 t) + b_2 \sin(\omega_0 t).$$

Einsetzen in die Gleichung (4.8) liefert (man beachte $c/m = \omega^2$):

$$\begin{aligned} \sin(\omega_0 t) &\stackrel{!}{=} y_s''(t) + \omega^2 y_s(t) \\ &= -b_1 \omega_0^2 \cos(\omega_0 t) - b_2 \omega_0^2 \sin(\omega_0 t) + \omega^2 (b_1 \cos(\omega_0 t) + b_2 \sin(\omega_0 t)) \\ &= (\omega^2 - \omega_0^2) b_1 \cos(\omega_0 t) + (\omega^2 - \omega_0^2) b_2 \sin(\omega_0 t). \end{aligned}$$

Diese Gleichheit ist im Fall $\omega_0 \neq \omega$ erfüllt für

$$b_1 = 0 \quad \text{und} \quad b_2 = \frac{1}{\omega^2 - \omega_0^2}$$

und wir haben eine spezielle Lösung gefunden als

$$y_s(t) = \frac{1}{\omega^2 - \omega_0^2} \sin(\omega_0 t).$$

Man beachte, dass deren Amplitude $1/(\omega^2 - \omega_0^2)$ für ω_0 in der Nähe von ω sehr groß wird. Dieser Effekt wird auch gerne mit „Resonanzkatastrophe“ bezeichnet. Im Fall $\omega = \omega_0$ funktioniert dieser Ansatz nicht.

Schließlich können wir damit alle Lösungen von (4.8) hinschreiben:

$$y(t) = y_s(t) + y_h(t) = \frac{1}{\omega^2 - \omega_0^2} \sin(\omega_0 t) + c_1 \cos(\omega t) + c_2 \sin(\omega t) \quad \text{mit } c_1, c_2 \in \mathbb{R}.$$

Index

- abgeschlossene Menge, 32
- Ableitung
 - Gradient, 38
 - Hesse-Matrix, 47
 - Jacobi-Matrix, 40
 - partielle, 35
 - Rechenregeln, 43
- Algorithmus, Gauß-, 6
- allgemeine Lösung einer DGL, 78
- Anfangswert einer DGL, 79
- Anfangswertproblem, 79

- beschränkte Folge, 28
- beschränkte Menge, 32

- charakteristisches Polynom, 14

- Determinante, 7
 - Entwicklungsformel
 - für Spalten, 9
 - für Zeilen, 8
 - Sarrus-Regel, 7
- diagonalisierbare Matrix, 17
- Diagonalisierung einer Matrix, 18
- Diagonalmatrix, 17
- Differentialgleichung, 78
 - allgemeine Lösung, 78
 - Anfangswert, 79
 - Anfangswertproblem, 79
 - homogen lineare, 80
 - inhomogen lineare, 80
 - lineare, 79
 - 1. Ordnung, 83
 - 2. Ordnung mit konstanten Koeffizienten, 86
 - Partikulärlösung, 83
 - spezielle Lösung, 83
 - zugehörige homogene DGL, 83
 - mit getrennten Variablen, 80
 - nichtlineare, 80
 - Ordnung einer, 78
- Divergenz eines Vektorfelds, 42
- Dreiecksmatrix, 12

- Eigenraum, 14
- Eigenvektor, 13
- Eigenwert, 13
- Einheitsmatrix, 7
- Elementarumformungen, 5
- Entwicklungsformel
 - für Spalten, 9
 - für Zeilen, 8
- erweiterte Koeffizientenmatrix, 2
- Extremstellen unter Nebenbedingung, 55
- Extremum, lokales/globales, 45

- Fadenpendel, 80
- Federpendel, 79
- Feld
 - Skalar-, 26
 - Vektor-, 25, 26
- Fibonacci-Zahlen, 22
- Folge
 - beschränkte, 28
 - Fibonacci-, 22
 - Grenzwert, 28
 - im \mathbb{R}^m , 27

Index

- konvergente, 28
- Koordinaten-, 27
- Fubini, Satz von, 61
- Fundamentallösungen, 87
- Funktion
 - Gradient einer, 38
 - Koordinaten-, 27
 - partiell differenzierbare
 - reellwertig, 35
 - vektorwertig, 39
 - Stamm-, 74
 - stetig in einem Punkt, 28
 - stetig partiell differenzierbare
 - skalarwertig, 36
 - vektorwertig, 40
 - stetige, 28
- Funktionsgrenzwert, 31
- Gauß-Algorithmus, 6
- Geschwindigkeit eines Weges, 73
- getrennte Variablen, DGL mit, 80
- globales Extremum, 45
- globales Minimum/Maximum, 45
- Gradient, 38
- Grenzwert
 - einer Folge, 28
 - Funktions-, 31
- Hülle, lineare, 16
- Hesse-Matrix eines Skalarfelds, 47
- homogene DGL, zugehörige zu linearer DGL, 83
- homogene lineare DGL, 80
- indefinite Matrix, 48
- inhomogene lineare DGL, 80
- Integral
 - über projizierbare Mengen, 61
 - Kurven-, 71
 - wegunabhängiges, 74
 - Weg-, 71
 - wegunabhängiges, 74
- inverse Matrix, 7
 - Berechnung von, 7
- invertierbare Matrix, 7
- Jacobi-Matrix, 40
- Koeffizientenmatrix, 1, 2
 - erweiterte, 2
- konservatives Kraftfeld, 74
- konvergente Folge, 28
- Koordinaten
 - Kugel-, 69
 - Polar-, 63
 - Zylinder-, 68
- Koordinatenfolge, 27
- Koordinatenfunktion, 27
- kritischer Punkt eines Skalarfelds, 47
- Kugel, 33
- Kugelkoordinaten, 69
- Kurve, 71
- Kurvenintegral, 71
 - eines Vektorfeldes, 73
 - wegunabhängiges, 74
- Lagrange, Methode von, 55
- Lagrange-Multiplikator, 55
- LGS, 2
- lineare Differentialgleichung, 79
 - 1. Ordnung, 83
 - 2. Ordnung, 86
 - mit konstanten Koeffizienten, 86
 - homogene, 80
 - inhomogene, 80
 - Partikulärlösung, 83
 - spezielle Lösung, 83
 - zugehörige homogene DGL, 83
- lineare Hülle, 16
- lineares Gleichungssystem, 2
- lokales Extremum, 45
- lokales Minimum/Maximum, 45
- $M_{n,m}(\mathbb{R})$, 1
- Masse eines Körpers, 66
- Matrix, 1
 - charakteristisches Polynom, 14
 - Diagonal-, 17

- diagonalisierbare, 17
- Diagonalisierung, 18
- Eigenraum, 14
- Eigenvektor, 13
- Einheits-, 7
- Hesse-, 47
- indefinite, 48
- inverse, 7
 - Berechnung von, 7
- invertierbare, 7
- Jacobi-, 40
- Koeffizienten-, 2
 - erweiterte, 2
- negativ definite, 48
- obere Dreiecks-, 12
- positiv definite, 48
- reguläre, 7
- singuläre, 7
- Streichungs-, 8
- symmetrische, 22
- Maximum
 - globales, 45
 - lokales, 45
- Menge
 - abgeschlossene, 32
 - beschränkte, 32
 - offene, 33
 - projizierbare, 59
 - Standard-, 60
- Methode von Lagrange, 55
- Minimum
 - globales, 45
 - lokales, 45
- Mittelpunkt, 33
- negativ definite Matrix, 48
- nichtlineare Differentialgleichung, 80
- obere Dreiecksmatrix, 12
- offene Menge, 33
- Ordnung einer DGL, 78
- Parametrisierung, 53
- partiell differenzierbare Funktion
 - reellwertig, 35
 - vektorwertig, 39
- partielle Ableitung, 35
 - höherer Ordnung, 41
- Partikulärlösung einer linearen DGL, 83
- Pivotelement, 2
- Polarkoordinaten, 63
- positiv definite Matrix, 48
- Potenzial, 74
- projizierbare Menge, 59
- Punkt, kritischer, 47
- Quelldichte, 42
- $\mathbb{R}^{n \times m}$, 1
- Radius, 33
- Regel von Sarrus, 7
- reguläre Matrix, 7
- Sarrus-Regel, 7
- Satz
 - von Fubini, 61
 - von Lagrange, 55
 - von Schwarz, 41
 - von Weierstraß, 32
- Schwarz, Satz von, 41
- Schwerpunkt, 66
- singuläre Matrix, 7
- Skalarfeld, 26
- Spann, 16
- spezielle Lösung einer linearen DGL, 83
- Stammfunktion, 74
- Standardmenge, 60
- stetig in einem Punkt, 28
- stetig partiell differenzierbare Funktion
 - skalarwertig, 36
 - vektorwertig, 40
- Stetigkeit einer Funktion, 28
- Streichungsmatrix, 8
- Substitutionsregel
 - im \mathbb{R}^3 , 67
 - im \mathbb{R}^2 , 63
- symmetrische Matrix, 22

Index

Umformungen, Elementar-, 5

van-der-Waals-Gleichung, 25, 29, 36

Variation-der-Konstanten-Formel, 85

Vektorfeld, 25, 26

 Divergenz eines, 42

 Wegintegral, 73

Volumen eines Körpers, 66

Weg, 71

 Geschwindigkeit, 73

Wegintegral, 71

 eines Vektorfeldes, 73

 wegunabhängiges, 74

wegunabhängiges Kurvenintegral, 74

wegunabhängiges Wegintegral, 74

Weierstraß, Satz von, 32

x -projizierbare Menge, 59

y -projizierbare Menge, 59

Zahl, Fibonacci-, 22

Zeilenstufenform, 2

Zylinderkoordinaten, 68