

# Einführung in die Mathematik

Tammo tom Dieck

Mathematisches Institut  
Georg-August-Universität  
Göttingen 1997

Version vom 4. Mai 2004

# Inhaltsverzeichnis

<b>1</b>	<b>Elementarmathematik</b>	<b>5</b>
1	Zahlen . . . . .	5
1.1.	Rechenregeln . . . . .	5
1.2.	Potenzen . . . . .	5
1.3.	Zahlensorten . . . . .	6
1.4.	Summen- und Produktzeichen . . . . .	6
1.5.	Anordnung der Zahlen . . . . .	8
1.6.	Der absolute Betrag . . . . .	9
1.7.	Binomialkoeffizienten . . . . .	10
1.8.	Die binomische Formel . . . . .	12
1.9.	Beweisprinzip: Die vollständige Induktion . . . . .	12
2	Die Mengensprache . . . . .	13
2.1.	Mengen . . . . .	13
2.2.	Logische Beziehungen . . . . .	13
2.3.	Produktmenge . . . . .	14
2.4.	Funktionen. Abbildungen . . . . .	14
3	Die Ebene . . . . .	15
3.1.	Die Zahlenebene . . . . .	15
3.2.	Der Kreis . . . . .	15
3.3.	Geraden in der Ebene . . . . .	16
3.4.	Vektoren . . . . .	17
3.5.	Koordinatensysteme . . . . .	18
3.6.	Gleichungen . . . . .	19
3.7.	Parabel . . . . .	19
3.8.	Ellipse . . . . .	20
3.9.	Hyperbel . . . . .	20
3.10.	Strecken . . . . .	20
3.11.	Lineare Abbildungen . . . . .	21
3.12.	Determinante . . . . .	22
3.13.	Komplexe Zahlen . . . . .	22
3.14.	Polarkoordinaten . . . . .	23
<b>2</b>	<b>Differential- und Integralrechnung</b>	<b>24</b>
1	Funktionen . . . . .	24
1.1.	Der Begriff einer Funktion . . . . .	24
1.2.	Das Rechnen mit Polynomen . . . . .	26
1.3.	Verkettung . . . . .	27
1.4.	Intervalle . . . . .	27
1.5.	Weitere Funktionen . . . . .	28
1.6.	Graph einer Funktion . . . . .	28
1.7.	Interpolation . . . . .	29
1.8.	Die Umkehrfunktion . . . . .	29
2	Stetige Funktionen . . . . .	30
2.1.	Stetigkeit . . . . .	31
2.2.	Die Hauptsätze . . . . .	31
2.3.	Grenzwerte . . . . .	33

3	Folgen. Reihen. Grenzwerte . . . . .	33
	3.1. Folgen . . . . .	33
	3.2. Konvergenz. Grenzwert . . . . .	34
	3.3. Konvergenzkriterien für Folgen . . . . .	36
	3.4. Reihen . . . . .	37
	3.5. Konvergenzkriterien für Reihen . . . . .	38
4	Differentialrechnung . . . . .	39
	4.1. Der Differenzenquotient . . . . .	39
	4.2. Die Ableitung . . . . .	40
	4.3. Ableitungsregeln . . . . .	41
	4.4. Umkehrfunktionen . . . . .	43
	4.5. Der Mittelwertsatz . . . . .	43
	4.6. Das lokale Verhalten . . . . .	45
	4.7. Konvex. Konkav . . . . .	45
	4.8. Verallgemeinerter Mittelwertsatz . . . . .	45
	4.9. Grenzwertregeln . . . . .	46
	4.10. Taylor-Polynome und Reihen . . . . .	46
	4.11. Potenzreihen . . . . .	47
	4.12. Partielle Ableitungen . . . . .	47
5	Exponentialfunktion und Logarithmus . . . . .	49
	5.1. Potenzen mit reellen Exponenten . . . . .	49
	5.2. Das Potenzgesetz . . . . .	50
	5.3. Die Exponentialfunktion . . . . .	51
	5.4. Die Differentialgleichung $y' = \alpha y$ . . . . .	51
	5.5. Die Funktionalgleichung der Exponentialfunktion . . . . .	52
	5.6. Der Logarithmus . . . . .	52
	5.7. Die allgemeine Potenz . . . . .	53
	5.8. Berechnung der Exponentialfunktion . . . . .	54
	5.9. Logarithmusreihe und Binomialreihe . . . . .	55
	5.10. Wachstumsmodelle . . . . .	56
6	Trigonometrische Funktionen . . . . .	57
	6.1. Sinus und Cosinus . . . . .	57
	6.2. Die Differentialgleichung $y'' = -y$ . . . . .	58
	6.3. Additionstheoreme . . . . .	59
	6.4. Nullstellen und Perioden . . . . .	59
	6.5. Reihenentwicklung . . . . .	60
	6.6. Tangens und Cotangens . . . . .	60
	6.7. Umkehrfunktionen . . . . .	60
7	Integralrechnung . . . . .	61
	7.1. Das Integral . . . . .	61
	7.2. Das Rechnen mit Integralen . . . . .	63
	7.3. Mittelwertsatz der Integralrechnung. . . . .	64
	7.4. Der Hauptsatz . . . . .	64
	7.5. Anwendungen des Hauptsatzes . . . . .	65
	7.6. Kurvenlänge . . . . .	66

<b>3</b>	<b>Analytische Geometrie. Lineare Algebra</b>	<b>68</b>
1	Lineare Gleichungen . . . . .	68
	1.1. Gleichungssysteme . . . . .	68
	1.2. Matrizen . . . . .	69
	1.3. Zeilenstufenmatrizen . . . . .	70
	1.4. Elementare Zeilenumformungen . . . . .	72
	1.5. Der Gaußsche Algorithmus . . . . .	73
2	Vektorrechnung . . . . .	73
	2.1. Der Standardvektorraum . . . . .	73
	2.2. Skalarprodukt. Länge. Winkel . . . . .	74
	2.3. Hyperebenen . . . . .	77
	2.4. Geraden . . . . .	78
3	Die Methode der kleinsten Quadrate . . . . .	80
4	Koordinatensysteme . . . . .	82
	4.1. Linearkombinationen . . . . .	82
	4.2. Basis. Koordinatensystem . . . . .	84
	4.3. Basiswechsel . . . . .	85
5	Matrizenrechnung . . . . .	86
	5.1. Das Matrizenprodukt . . . . .	86
	5.2. Basiswechsel . . . . .	88
	5.3. Koordinatenwechsel . . . . .	88
	5.4. Inverse Matrizen . . . . .	89
	5.5. Zeilenumformungen . . . . .	89
	5.6. Gleichungssysteme . . . . .	90
6	Lineare Abbildungen . . . . .	91
	6.1. Der Begriff einer linearen Abbildung . . . . .	91
	6.2. Beschreibung durch Matrizen . . . . .	92
	6.3. Drehungen und Spiegelungen der Ebene . . . . .	93
	6.4. Scherungen. Dehnungen und Stauchungen . . . . .	94
7	Determinanten . . . . .	95
	7.1. Determinantenfunktion . . . . .	95
	7.2. Der Entwicklungssatz . . . . .	97
	7.3. Die Cramersche Regel . . . . .	97
	7.4. Der Produktsatz . . . . .	98
	7.5. Determinante von Blockmatrizen . . . . .	99
	7.6. Volumen . . . . .	99
	7.7. Orientierung . . . . .	99
8	Das Vektorprodukt . . . . .	100
9	Eigenwerte . . . . .	102
	9.1. Eigenvektoren. Eigenwerte . . . . .	102
	9.2. Das charakteristische Polynom . . . . .	103

# 1 Elementarmathematik

## 1 Zahlen

### 1.1. Rechenregeln

Was sind eigentlich Zahlen? Auf diese Frage gibt es keine Antwort. Aber jeder hat eine Vorstellung von den Zahlen und kann mit ihnen umgehen, das heißt mit ihnen rechnen, sie nach allgemeinverbindlichen Regeln addieren und multiplizieren. Regeln über die Addition und die Multiplikation sind:

$$(a + b) + c = a + (b + c), \quad (ab)c = a(bc)$$

$$a + b = b + a, \quad ab = ba$$

$$a(b + c) = ab + ac, \quad (a + b)c = ac + bc.$$

In der ersten Zeile stehen die *Assoziativitätsgesetze*, in der zweiten die *Kommutativitätsgesetze* und in der dritten die *Distributivgesetz*. Das Assoziativgesetz der Addition hat als Konsequenz, daß man in einer Summe von mehreren Summanden keine Klammern setzen muß. Das Kommutativgesetz besagt, daß es auf die Reihenfolge der Summanden nicht ankommt. Ebenso für die Faktoren eines Produkts. Weitere wichtige Regeln besagen, daß die Addition eine „Umkehrung“ Subtraktion und die Multiplikation eine „Umkehrung“ Division hat. Das bedeutet: Zu je zwei Zahlen  $a, b$  gibt es genau eine Zahl  $x$ , bezeichnet mit  $x = b - a$ , die die Gleichung  $a + x = b$  erfüllt; zu je zwei Zahlen  $a \neq 0$  und  $b$  gibt es genau eine Zahl, bezeichnet  $x = \frac{b}{a}$  oder  $x = b/a$ , die die Gleichung  $ax = b$  erfüllt. Der Gebrauch dieser Regeln, der Klammern, des Minuszeichens, der Bruchstriche ist in Fleisch und Blut übergegangen und wird hiermit als bekannt vom Tisch gewischt. Aber — viele dieser Regeln gleichzeitig angewendet können zu komplizierten Formeln führen, und diese Formeln sind dann durchaus eine Überlegung wert. Auch führt man für das Resultat gewisser Regeln Abkürzungen ein und sammelt dann Regeln über die abkürzenden Symbole — in Wahrheit gibt es also des Leidens kein Ende, und man sieht sich früher oder später gezwungen, das vom Tisch Gewischte mühsam wieder aufzuklauben.

### 1.2. Potenzen

Ist  $n$  eine natürliche Zahl  $1, 2, 3, \dots$ , so wird mit  $a^n$  das Produkt von  $n$  Faktoren  $a \cdot \dots \cdot a$  der Zahl  $a$  bezeichnet und  $a$  hoch  $n$  gelesen und  $n$ -te *Potenz* von  $a$  genannt. Also zum Beispiel  $2^2 = 4$ ,  $2^3 = 8$ ,  $10^3 = 1000$ . Danach wird  $a^0 = 1$  vereinbart und für  $a \neq 0$  noch  $a^{-n} = \frac{1}{a^n}$  gesetzt. Insbesondere ist also

$$a \cdot a^{-1} = a^{-1} \cdot a = 1, \quad \frac{b}{a} = b \cdot a^{-1} = a^{-1} \cdot b.$$

Zum Beispiel ist  $\frac{25}{24} = 5^2 2^{-3} 3^{-1}$ . Für die Potenzen gelten dann die Regeln (*Potenzgesetze*)

$$(1.1) \quad a^m a^n = a^{m+n}, \quad a^n b^n = (ab)^n, \quad (a^m)^n = a^{mn}.$$

Darin dürfen  $m$  und  $n$  beliebige ganze Zahlen sein. Beispiel:

$$(2^2)^3 = 2^{2 \cdot 3} = 64, \quad 2^{(2^3)} = 2^8 = 256.$$

In einem Symbol  $a^n$  heißt  $a$  die *Basis* und  $n$  der *Exponent*. Das gebräuchliche Dezimalsystem beruht auf den Potenzen der Zahl 10. So ist 1996 die Abkürzung für  $10^3 + 9 \cdot 10^2 + 9 \cdot 10^1 + 6 \cdot 10^0$ .

### 1.3. Zahlensorten

Es gibt verschiedene Sorten von Zahlen mit eigenen Namen. Für die Zahlenmenge einer Sorte verwenden wir die folgenden Bezeichnungen<sup>1</sup>.

$\mathbb{N} = \{1, 2, 3, 4, \dots\}$  Menge der *natürlichen* Zahlen.

$\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$  Menge der natürlichen Zahlen unter Einschluß der Null.

$\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$  Menge der *ganzen* Zahlen.

$\mathbb{Q}$  Menge der *rationalen* Zahlen der Form  $\frac{p}{q}$  mit ganzen Zahlen  $p$  und  $q \neq 0$  („Brüche“).

$\mathbb{R}$  Menge der *reellen* Zahlen.

$\mathbb{R}_+$  Menge der positiven reellen Zahlen.

$\mathbb{R}_-$  Menge der negativen reellen Zahlen.

Reelle Zahlen sind alle endlichen und unendlichen Dezimalbrüche. Einige haben besondere Namen, etwa  $\sqrt{2}$  oder  $\pi = 3,14159265358\dots$ . Die reellen Zahlen entsprechen den Punkten der *Zahlengeraden*. Zahlen, die nicht rational sind, heißen *irrational*;  $\sqrt{2}$  ist irrational.

Endliche oder unendliche Dezimalbrüche liefern eine Art Vorstellung von den Zahlen und ihrer Vielheit. Aber sie sind theoretisch und praktisch nicht recht brauchbar, zum Beispiel kann man im allgemeinen nicht die Dezimalentwicklung einer Summe oder eines Produktes von gegebenen Dezimalentwicklungen ausrechnen.

In diesem Abschnitt berufen wir uns auf ein Vorverständnis. Um zum Beispiel zu verstehen und zu begründen, warum der unendliche Dezimalbruch  $0,99999\dots$  mit lauter Neunen nach dem Komma gleich der Zahl 1 ist<sup>2</sup>, muß man sich im Einzelnen mit dem Grenzwertbegriff vertraut machen.

### 1.4. Summen- und Produktzeichen

Oft hat man Summen und Produkte von sehr vielen Zahlen zu bilden. Wie schreibt man diese in allgemeinen Formel kurz und bündig auf? Die Zahlen werden von einer gewissen Stelle an durchnummeriert; so bezeichnet etwa

$$a_1, a_2, a_3, \dots, a_n$$

ein von 1 bis  $n$  durchnummeriertes System von Zahlen, wobei die unten angehängte Nummer *Index* genannt wird. Die Summe dieser Zahlen kann man salopp mit

$$a_1 + a_2 + a_3 + \dots + a_n$$

<sup>1</sup>Für die Mengennotation sein auf den späteren Abschnitt über Mengensprache verwiesen

<sup>2</sup>und nicht eventuell doch „etwas kleiner“

notieren. Die drei Punkte unterstellen, daß man weiß, wie es weitergeht. Da dieses aber nicht immer der Fall ist, führt man ein neues Symbol, das *Summenzeichen*, ein. Es ist erklärt durch

$$\sum_{k=1}^n a_k = a_1 + a_2 + a_3 + \cdots + a_n.$$

Das große Sigma  $\sum$  darin wird als „Summe über“ gelesen; unten am Sigma wird der Beginn der Numerierung oben das Ende der Numerierung notiert; danach steht dann die „allgemeine Bezeichnung“ eines Summanden, hier  $a_k$ . Der Buchstabe  $k$  kann durch irgendeinen anderen ersetzt werden, ohne daß sich etwas an dem Summenwert ändert. Hier ein Beispiel

$$\sum_{k=3}^7 a_k = \sum_{j=3}^7 a_j = \sum_{l=2}^6 a_{l+1} = a_3 + a_4 + a_5 + a_6 + a_7.$$

Sinngemäß kann man an dem Summenzeichen  $\sum$  auch noch andere Information unterbringen; etwas Gewöhnung und gutwillige Interpretation ist angebracht. Ist etwa  $S$  eine endliche Menge (*Indexmenge*) und  $(a_s \mid s \in S)$  eine Familie von Zahlen  $a_s$ , die mit den Elementen von  $S$  aufgelistet (indiziert) sind, so wird ihre Summe mit

$$\sum_{s \in S} a_s$$

bezeichnet.

Wie multipliziert man zwei solche großen Summen nach dem Distributivgesetz? Wie kann man also etwa die sich nach Auflösen der Klammer in

$$\left( \sum_{j=0}^m a_j \right) \left( \sum_{k=0}^n b_k \right)$$

ergebende Summe aller Terme der Form  $a_j b_k$  hinschreiben? Man kann zum Beispiel eine Doppelsumme schreiben

$$\sum_{j=0}^m \sum_{k=0}^n a_j b_k = \sum_{k=0}^n \sum_{j=0}^m a_j b_k.$$

Oft ist es aber zweckmäßig, die Summanden in bestimmter Weise zu gruppieren. Betrachten wir stattdessen Summen der Form

$$\sum_{k=0}^n a_k x^k = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

worin  $x^k$  natürlich eine  $k$ -te Potenz von  $x$  sein soll. Wenn man zwei solche Summen miteinander multipliziert, ist es ratsam, die Summanden  $a_j x^j b_k x^k = a_j b_k x^{j+k}$  nach dem Exponenten der Potenz von  $x$  zu ordnen. Etwas sinnieren statt fernsehen überzeugt dann von der folgenden Formel

$$(1.2) \quad \left( \sum_{j=0}^m a_j x^j \right) \left( \sum_{k=0}^n b_k x^k \right) = \sum_{i=0}^{m+n} \left( \sum_{l=0}^i a_l b_{i-l} \right) x^i.$$

Für ein Produkt von vielen Zahlen verwendet man analog ein *Produktzeichen* mit einem großen Pi

$$\prod_{j=1}^n a_j = a_1 \cdot a_2 \cdot a_3 \cdot \dots \cdot a_n.$$

Summen und Produktzeichen werden hauptsächlich in allgemeinen Überlegungen gebraucht. Wir geben hier ein Beispiel für das Arbeiten mit dem Summenzeichen. Darin verwenden wir die Gleichung  $\frac{1}{j(j+1)} = \frac{1}{j} - \frac{1}{j+1}$ .

$$\begin{aligned} \sum_{j=1}^n \frac{1}{j(j+1)} &= \sum_{j=1}^n \left( \frac{1}{j} - \frac{1}{j+1} \right) \\ &= \sum_{j=1}^n \frac{1}{j} - \sum_{j=1}^n \frac{1}{j+1} \\ &= 1 + \sum_{j=2}^n \frac{1}{j} - \sum_{j=1}^{n-1} \frac{1}{j+1} - \frac{1}{n+1} \\ &= 1 + \sum_{j=2}^n \frac{1}{j} - \sum_{j=2}^n \frac{1}{j} - \frac{1}{n+1} \\ &= 1 - \frac{1}{n+1} = \frac{n}{n+1}. \end{aligned}$$

### 1.5. Anordnung der Zahlen

Die reellen Zahlen sind der Größe nach geordnet. Die Größenbeziehung wird durch die folgenden Symbole notiert:

$a < b, b > a$  bedeutet:  $a$  ist *kleiner* als  $b$ ,  $b$  ist *größer* als  $a$ .

Auf der Zahlengeraden liegt nach der üblichen Veranschaulichung die größere Zahl rechts von der kleineren. Die Relation  $0 < a$  bedeutet:  $a$  ist *positiv*; die Relation  $0 > a$  bedeutet:  $a$  ist *negativ*. Für eine reelle Zahl  $a$  gilt genau eine der drei Relationen  $a > 0$ ,  $a = 0$ ,  $a < 0$ . Wir verwenden weiterhin auch das folgende Symbol:

$a \leq b, b \geq a$  bedeutet:  $a$  ist *kleiner oder gleich*  $b$ ,  $b$  ist *größer oder gleich*  $a$ .

**(1.3)** Für das Anordnungszeichen  $<$  gelten die folgenden Rechenregeln:

$$\begin{aligned} a < b, b < c &\Rightarrow a < c \\ a < b, c < d &\Rightarrow a + c < b + d \\ a < b, c > 0 &\Rightarrow ac < bc \\ a < b, c < 0 &\Rightarrow ac > bc \\ 0 < a < b &\Rightarrow 0 < \frac{1}{b} < \frac{1}{a}. \end{aligned}$$

Wir benutzen darin  $\Rightarrow$  als Symbol für die logische Folgerung. ◇



Im Zusammenhang mit den Zeichen  $=$ ,  $<$ ,  $\leq$  hat sich folgende Sprechweise eingebürgert: Eine irgendwie geartete Relation der Form  $x = y$  heißt *Gleichung*, eine der Form  $x < y$  oder  $x \leq y$  *Ungleichung*. Anwendung der eben aufgeführten Rechenregeln bezeichnet man dann als *Rechnen mit Ungleichungen*. In der Praxis kommt es sehr häufig vor, daß man nicht mit festen Zahlen arbeiten kann, sondern gewisse Schwankungen, Bandbreiten, Fehlergrenzen und Limitierungen ansetzen oder berücksichtigen muß. Darum ist das Rechnen mit Ungleichungen von fundamentaler Bedeutung. Es versteht sich, daß dieses gleichermaßen für die theoretische Seite der Mathematik gilt. Das wird insbesondere bei allen Gelegenheiten deutlich werden, wo Grenzwerte eine Rolle spielen, also in der Differential- und Integralrechnung.

Wir notieren noch eine Folgerung aus den Regeln, die sich durch vollständige Induktion daraus ergibt:

**(1.4) Notiz.** Sei  $0 \leq a < b$ . Dann gilt für jede natürliche Zahl  $n$   $0 \leq a^n < b^n$ .  $\square$

Sind  $a_1, \dots, a_n$  reelle Zahlen, so wird der Zahlenwert der größten darunter das *Maximum*, der Wert der kleinsten darunter das *Minimum* von  $(a_1, \dots, a_n)$  genannt. Wir verwenden die Bezeichnungen

$$\max(a_1, \dots, a_n), \quad \min(a_1, \dots, a_n).$$

Unendlich viele Zahlen haben nicht immer ein Maximum oder ein Minimum. So hat  $(1, \frac{1}{2}, \frac{1}{3}, \dots)$  zwar das Maximum 1 aber kein Minimum.

**(1.5) Beispiel.** Als eine Anwendung des Rechnens mit Ungleichungen folgt durch vollständige Induktion für  $x > -1$  und natürliche Zahlen  $n$  die *Bernoullische Ungleichung*

$$(1 + x)^n \geq 1 + nx.$$

Für  $x \geq 0$  erkennt man sie auch direkt aus der binomischen Formel (siehe dazu 1.8).  $\diamond$

## 1.6. Der absolute Betrag

Der (*absolute*) *Betrag*  $|a|$  einer reellen Zahl wird durch

$$|a| = \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a \leq 0 \end{cases}, \quad |a| = \max(a, -a)$$

erklärt. So gilt etwa  $|0| = 0$ ,  $|5| = |-5| = 5$ .

**(1.6)** Für das Betragszeichen gelten die folgenden Regeln:

$$\begin{aligned} |a| &\geq 0 \\ |a + b| &\leq |a| + |b| \\ |ab| &= |a||b|. \end{aligned}$$

So ist  $|2| = |5 + (-3)| < |5| + |3| = 8$  und  $8 = |5 + 3| = |5| + |3|$ . Die zweite dieser Regeln heißt *Dreiecksungleichung*. Sie läßt sich durch vollständige Induktion (siehe Abschnitt 1.9) auf eine Summe mehrerer Zahlen verallgemeinern

$$|a_1 + a_2 + \dots + a_n| \leq |a_1| + |a_2| + \dots + |a_n|.$$

Das Betragszeichen dient unter anderem dazu, Fehlerschranken anzugeben. So bedeutet zum Beispiel

$$|a - a_0| < \frac{1}{1000},$$

daß die Zahl  $a$  höchstens um  $\frac{1}{1000}$  von der Zahl  $a_0$  abweicht. Folgende Aussagen sind äquivalent:

$$|a| < \varepsilon \quad \Leftrightarrow \quad -\varepsilon < a < \varepsilon.$$

Diese Umformung wird im folgenden oft stillschweigend verwendet. Die Dreiecksungleichung liefert  $|b| = |a + b - a| \leq |b - a| + |a|$ , also  $|b| - |a| \leq |b - a|$ . Durch Vertauschen von  $a$  und  $b$  und Multiplikation mit  $-1$  erhält man  $|b| - |a| \geq -|b - a|$ . Insgesamt folgt also

$$||b| - |a|| \leq |b - a|.$$

Man überlege sich  $\max(x, y) = \frac{1}{2}(x + y + |x - y|)$ .

### 1.7. Binomialkoeffizienten

Sei  $n$  eine natürliche Zahl. Das Produkt der Zahlen von 1 bis  $n$  wird  $n$ -Fakultät genannt und  $n!$  bezeichnet:

$$(1.7) \quad 1 \cdot 2 \cdot 3 \cdot \dots \cdot n = n!.$$

Wir vereinbaren außerdem, daß  $0! = 1$  sein soll. Es gilt also

$$1! = 1, \quad 2! = 2, \quad 3! = 6, \quad 4! = 24, \quad 5! = 120, \quad 6! = 720.$$

Welche Bedeutung hat die Zahl  $n$ -Fakultät? Sie tritt bei vielen kombinatorischen Abzählproblemen auf.

**(1.8) Notiz.** *Es gibt  $n!$  Möglichkeiten,  $n$  Objekte der Reihe nach anzuordnen.*

BEWEIS. Für das erste Objekt gibt es  $n$  Möglichkeiten. Haben wir dieses gewählt, so stehen noch  $n - 1$  für das zweite Objekt zur Wahl. Usw.  $\square$

Hier sind die  $6 = 3!$  Anordnungen von 1, 2, 3:

$$123, 132, 213, 231, 312, 321.$$

Für  $0 \leq k \leq n$  kürzen wir ab

$$(1.9) \quad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

und lesen diesen Ausdruck  $n$  über  $k$ . Zahlen dieser Form heißen *Binomialkoeffizienten*, weil sie in der alsbald besprochenen binomischen Formel vorkommen. Wir werden gleich sehen, daß es sich um natürliche Zahlen handelt. Es ist bequem, unter dem Symbol  $\binom{n}{k}$  die Zahl Null zu verstehen, wenn  $k < 0$  oder  $k > n$  ist. Der Definition nach gilt

$$\binom{n}{k} = \binom{n}{n-k}, \quad \binom{n}{1} = n, \quad \binom{n}{2} = \frac{n(n-1)}{2}.$$

Welche Bedeutung haben diese Zahlen?

**(1.10) Notiz.** Es gibt  $\binom{n}{k}$  Teilmengen mit  $k$  Elementen in einer Menge mit  $n$  Elementen.

BEWEIS. Wir greifen nacheinander  $k$  Elemente heraus. Für das erste gibts es  $n$  Möglichkeiten, usw. Insgesamt erhalten wir  $n(n-1)\cdots(n-k+1)$  Möglichkeiten. Nun liefern aber verschiedene Anordnungen der  $k$  herausgegriffenen Elemente dieselbe Menge. Also müssen wir noch durch die Anzahl  $k!$  der Anordnungen dividieren. Indem man in (1.8) geeignet kürzt, sieht man, daß

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!}$$

ist. □

Es gibt  $\binom{49}{6} = 13\,983\,816$  Möglichkeiten, beim Zahlenlotto 6 Zahlen aus 49 herauszugreifen. Die  $\binom{4}{2} = 6$  zweielementigen Teilmengen von  $\{1, 2, 3, 4\}$  sind

$$\{1, 2\}, \quad \{1, 3\}, \quad \{1, 4\}, \quad \{2, 3\}, \quad \{2, 4\}, \quad \{3, 4\}.$$

**(1.11) Notiz.** Die Binomialkoeffizienten erfüllen die sogenannte Pascalsche Formel

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

BEWEIS. Wir formen die rechte Seite nach den Regeln der Bruchrechnung um.

$$\begin{aligned} \binom{n}{k} + \binom{n}{k-1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n-k+1)!} \\ &= \frac{n!}{(k-1)!(n-k)!} \cdot \left( \frac{1}{k} + \frac{1}{n-k+1} \right) \\ &= \frac{n!}{(k-1)!(n-k)!} \cdot \frac{n+1}{k(n-k+1)} \\ &= \frac{(n+1)!}{k!(n-k+1)!} \\ &= \binom{n+1}{k} \end{aligned}$$

In der vorstehenden Rechnung sei  $1 \leq k \leq n$ . Aus den Vereinbarungen erkennt man, daß die behauptete Gleichheit auch für  $k \leq 0$  und  $k \geq n+1$  gilt. □

Die durch die Pascalsche Formel gefundene schrittweise (=rekursive) Berechnung der Binomialkoeffizienten wird oft im *Pascalschen Dreieck* notiert

$$\begin{array}{cccccc} & & & & & & 1 \\ & & & & & & & 1 \\ & & & & & & & & 1 \\ & & & & & & & & & 1 \\ & & & & & & & & & & 1 \\ & & & & & & & & & & & 1 \\ & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & & & 1 \\ & & & & & & & & & & & & & & & & 1 \end{array}$$

In jeder neuen Zeile ist jeweils eine Zahl die Summe der unmittelbar schräg darüberstehenden; die Seitenkanten bestehen aus Einsen.

### 1.8. Die binomische Formel

Die binomische Formel berechnet  $(x + y)^n$  für jede natürliche Zahl  $n$  durch Auflösen der Klammern nach dem Distributivgesetz. Durch direktes Ausrechnen erhält man (vergleiche das Pascalsche Dreieck)

$$\begin{aligned}(x + y)^1 &= x + y \\(x + y)^2 &= x^2 + 2xy + y^2 \\(x + y)^3 &= x^3 + 3x^2y + 3xy^2 + y^3 \\(x + y)^4 &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4 \\(x + y)^5 &= x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5\end{aligned}$$

Die Frage ist: Wie geht es weiter?

Die folgende Aussage wird als *Binomische Formel* bezeichnet:

**(1.12) Satz.** *Für alle natürliche Zahlen  $n$  gilt*

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}.$$

BEWEIS. Wir nehmen einmal an, die Gleichung sei für eine bestimmte natürliche Zahl  $n$  richtig. Wir multiplizieren beide Seiten der Gleichung mit  $x + y$ . Links ergibt sich  $(x + y)^{n+1}$ . Die rechte Seite rechnen wir nach dem Distributivgesetz und der Pascalschen Formel um. Zunächst ist

$$\left( \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \right) (x + y) = \sum_{k=0}^n \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^n \binom{n}{k} x^k y^{n-k+1}.$$

Der Ausdruck  $x^k y^{n+1-k}$  kommt für  $1 \leq k \leq n$  in beiden Summen vor, und zwar mit den Koeffizienten

$$\binom{n}{k} \quad \text{und} \quad \binom{n}{k-1}.$$

Diese fassen wir nach der Pascalschen Formel zusammen. Es bleiben dann noch die Fälle  $k = 0$  und  $k = n + 1$ , die beide insgesamt mit dem Koeffizienten 1 auftreten. Nach dieser Umformung wird also aus der Doppelsumme genau die rechte Seite der binomischen Formel, wenn man darin  $n$  durch  $n + 1$  ersetzt. Wir haben also gezeigt: Falls die Formel für  $n$  gilt, so auch für  $n + 1$ . Da die Formel offenbar für den Anfangswert  $n = 1$  gilt, so gilt sie nach dem Beweisprinzip der vollständigen Induktion allgemein.  $\square$

### 1.9. Beweisprinzip: Die vollständige Induktion

Es sei für jede natürliche Zahl  $n$  eine Aussage  $A(n)$  gegeben (zum Beispiel eine von  $n$  abhängige Gleichung wie die binomische Formel). Ist die Aussage für  $n = 1$  richtig und folgt aus der Richtigkeit von  $A(n)$  diejenige von  $A(n + 1)$  („Schluß von  $n$  auf  $n + 1$ “, der „Induktionsschritt“), so gilt die Aussage allgemein.  $\diamond$

Der intuitive Sinn dieses Prinzips ist klar: Es gilt  $A(1)$  nach Voraussetzung, also  $A(2)$ , also  $A(3)$ , also ... ad infinitum. Nur kann man in endlicher Lebenszeit nicht

unendlich viele Schlüsse wirklich durchführen, und da hilft man sich mit dem *deus ex machina* „Prinzip“.

Als Anwendung des Prinzips erkennt man aus der Pascalschen Formel, daß die Binomialkoeffizienten natürliche Zahlen sind. Das folgt allerdings auch aus der Interpretation (1.9).

## 2 Die Mengensprache

Es hat sich weltweit durchgesetzt, mathematische Begriffe und Objekte in der Mengensprache zu formulieren. Vom Folgenden muß man jedenfalls für diesen Text den ersten Abschnitt zur Kenntnis nehmen.

### 2.1. Mengen

Eine *Menge* ist eine gedankliche Zusammenfassung von Dingen zu einem neuen Objekt. Die zusammengefaßten Dinge heißen dann die *Elemente* der so erklärten Menge. Elemente einer Menge können etwa Zahlen, Funktionen, Kreise, andere Mengen, ... sein.

Das Symbol  $a \in M$  bedeutet:  $a$  ist *Element der Menge*  $M$ . Dafür sagt man auch:  $a$  liegt in  $M$ , oder:  $a$  ist in  $M$  enthalten. Eine Menge ist durch ihre Elemente bestimmt. So bedeutet  $a \in \mathbb{Z}$ , daß  $a$  eine ganze Zahl ist.

Es gibt verschiedene Möglichkeiten, Mengen zu beschreiben und zu erklären.

- (1) Aufzählung aller ihrer Elemente in geschweiften Klammern.
- (2) Zusammenfassung von Dingen mit einer bestimmten Eigenschaft  $E$ . Symbolisch

$$M = \{x \mid x \text{ hat die Eigenschaft } E\}.$$

So ist etwa  $\{2, 3, 6, 7\}$  die Menge, deren Elemente 2, 3, 6 und 7 sind. Ferner ist

$$\{x \mid x \in \mathbb{R}, 3 < x < 5\}$$

die Menge aller reellen Zahlen, die zwischen 3 und 5 liegen.

### 2.2. Logische Beziehungen

Es gibt verschiedene logische Beziehungen zwischen Mengen, für die eigene Symbole verwendet werden.

Eine Menge  $A$  heißt *Teilmenge* oder *Untermenge* der Menge  $B$ , wenn jedes Element von  $A$  auch Element von  $B$  ist. Dieser Sachverhalt wird durch das Symbol  $A \subset B$  ausgedrückt (der Fall  $A = B$  ist dabei nicht ausgeschlossen). Beispielsweise gilt

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{R}.$$

Es gilt genau dann  $A = B$ , wenn sowohl  $A \subset B$  ist als auch  $B \subset A$ . Darin steckt eine *Methode*, die Gleichheit zweier Mengen  $A$  und  $B$  zu beweisen: Man nehme ein beliebiges Element aus  $A$  und zeige, daß es auch in  $B$  liegt; und umgekehrt.

Der *Durchschnitt*  $A \cap B$  zweier Mengen  $A$  und  $B$  ist die Menge der Elemente, die sowohl in  $A$  als auch in  $B$  liegen. Man bezeichnet  $A \cap B$  auch als *Schnittmenge*.

Die *Vereinigung*  $A \cup B$  zweier Mengen  $A$  und  $B$  ist die Menge der Elemente, die entweder in  $A$  oder in  $B$  oder in beiden enthalten ist.

Beispiel. Ist  $A = \{x \mid -1 < x < 2\}$  und  $B = \{y \mid 1 < y < 3\}$ , so gilt  $A \cap B = \{z \mid 1 < z < 2\}$  und  $A \cup B = \{z \mid -1 < z < 3\}$ .

Als eine Besonderheit betrachtet man die mit  $\emptyset$  bezeichnete *leere Menge*, die überhaupt kein Element enthält. Gilt  $A \cap B = \emptyset$ , so heißen die Mengen  $A$  und  $B$  *disjunkt*. So ist  $\mathbb{R} = \mathbb{R}_+ \cup \{0\} \cup \mathbb{R}_-$  eine Zerlegung in paarweise disjunkte Teilmengen.

### 2.3. Produktmenge

Sind  $X$  und  $Y$  Mengen, so wird mit  $X \times Y$  die Menge aller geordneten Paare  $(x, y)$  mit  $x \in X$  und  $y \in Y$  bezeichnet und (*cartesisches*) *Produkt* von  $X$  und  $Y$  genannt.

### 2.4. Funktionen. Abbildungen

Seien  $A$  und  $B$  Mengen. Eine *Abbildung von  $A$  nach  $B$*  ist eine Vorschrift, die jedem  $a \in A$  ein  $b \in B$  zuordnet. Die Vorschrift wird meist durch ein Funktionsymbol  $b = f(a)$  geschrieben. Man sagt auch,  $a$  werde auf  $f(a)$  abgebildet. In Symbolen schreiben wir dann

$$f: A \rightarrow B, \quad a \mapsto f(a).$$

Manchmal schreiben wir auch  $f$  neben oder über den Pfeil. Dabei heißt  $A$  der *Definitionsbereich* oder die *Quelle* der Abbildung,  $B$  der *Bildbereich* oder das *Ziel* der Abbildung. Die „Vorschrift“ ist durch die Teilmenge

$$\{(a, b) \mid b = f(a), a \in A\} \subset A \times B$$

bestimmt und kann mengentheoretisch mit ihr gleichgesetzt werden. Zu einer Abbildung gehören also die *drei* Daten  $A$ ,  $B$  und  $f$ , wenn man präzise sein will. In der Praxis erlaubt man sich allerdings gewisse Freiheiten. Statt Abbildung sagt man auch *Funktion*, insbesondere dann, wenn die Zielmenge aus Zahlen besteht. In einigen Fällen sind auch Worte wie *Operator* oder *Funktional* gebräuchlich. Die Abbildung  $A \times B \rightarrow A$ ,  $(a, b) \mapsto a$  heißt *Projektion auf den Faktor  $A$* . Die Abbildung  $\text{id}(A) = \text{id}: A \rightarrow A$ ,  $a \mapsto a$  heißt die *identische Abbildung* von  $A$ . Die *Verkettung* der Abbildungen  $f: A \rightarrow B$  und  $g: B \rightarrow C$  ist die Abbildung  $A \rightarrow C$ , die  $a$  auf  $g(f(a))$  abbildet; sie wird mit  $gf$  oder  $g \circ f$  bezeichnet.

Zu einer Inklusion  $A \subset B$  gehört die Abbildung  $i: A \rightarrow B$ ,  $a \mapsto a$ , die wir ebenfalls *Inklusion* nennen und auch mit  $i: A \subset B$  mißbräuchlich bezeichnen. Eine gegebene Abbildung  $f: B \rightarrow Y$  besitzt die *Einschränkung*  $f \circ i: A \rightarrow Y$  auf die Teilmenge  $A \subset B$ , die auch  $f|_A$  notiert wird.

Eine Abbildung  $f: A \rightarrow B$  heißt *injektiv*, wenn aus  $a_1 \neq a_2$  immer  $f(a_1) \neq f(a_2)$  folgt, und *surjektiv*, wenn zu jedem  $b \in B$  ein  $a \in A$  mit  $b = f(a)$  existiert. Sie heißt *bijektiv*, wenn sie sowohl surjektiv als auch injektiv ist. Eine Abbildung  $f: A \rightarrow B$  ist genau dann bijektiv, wenn sie eine *Umkehrabbildung*  $g: B \rightarrow A$  hat, d. h. wenn eine Abbildung  $g$  existiert mit den Eigenschaften  $gf = \text{id}(A)$  und  $fg = \text{id}(B)$ . Ist  $C \subset A$ , so heißt  $f(C) = \{b \mid \text{es gibt } c \in C \text{ mit } b = f(c)\}$  das *Bild* von  $C$  bei  $f$ . Ist  $D \subset B$ , so heißt  $f^{-1}(D) = \{a \mid f(a) \in D\}$  das *Urbild* von  $D$  bei  $f$ . Das Urbild der einelementigen Teilmenge  $\{c\}$  wird  $f^{-1}(c)$  geschrieben.

Zwei Mengen heißen *gleichmächtig*, wenn es eine Bijektion (= eine bijektive Abbildung) zwischen ihnen gibt. Die *Mächtigkeit* einer Menge  $M$  ist die „Anzahl“ ihrer Elemente (was immer das für unendliche Mengen bedeuten mag); sie wird mit  $|M|$

bezeichnet. Eine Menge heißt *abzählbar unendlich*, wenn es eine Bijektion zur Menge der natürlichen Zahlen gibt.

Die Menge der ganzen Zahlen und die Menge der rationalen Zahlen ist abzählbar, dagegen nicht die Menge der reellen Zahlen.

### 3 Die Ebene

#### 3.1. Die Zahlenebene

Die *Zahlenebene*  $\mathbb{R}^2$  ist die Menge aller geordneten Paare  $(x, y)$  reeller Zahlen  $x$  und  $y$ . Das Wort *geordnet* bezieht sich hier darauf, daß es auf die Reihenfolge ankommt: Für  $x \neq y$  sind die Paare  $(x, y)$  und  $(y, x)$  verschieden. Ein Zahlenpaar nennt man in diesem Kontext *Punkt* der Ebene. Die Zahlenebene ist die rechnerische Grundlage der ebenen Geometrie.

Die Menge  $\mathbb{R}^2$  wird durch die Menge der Punkte einer Ebene veranschaulicht. Dabei ist in der Ebene ein Punkt, der *Ursprung* oder *Nullpunkt* ausgezeichnet, und die Lage eines Punktes wird in einem rechtwinkligen Koordinatensystem durch ein Zahlenpaar festgelegt. Die Punkte der Form  $(x, 0)$  bilden die *Abszisse*, oder auch  $x$ -Achse genannt. Die Punkte der Form  $(0, y)$  bilden die *Ordinate*, oder auch  $y$ -Achse genannt. Beide Achsen heißen *Koordinatenachsen* und bilden ein *Koordinatensystem*. Wird ein „allgemeiner“ Punkt mit  $(x_1, x_2)$  bezeichnet, so spricht man natürlich von  $x_1$ - und  $x_2$ -Achse. Der Punkt  $(0, 0)$  heißt *Ursprung* oder *Nullpunkt* des Koordinatensystems. Man nennt  $x$  die erste und  $y$  die zweite *Koordinate* des Punktes  $P = (x, y)$ .

Die sogenannte *analytische Geometrie* hat als Ziel, geometrische Objekte, Begriffe und Sachverhalte rein algebraisch mit Hilfe von Zahlen und Funktionen zu untersuchen. Wir werden sogleich einige Beispiel dieser Methode kennenlernen.

#### 3.2. Der Kreis

Der aus der Elementargeometrie bekannte *Satz des Pythagoras* wird als Motivierung dafür benutzt, den Abstand zweier Punkte der Zahlenebene zu definieren. Und zwar wird der *Abstand* von  $(x_1, y_1)$  und  $(x_2, y_2)$  als die Zahl

$$r = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

erklärt.

Ein *Kreis* in der Ebene vom *Radius*  $r$  um den *Mittelpunkt*  $m = (x_1, y_1)$  ist die Menge der Punkte, die von  $m$  den Abstand  $r$  haben. In der mengentheoretischen Notation ist dieser Kreis also die Menge

$$\{(x, y) \mid r = \sqrt{(x - x_1)^2 + (y - y_1)^2}\}.$$

Anders formuliert: Die Punkte  $(x, y)$ , die auf dem Kreis um  $(x_1, y_1)$  mit dem Radius  $r$  liegen, sind die Lösungen der *Kreisgleichung*

$$(3.1) \quad (x - x_1)^2 + (y - y_1)^2 = r^2.$$

Statt Kreis sagt man zur Verdeutlichung auch *Kreislinie* — zur Unterscheidung vom *Vollkreis*

$$\{(x, y) \mid r \geq \sqrt{(x - x_1)^2 + (y - y_1)^2}\}$$

aller Punkte, die von  $m$  höchstens den Abstand  $r$  haben.

### 3.3. Geraden in der Ebene

Seien  $a, b, c$  reelle Zahlen und  $a, b$  nicht beide gleich Null. Eine Punktmenge der Form

$$G = \{(x, y) \mid ax + by = c\}$$

heißt dann *Gerade* in der Ebene und

$$ax + by = c$$

ihre *Geradengleichung*. Ist  $(x, y) \in G$ , so sagen wir, die Gerade  $G$  *läuft durch* den Punkt  $(x, y)$ , der Punkt  $(x, y)$  *liegt auf* der Geraden  $G$ .

Sind zwei Geradengleichungen

$$a_1x + b_1y = c_1, \quad a_2x + b_2y = c_2$$

gegeben und  $G_1, G_2$  die zugehörigen Geraden, so sagt die Anschauung der ebenen Geometrie, daß die folgenden drei Fälle auftreten können.

1. Fall:  $G_1 = G_2$ .
2. Fall:  $G_1 \cap G_2 = \emptyset$ .
3. Fall:  $G_1 \cap G_2 = \{(x_0, y_0)\}$ .

Im zweiten Fall sagen wir,  $G_1$  und  $G_2$  seien *parallel*. Im dritten Fall ist  $(x_0, y_0)$  der eindeutig bestimmte *Schnittpunkt* der beiden Geraden. Wir erläutern nun, was diese drei Fälle rechnerisch bedeuten und wie man den Schnittpunkt aus den Geradengleichungen ausrechnet.

Die Bestimmung der Schnittpunkte von zwei Geraden führt auf die Betrachtung von *linearen Gleichungssystemen*

$$(3.2) \quad \begin{aligned} ax + by &= p \\ cx + dy &= q. \end{aligned}$$

Die *Lösungsmenge* des Systems (3.2) ist die Menge aller  $(x, y) \in \mathbb{R}^2$ , die beide Gleichungen erfüllen. Das Schema

$$(3.3) \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

heißt die *Koeffizientenmatrix*  $A$  des Systems (3.2).

Wir wollen im folgenden voraussetzen, daß jede Zeile in (3.2) eine Gerade beschreibt, d. h. weder  $a, b$  noch  $c, d$  sind beide Null. Zunächst überlegen wir uns:

**(3.4) Notiz.** *Genau dann gilt  $ad - bc = 0$ , wenn es eine Zahl  $\lambda \neq 0$  gibt, mit der  $\lambda a = c$ ,  $\lambda b = d$  gilt.*

BEWEIS. Gibt es ein derartiges  $\lambda$ , so ist  $ad - bc = \lambda ab - \lambda ab = 0$ . Sei umgekehrt  $ad = bc$ . Ist  $b \neq 0$ , so setzen wir  $\lambda = \frac{d}{b}$ . Ist  $b = 0$ , so ist nach Voraussetzung  $a \neq 0$ . Wir setzen dann  $\lambda = \frac{c}{a}$ .  $\square$

**(3.5) Satz.** *Die Lösungsmenge von (3.2) besteht genau dann aus einem Punkt, wenn  $ad - bc \neq 0$  ist. Ist  $ad - bc = 0$ , so ist die Lösungsmenge entweder leer oder eine Gerade.*



BEWEIS. Sei  $ad - bc \neq 0$  und  $(x, y)$  eine Lösung. Wir multiplizieren die erste Gleichung in (3.2) mit  $d$ , die zweite mit  $b$ , subtrahieren dann die zweite von der ersten und erhalten  $(ad - bc)x = pd - qb$ . Auf diese Weise sehen wir, daß eine Lösung  $(x, y)$  im Falle  $ad - bc \neq 0$  notwendig die Form

$$(3.6) \quad x = \frac{pd - qb}{ad - bc}, \quad y = \frac{pa - pc}{ad - bc}$$

hat. Diese Werte erfüllen auch wirklich die Gleichungen (3.2), wie eine Probe zeigt.

Ist  $ad - bc = 0$ , so gibt es ein  $\lambda \neq 0$  wie in (3.4). Jede Lösung des Systems (3.2) ist dann auch eine Lösung des Systems

$$\begin{aligned} \lambda ax + \lambda by &= \lambda p \\ cx + dy &= q, \end{aligned}$$

und umgekehrt. Wegen  $\lambda a = c$ ,  $\lambda b = d$  widersprechen sich im Fall  $\lambda p \neq q$  die Gleichungen — es gibt keine Lösung. Im Fall  $\lambda p = q$  sind beide Zeilen gleich und die Lösungsmenge ist eine Gerade.  $\square$

Wir interpretieren den Satz (3.5) geometrisch und erhalten folgende Aussagen über die Geraden

$$L = \{(x, y) \mid ax + by = p\} \quad \text{und} \quad L' = \{(x, y) \mid cx + dy = q\} :$$

- (1) Genau dann ist  $L = L'$ , wenn es eine reelle Zahl  $\lambda \neq 0$  so gibt, daß  $\lambda a = c$ ,  $\lambda b = d$ ,  $\lambda p = q$  ist.
- (2) Genau dann ist  $L \cap L' = \emptyset$ , wenn es eine reelle Zahl  $\lambda \neq 0$  gibt, so daß  $\lambda a = c$ ,  $\lambda b = d$ ,  $\lambda p \neq q$  ist.
- (3) Genau dann schneiden sich  $L$  und  $L'$  in einem Punkt, wenn  $ad - bc \neq 0$  ist. Der Schnittpunkt ist durch (3.6) gegeben.

Als weitere Folgerung erhalten wir:

**(3.7) Satz.** *Seien  $(x_1, y_1)$  und  $(x_2, y_2)$  zwei verschiedene Punkte, d. h.  $y_2 - y_1$  und  $x_2 - x_1$  seien nicht beide gleich Null. Dann gibt es genau eine Gerade, die durch diese Punkte läuft. Sie wird durch die Gleichung*

$$(y_2 - y_1)x - (x_2 - x_1)y = x_1y_2 - x_2y_1$$

gegeben.

BEWEIS. Haben zwei Geraden mehr als einen gemeinsamen Punkt, so sind sie gleich, wie wir eben überlegt haben. Man bestätigt durch Einsetzen, daß  $(x_1, y_1)$  und  $(x_2, y_2)$  auf der angegebenen Geraden liegen.  $\square$

Die Geradengleichung  $x = 0$  beschreibt die  $y$ -Achse, die Geradengleichung  $y = 0$  die  $x$ -Achse.

### 3.4. Vektoren

Die Elemente von  $\mathbb{R}^2$  haben eine zweite Bedeutung. Sie ergibt sich aus zwei Rechenoperationen für Zahlenpaare. Wir erklären eine *Addition* von Zahlenpaaren durch<sup>3</sup>

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2)$$

<sup>3</sup> $A := B$  bedeutet,  $A$  wird durch  $B$  definiert.

und eine Multiplikation mit einer reellen Zahl  $\lambda$  (auch *Skalarmultiplikation* genannt) durch

$$\lambda(x, y) := (\lambda x, \lambda y).$$

Wenn wir in dieser Weise mit Zahlenpaaren rechnen, heißen sie *Vektoren*. Die zeichnerische Veranschaulichung ist die folgende: Der *Vektor*  $(x, y)$  wird durch einen Pfeil von  $(0, 0)$  nach  $(x, y)$  dargestellt. Die Summe von  $u = (x_1, y_1)$  und  $v = (x_2, y_2)$  erhalten wir, indem wir das von den beiden Vektoren  $u$  und  $v$  aufgespannte Parallelogramm betrachten; die vom Nullpunkt ausgehende Diagonale ist die Summe (siehe dazu Abschnitt III.2).

**(3.8) Notiz.** *Ist  $(a, b) \neq 0$ , so liegen die Punkte  $\lambda(a, b) = (\lambda a, \lambda b)$  alle auf der Geraden durch  $(0, 0)$  und  $(a, b)$  und füllen sie ganz aus.*

BEWEIS. Die Gleichung der Geraden ist  $bx - ay = 0$ . Nun wenden wir (3.4) an.  $\square$

Statt durch Gleichungen lassen sich Geraden durch *Parameterdarstellungen* angeben. Sind  $(a, b)$  und  $(c, d)$  gegeben, und ist  $(a, b) \neq (0, 0)$ , so ist die Menge

$$L = \{(c, d) + \lambda(a, b) \mid \lambda \in \mathbb{R}\}$$

eine Gerade durch  $(c, d)$ . Wir nennen  $(a, b)$  ihre *Richtung*. Eine kleine Rechnung zeigt, daß  $L$  die Lösungsmenge der Gleichung

$$-bx + ay = ad - bc$$

ist.

Die Unterscheidung von „Punkten“ und „Vektoren“ ist an dieser Stelle rein sprachlicher Natur und mag etwas künstlich erscheinen. Man spricht von Punkten, wenn man an einer Stelle, an einem Ort in der Ebene interessiert ist. Man spricht von Vektoren, wenn man an den Rechenobjekten interessiert ist.

### 3.5. Koordinatensysteme

Für viele geometrische Untersuchungen ist es wichtig, andere Koordinatensysteme zu betrachten. Wir nennen ein geordnetes Paar von Vektoren  $(a, b), (c, d)$  ein *Koordinatensystem* oder eine *Basis* der Ebene, wenn sie nicht auf einer Geraden durch den Nullpunkt liegen. Nach (3.4) und (3.8) ist das zu der Bedingung  $ad - bc \neq 0$  äquivalent.

**(3.9) Satz.** *Sei  $(a, b), (c, d)$  ein Koordinatensystem. Jeder Punkt  $(u, v) \in \mathbb{R}^2$  besitzt eine eindeutig bestimmte Darstellung der Form*

$$(u, v) = \lambda(a, b) + \mu(c, d).$$

BEWEIS. Nach den in Unterabschnitt 3.4 festgelegten Rechenregeln ist die Gleichung des Satzes äquivalent zu den beiden Gleichungen

$$\begin{aligned} (3.10) \quad u &= a\lambda + c\mu \\ v &= b\lambda + d\mu. \end{aligned}$$

Bei gegebenem  $(u, v)$  haben wir die Gleichungen nach  $(\lambda, \mu)$  aufzulösen. Nach (3.5) hat das Gleichungssystem wegen  $ad - bc \neq 0$  genau eine Lösung  $(\lambda, \mu)$ .  $\square$

Wir nennen  $(\lambda, \mu)$  die *Koordinaten* des Punktes  $(u, v)$  *bezüglich des Koordinatensystems*  $(a, b), (c, d)$ . Das ursprünglich gegebene rechtwinklige Koordinatensystem wird in der jetzt verwendeten Terminologie durch die sogenannte *Standardbasis*

$$e_1 = (1, 0), \quad e_2 = (0, 1)$$

gegeben. Wegen

$$(\lambda, \mu) = \lambda e_1 + \mu e_2$$

sind dann  $(\lambda, \mu)$  auch die Koordinaten des Punktes  $(\lambda, \mu)$  bezüglich  $e_1, e_2$ .

Die Gleichungen (3.10) zeigen, wie sich die Koordinaten des  $(e_1, e_2)$ -Systems aus den Koordinaten des  $((a, b), (c, d))$ -Systems ausrechnen lassen. Wir bezeichnen deshalb (3.10) auch als *Koordinatentransformation*. Die umgekehrte Koordinatentransformation wird durch Auflösen der Gleichungen gemäß (3.6) nach  $\lambda, \mu$  gegeben. Wir nennen im gegenwärtigen Kontext (3.3) die *Matrix des Koordinatensystems* und der Koordinatentransformation (3.10).

Besonders interessant ist ein um den Winkel  $\varphi$  gegenüber dem ursprünglichen System gegen den Uhrzeigersinn gedrehtes System. Es wird durch die *Drehmatrix*

$$(3.11) \quad D(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

beschrieben (siehe dazu Abschnitt III.5).

### 3.6. Gleichungen

Die analytische Geometrie beschreibt, wie schon gesagt, ebene geometrische Objekte als Punktmenge. Oft sind Punktmenge als Lösungsmengen einer Gleichung (oder eines Gleichungssystems) gegeben. Das allgemeine Schema ist folgendes. Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  eine Funktion. Dann heißt

$$N(f) = \{(x, y) \mid f(x, y) = 0\}$$

die *Nullstellenmenge von f*. Ist zum Beispiel  $f(x, y) = \alpha x + \beta y - \gamma$ , so ist  $N(f)$  eine Gerade (falls  $(\alpha, \beta) \neq (0, 0)$ ).

Ist eine Koordinatentransformation (3.10) gegeben, so liefert  $f$  eine neue Funktion  $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,

$$g(\lambda, \mu) = f(a\lambda + c\mu, b\lambda + d\mu).$$

Dieselbe Menge  $N(f)$  wird jetzt bezüglich der neuen Koordinaten  $(\lambda, \mu)$  durch  $N(g)$  beschrieben. Der Kreis vom Radius  $r$  um den Nullpunkt ist  $N(f)$  für  $f(x, y) = x^2 + y^2 - r^2$ .

An dieser Stelle muß man aufpassen: Man kann den Bezeichnungen nicht ansehen, ob  $(\lambda, \mu)$  die Koordinaten eines Punktes im  $\mathbb{R}^2$  bezüglich eines anderweitig festgelegten Koordinatensystems sind, oder ob  $(\lambda, \mu)$  einfach ein Punkt der Ebene ist. Für das Koordinatensystem  $e_1, e_2$  stimmen beide Bedeutungen überein.

### 3.7. Parabel

Die Normalform ihrer Gleichung ist  $y = ax^2$  für  $a > 0$ . Die Menge aller  $(x, y) \in \mathbb{R}^2$ , die dieser Gleichung genügen, ist eine *Parabel* als Punktmenge. Ihr *Brennpunkt* ist  $F = (0, 1/4a)$  und ihre *Leitlinie* die Gerade  $y = -1/4a$ . Die Punkte  $(x, y)$  der Parabel

haben zum Brennpunkt und zur Leitlinie gleichen Abstand, und diese Eigenschaft charakterisiert die Parabel. Wir rechnen dazu: Der Abstand zum Brennpunkt ist  $(x^2 + (y - 1/4a)^2)^{1/2}$  und der zur Leitlinie ist  $y + 1/4a$ . Diese Abstände sind genau dann gleich, wenn  $x^2 + (y - 1/4a)^2 = (y + 1/4a)^2$  ist, und diese Gleichung ist äquivalent zu  $y = ax^2$ .

### 3.8. Ellipse

Die Normalform ihrer Gleichung ist

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad a \geq b > 0.$$

Darin ist  $2a$  die *Länge der großen Achse*,  $2e = 2\sqrt{a^2 - b^2}$  die *Exzentrizität*, und die Punkte  $F_1 = (-e, 0)$  und  $F_2 = (e, 0)$  sind die *Brennpunkte* der Ellipse. Man rechnet nach, daß die genannte Ellipse die Menge derjenigen Punkte  $(x, y)$  ist, die als Summe der Abstände von den beiden Brennpunkten den Wert  $2a$  haben. Ist  $p = (p_1, p_2)$  ein Punkt der Ellipse, so ist die Gerade mit der Gleichung

$$\frac{p_1 x}{a^2} + \frac{p_2 y}{b^2} = 1$$

die *Tangente* an die Ellipse durch  $p$ .

### 3.9. Hyperbel

Die Normalform ihrer Gleichung ist

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1, \quad a, b > 0.$$

Ihre *Exzentrizität* ist  $2e = 2\sqrt{a^2 + b^2}$ , und die *Brennpunkte* sind wieder  $F_1 = (-e, 0)$  und  $F_2 = (e, 0)$ . Die Hyperbel ist der geometrische Ort derjenigen Punkte, deren Differenz der Abstände zu den beiden Brennpunkte gleich  $\pm 2a$  ist. Die Gleichung der Hyperbel kann in der Form

$$\left(\frac{x}{a} + \frac{y}{b}\right) \left(\frac{x}{a} - \frac{y}{b}\right) = 1$$

geschrieben werden. Die Geraden

$$u = \frac{x}{a} + \frac{y}{b} = 0, \quad v = \frac{x}{a} - \frac{y}{b} = 0$$

heißen die *Asymptoten* dieser Hyperbel. Die beiden Äste der Hyperbel nähern sich den Asymptoten auf ihrer Wanderung ins Unendliche. Das sieht man am besten, wenn man die Asymptoten als Koordinatenachsen verwendet: In den Koordinaten  $u$  und  $v$  wird die Hyperbel durch die Gleichung  $uv = 1$  gegeben.

### 3.10. Strecken

Die Menge  $\{t(a, b) + (1 - t)(c, d) \mid 0 \leq t \leq 1\}$  heißt *Verbindungsstrecke* von  $(a, b)$  nach  $(c, d)$ . Eine Menge  $K \subset \mathbb{R}^2$  heißt *konvex*, wenn sie mit je zwei Punkten auch deren gesamte Verbindungsstrecke enthält. Die Menge  $\{(x, y) \mid x^2 + y^2 \leq r^2\}$  ist konvex. Sei  $L \subset \mathbb{R}^2$  eine Gerade. Das Komplement  $\mathbb{R}^2 \setminus L$  ist nicht konvex, jedoch Vereinigung zweier konvexer Mengen. Diese heißen die durch  $L$  bestimmten *Halbebenen*.

### 3.11. Lineare Abbildungen

Gleichungen der Form (3.10) lassen sich auch als Vorschrift für eine Abbildung

$$(3.12) \quad F: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (\lambda, \mu) \mapsto (a\lambda + b\mu, c\lambda + d\mu)$$

auffassen, und (3.3) heißt in diesem Fall die *Matrix der Abbildung*  $F$ . Achtung: Wir haben gegenüber (3.10)  $b$  und  $c$  vertauscht, um uns den späteren Vereinbarungen anzupassen. Die Abbildung  $F$  hat die Eigenschaft, die Vektoren  $e_1, e_2$  auf die Vektoren  $(a, c)$  und  $(b, d)$  abzubilden. Sie bildet ferner Geraden auf Geraden ab und heißt deshalb *lineare* Abbildung. Ist  $N(f)$  die Nullstellenmenge von  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ , so ist das *Bild* von  $N(f)$  bei  $F$  definitionsgemäß die Menge

$$FN(f) := \{F(\lambda, \mu) \mid (\lambda, \mu) \in N(f)\}.$$

Diese Menge wird selbst wieder durch eine Gleichung beschrieben. Wir nehmen zur Bestimmung dieser Gleichung an, daß  $F$  eine Umkehrabbildung  $G$  hat, d. h. es gilt

$$G(F(\lambda, \mu)) = (\lambda, \mu), \quad F(G(\lambda, \mu)) = (\lambda, \mu).$$

Dann gilt für  $(\lambda, \mu) \in N(f)$ :

$$0 = f(\lambda, \mu) = f(GF(\lambda, \mu)) = 0.$$

Daraus sehen wir, daß  $FN(f)$  die Nullstellenmenge der Verkettung  $f \circ G$  ist.

Die Matrix der identischen Abbildung  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $(x, y) \mapsto (x, y)$  ist die sogenannte *Einheitsmatrix*

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

Ist

$$A' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$$

die Matrix einer weiteren linearen Abbildung  $G$ , so hat die Verkettung  $F \circ G$  die Matrix

$$AA' = A \circ A' = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}.$$

Das muß man durch Einsetzen in die Formeln nachrechnen. Wir bezeichnen die neue Matrix mit  $AA'$  oder  $A \circ A'$  und nennen sie das *Produkt* oder die *Verkettung* von  $A$  und  $A'$ .

Durch eine Rechnung wie in (3.5) wird gezeigt:

**(3.13) Satz.** *Die lineare Abbildung (3.13) hat genau dann eine Umkehrabbildung, wenn  $D := ad - bc \neq 0$  ist. Die Umkehrabbildung ist eine lineare Abbildung mit der Matrix*

$$\begin{pmatrix} \frac{d}{D} & -\frac{b}{D} \\ -\frac{c}{D} & \frac{a}{D} \end{pmatrix}.$$

### 3.12. Determinante

An verschiedenen Stellen war der Wert  $ad - bc$  einer Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

bedeutsam. Dieser Wert heißt *Determinante* der Matrix  $A$  und wird mit

$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

bezeichnet. Der Absolutbetrag  $|\det(A)|$  ist der *Flächeninhalt* des von  $(a, b)$  und  $(c, d)$  aufgespannten Parallelogramms. Es gilt die Formel

$$\det(AB) = \det(A) \det(B),$$

die man durch Nachrechnen verifiziert.

### 3.13. Komplexe Zahlen

Für die gesamte Mathematik ist die Erweiterung des Bereichs der reellen Zahlen zu den komplexen Zahlen von grundlegender Bedeutung.

Keine reelle Zahl  $x$  genügt der Gleichung  $x^2 + 1 = 0$ . Man erfindet eine neue Zahl  $i$ , für die  $i^2 = -1$  gilt, bildet mit reellen Zahlen  $a$  und  $b$  Ausdrücke  $a + bi$ , addiert zwei solche gemäß der Regel

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

und multipliziert sie gemäß der Regel

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i.$$

Man rechnet also einfach wie gewohnt, aber jedesmal, wenn ein  $i^2$  vorkommt, kann man dieses durch  $-1$  ersetzen. Beispiel:  $(2 + i)(2 - i) = 2^2 - i^2 = 4 + 1 = 5$ . Man nennt  $z = a + bi$  eine *komplexe Zahl*,  $a = \operatorname{Re}(z)$  ihren *Realteil*,  $b = \operatorname{Im}(z)$  ihren *Imaginärteil* und  $i$  die *imaginäre Einheit*. Die Gesamtheit der komplexen Zahlen wird mit  $\mathbb{C}$  bezeichnet. Geometrisch wird  $z = a + bi$  als Punkt  $(a, b) \in \mathbb{R}^2$  der Zahlenebene veranschaulicht. Die Zahl  $\bar{z} = a - bi$  heißt die zu  $z = a + bi$  *konjugierte Zahl*. Es gelten die Regeln  $\overline{u + v} = \bar{u} + \bar{v}$  und  $\overline{u \cdot v} = \bar{u} \cdot \bar{v}$ . Wegen  $z \cdot \bar{z} = (a + bi)(a - bi) = a^2 + b^2$  setzen wir  $z \cdot \bar{z} = |z|^2$  und nennen  $|z|$  die *Norm* oder den *Betrag* von  $z$ . Es gelten die *Dreiecksungleichung*  $|u + v| \leq |u| + |v|$  sowie  $|uv| = |u||v|$ . Ist  $z \neq 0$ , so ist  $w = \bar{z}|z|^{-2}$  das multiplikative Inverse zu  $z$ , d. h. es gilt  $zw = 1$ . Wir haben also die Formel

$$\frac{1}{a + bi} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i.$$

Geometrisch wird die komplexe Zahl  $a + bi$  durch den Punkt  $(a, b)$  der Zahlenebene veranschaulicht. Wenn man in dieser Weise mit den Punkten der Ebene als komplexen Zahlen rechnet, spricht man auch von der *Gaußschen Zahlenebene*.

Eine wesentliche Eigenschaft der komplexen Zahlen wird durch den folgenden Satz (früher *Fundamentalsatz der Algebra* genannt) festgehalten.

**(3.14) Satz.** *Eine Gleichung der Form*

$$a_0 + a_1z + a_2z^2 + \cdots + a_nz^n = 0$$

*hat immer (mindestens) eine komplexe Zahl  $z$  als Lösung.* □

### 3.14. Polarkoordinaten

Die Lage eines Punktes  $(x, y)$  in der Ebene läßt sich durch seinen Abstand  $r$  vom Nullpunkt sowie durch den Winkel  $\varphi$  zwischen der positiven  $x$ -Achse und dem Halbstrahl von  $(0, 0)$  nach  $(x, y)$  festlegen. Man nennt  $(r, \varphi)$  die *Polarkoordinaten* des Punktes  $(x, y)$ . Der Winkel wird im Bogenmaß gemessen. Unter Verwendung der trigonometrischen Funktionen Sinus und Cosinus (siehe dazu II.6) gilt

$$(3.15) \quad x = r \cos \varphi, \quad y = r \sin \varphi.$$

Ist  $(x, y) \neq (0, 0)$ , so sind  $r$  und  $\varphi$  im Bereich  $0 \leq \varphi < 2\pi$  durch  $(x, y)$  eindeutig bestimmt.

Eine komplexe Zahl läßt sich dann in der Form

$$z = r(\cos \varphi + i \sin \varphi)$$

mit  $r = |z|$  schreiben. Nach den Additionstheoremen für Sinus und Cosinus gilt für das Produkt

$$r_1(\cos \varphi_1 + i \sin \varphi_1) \cdot r_2(\cos \varphi_2 + i \sin \varphi_2) = r_1 r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)).$$

Das liefert eine geometrische Deutung der Multiplikation.

## 2 Differential- und Integralrechnung

### 1 Funktionen

#### 1.1. Der Begriff einer Funktion

Funktionen beschreiben Vorgänge und Prozesse. Sie gehören daher zu den wichtigsten mathematischen Objekten. Den Funktionsbegriff haben wir abstrakt schon im Abschnitt über Mengensprache eingeführt. Wir wiederholen:

**(1.1) Definition.** Eine *Funktion* besteht aus drei Dingen:

- (1) Einer *Definitionsmenge*  $X$ , auch *Definitionsbereich* genannt.
- (2) Einer *Wertemenge*  $Y$ .
- (3) Einer Vorschrift, die jedem Element  $x \in X$  ein Element  $f(x) \in Y$  zuordnet.

Wir schreiben eine solche Funktion auch symbolisch in der folgenden Weise auf

$$f: X \rightarrow Y, \quad x \mapsto f(x).$$

Dabei nennen wir dann  $f(x)$  den *Wert* der Funktion an der *Stelle*  $x$ . In einem abstrakten Kontext sagt man statt Funktion auch *Abbildung* von  $X$  nach  $Y$ . Da  $x$  ein beliebiges Element aus der Menge  $X$  ist,  $x$  die Menge  $X$ , wie man sagt, „durchläuft“, so heißt  $x$  eine *Veränderliche*.  $\diamond$

Wir betrachten zunächst einmal den wichtigen Fall, daß  $X$  und  $Y$  die Menge der reellen Zahlen  $\mathbb{R}$  sind. In diesem Fall ist also nach Punkt (3) jeder reellen Zahl  $x$  eine reelle Zahl  $y = f(x)$  zuzuordnen. In einfachen und übersichtlichen Fällen wird die „Vorschrift“  $f$  eine Rechenvorschrift sein, die angibt, wie  $f(x)$  aus  $x$  ausgerechnet wird. Wir bemerken noch, daß im Fall  $f(x) = 0$  die Stelle  $x$  eine *Nullstelle* der Funktion  $f$  genannt wird. Wir geben einige einfache Beispiele.

**(1.2) Konstante Funktion.** Es sei  $c \in \mathbb{R}$  fest gegeben (eine *Konstante*). Durch  $f(x) = c$  wird die konstante Funktion mit dem Wert  $c$  definiert.  $\diamond$

**(1.3) Lineare Funktion.** Seien  $a$  und  $b$  reelle Zahlen. Eine Funktion der Form  $f(x) = ax + b$  heißt *lineare Funktion*. Die konstanten Funktionen sind der Spezialfall  $a = 0$ . (Die Funktion heißt „linear“, weil ihr Graph eine Gerade ist. Davon zu unterscheiden ist der Begriff „lineare Abbildung“, der in der linearen Algebra eingeführt wird; die linearen Abbildungen  $f: \mathbb{R} \rightarrow \mathbb{R}$  sind der Spezialfall  $b = 0$ ; im Fall  $b \neq 0$  spricht man in der linearen Algebra von „affiner Abbildung“.)  $\diamond$

**(1.4) Quadratische Funktionen.** Sind  $a \neq 0, b, c$  gegebene reelle Zahlen, so heißt eine Funktion der Form  $f(x) = ax^2 + bx + c$  eine *quadratische Funktion*. Die Umformung

$$ax^2 + bx + c = a \left( x + \frac{b}{2a} \right)^2 + \left( c - \frac{b^2}{4a^2} \right)$$

bezeichnet man als *quadratische Ergänzung*. Aus ihr erkennt man, daß im Fall  $a > 0$  der minimale Funktionswert an der Stelle  $x = -b/2a$  liegt.  $\diamond$

**(1.5) Potenzfunktionen.** Ist  $n \in \mathbb{N}$  eine natürliche Zahl, so heißt eine Funktion der Form  $f(x) = x^n$  eine *Potenzfunktion*.  $\diamond$



**(1.6) Polynome.** Die vorstehenden Beispiele sind alles Spezialfälle der Polynomfunktionen. Eine Funktion der Form

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = \sum_{k=0}^n a_kx^k$$

mit reellen Zahlen  $a_0, \dots, a_n$  heißt *Polynomfunktion*, und zwar vom *Grad*  $n$ , sofern  $a_n \neq 0$  ist<sup>4</sup>. Die Rechenvorschrift selbst nennt man auch einfach ein *Polynom*. Die Zahlen  $a_0, \dots, a_n$  heißen die *Koeffizienten* des Polynoms, und zwar  $a_j$  der Koeffizient von  $x^j$ . Durch die Funktion sind die Koeffizienten eindeutig bestimmt, das heißt gilt für alle  $x$  die Gleichung

$$\sum_{k=0}^n a_kx^k = \sum_{k=0}^n b_kx^k,$$

so folgt  $a_k = b_k$  für  $0 \leq k \leq n$ . (Diese Folgerung kann man schon dann ziehen, wenn die Gleichheit nur für  $n+1$  paarweise verschiedene Werte  $x = x_0, \dots, x_n$  besteht; zum Beweis betrachtet man das Differenzpolynom

$$d(x) = \sum_{k=0}^n (a_k - b_k)x^k;$$

es hat jedenfalls die  $n+1$  Nullstellen  $x_0, \dots, x_n$ ; ein Polynom vom Grad höchstens  $n$  kann aber höchstens  $n$  verschiedene Nullstellen haben, wie wir sogleich beim Rechnen mit Polynomen noch sehen werden.)  $\diamond$

**(1.7) Rationale Funktionen.** Eine Funktion der Form

$$f(x) = \frac{a_0 + a_1x + a_2x^2 + \cdots + a_mx^m}{b_0 + b_1x + b_2x^2 + \cdots + b_nx^n}$$

heißt *rationale Funktion*. Hier kann aber im allgemeinen nicht die gesamte Menge der reellen Zahlen als Definitionsbereich gewählt werden, da man ja durch Null nicht dividieren darf. Jede Menge  $A \subset \mathbb{R}$ , die keine Nullstelle des Nennerpolynoms enthält, ist als Definitionsbereich erlaubt.  $\diamond$

**(1.8) Rechnen mit Funktionen der Form  $X \rightarrow \mathbb{R}$ .** Man addiert oder multipliziert an jeder Stelle die Funktionswerte. Seien also

$$f: X \rightarrow \mathbb{R}, \quad g: X \rightarrow \mathbb{R}$$

Funktionen und sei  $c \in \mathbb{R}$ . Dann erklären wir neue Funktionen:

$$\begin{aligned} f + g &: x \mapsto f(x) + g(x) \\ fg = f \cdot g &: x \mapsto f(x)g(x) \\ cf &: x \mapsto cf(x) \\ \frac{f}{g} &: x \mapsto \frac{f(x)}{g(x)}. \end{aligned}$$

Die Quotientenbildung im letzten Fall ist natürlich nur erlaubt, wenn immer  $g(x) \neq 0$  ist. Dieser Fall ist uns schon bei der Definition der rationalen Funktionen begegnet. Sinngemäß nennen wir  $f + g$  die *Summe* und  $fg$  das *Produkt* der Funktionen  $f$  und  $g$ .  $\diamond$

<sup>4</sup>Dem Nullpolynom ordnet man manchmal den Grad  $-\infty$  zu.

## 1.2. Das Rechnen mit Polynomen

Durch wiederholte Anwendung der Rechenregeln (1.8) wird offenbar jede rationale Funktion aus konstanten und linearen Funktionen erhalten. Summe und Produkt von Polynomen ist wieder ein Polynom. Für die Addition sollte das unmittelbar klar sein. Für die Multiplikation benutzen wir die Formel I(1.2).

$$\left( \sum_{j=0}^m a_j x^j \right) \left( \sum_{k=0}^n b_k x^k \right) = \sum_{i=0}^{m+n} \left( \sum_{l=0}^i a_l b_{i-l} \right) x^i.$$

**(1.9) Definition.** Für Polynome gibt es ebenso wie für Zahlen eine *Division mit Rest*. Das soll bedeuten: Sind  $f(x)$  und  $g(x)$  Polynome, so gibt es weitere Polynome  $q(x)$  und  $r(x)$ , so daß die Gleichung

$$f(x) = q(x)g(x) + r(x)$$

besteht, worin der *Rest*  $r(x)$  gleich Null ist oder einen Grad hat, der kleiner als der von  $g(x)$  ist.  $\diamond$

Natürlich kann die Division „aufgehen“, das heißt  $r(x)$  kann das Nullpolynom sein; diesem sollte man deshalb einen Grad kleiner als Null zuordnen. Ist  $r(x) = 0$ , so sagt man, daß  $g(x)$  das Polynom  $f(x)$  *teilt*. Hat  $g$  einen größeren Grad als  $f$ , so ist die Definition mit  $q = 0$  und  $r = f$  erfüllt. Wir zeigen an einem Beispiel, wie man die Division mit Rest praktisch durchführt. Sei  $f(x) = x^4 - x^2 + 1$  und  $g(x) = x^2 - 2x$ . Durch die Subtraktion  $f(x) - x^2g(x)$  „beseitigt“ man die vierte Potenz. Das Resultat ist  $2x^3 - x^2 + 1$ . Davon subtrahiert man  $2xg(x)$  und beseitigt damit die dritte Potenz. Das Resultat ist  $3x^2 + 1$ . Davon subtrahiert man  $3g(x)$  und beseitigt damit die zweite Potenz. Das Resultat ist  $6x + 1$ . Also ergibt sich

$$x^4 - x^2 + 1 = (x^2 + 2x + 3)(x^2 - 2x) + (6x + 1),$$

wovon man sich auch durch eine „Probe“ überzeugen mag.

Dividiert man ein Polynom  $f(x)$  durch  $x - a$ , so ist der Rest eine Konstante. Ist  $a$  eine Nullstelle, so folgt durch Einsetzen von  $a$ , daß diese Konstante Null ist, das heißt:

**(1.10) Notiz.** *Ist  $a$  eine Nullstelle des Polynoms  $f(x)$ , so ist  $x - a$  ein Teiler von  $f(x)$ . Es gibt also eine Faktorisierung  $f(x) = (x - a)g(x)$ .*  $\square$

**(1.11) Beispiel.**  $x^n - a^n = (x - a) \left( \sum_{j=1}^{n-1} x^j a^{n-j-1} \right)$ .  $\diamond$

In der letzten Notiz kann man eventuell von  $g(x)$  einen weiteren Faktor der Form  $x - b$  abspalten. Wir sagen, ein Polynom *zerfällt in Linearfaktoren*, wenn es die Form

$$f(x) = c(x - a_1)(x - a_2) \cdots (x - a_n)$$

hat. Dabei wird nicht vorausgesetzt, daß die  $a_j$  paarweise verschieden sind. Kommt etwa der Wert  $a_1$  genau  $k$ -mal unter den  $a_j$  vor, so sagt man,  $a_1$  ist eine *k-fache* Nullstelle von  $f(x)$ . Wenn man das Produkt der Linearfaktoren ausmultipliziert, erkennt man, wie sich die Koeffizienten aus den Nullstellen (= Wurzeln) ausrechnen lassen. Ein Polynom vom Grad  $n$  hat also höchstens  $n$  verschiedene Nullstellen. Wenn man

auch komplexe Zahlen als Nullstellen zuläßt, so hat jedes Polynom eine Nullstelle und zerfällt deshalb in Linearfaktoren. Zum Beispiel ist dann  $x^2 + 1 = (x + i)(x - i)$ .

Summe und Produkt rationaler Funktionen ist wieder eine rationale Funktion. Das folgt, weil man mit den rationalen Funktionen nach den Regeln der „Bruchrechnung“ rechnet, also etwa: Erweitern, Kürzen, auf den Hauptnenner bringen. Beispiel: Es gilt

$$\frac{1}{1-x} + \frac{1}{1+x} = \frac{2}{1-x^2},$$

denn wegen  $1-x^2 = (1-x)(1+x)$  ist dieses der Hauptnenner, auf den man die Brüche links durch Erweiterung bringt. Haben Zähler und Nenner einer rationalen Funktion die Nullstelle  $a$ , so kann man das Zähler- und das Nennerpolynom durch  $x - a$  teilen und diesen Faktor kürzen. Dadurch kann man eventuell die zunächst nicht zugelassene Stelle  $x = a$  zum Definitionsbereich hinzufügen. Ist  $a$  eine Nullstelle des Nenners aber nicht des Zählers, so nennt man  $a$  eine *Polstelle* oder kurz einen *Pol* der rationalen Funktion.

### 1.3. Verkettung

Ein wichtiger Prozeß, aus gegebenen Funktionen neue herzustellen, ist die Verkettung von Funktionen: Der Funktionswert einer Funktion wird in eine andere Funktion „eingesetzt“. Genauer und formaler: Seien  $f: X \rightarrow Y$  und  $g: Y \rightarrow Z$  Funktionen. Die *Verkettung*  $g \circ f$  (gelesen:  $g$  nach  $f$ ,  $g$  verkettet mit  $f$ ) ist die Funktion

$$g \circ f: X \rightarrow Z, \quad x \mapsto g(f(x)).$$

Ist etwa  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto x - 2$  und  $g: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto x^2 + 3$ , so ist

$$g \circ f: \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto (x - 2)^2 + 3 = x^2 - 4x + 7.$$

### 1.4. Intervalle

Nicht immer ist die gesamte Menge der reellen Zahlen ein zweckmäßiger Definitionsbereich. Die wichtigsten Definitionsbereiche sind die Intervalle. Ein *Intervall* ist eine Teilmenge der reellen Zahlen, die mit je zwei Zahlen auch alle dazwischenliegenden enthält. Für Intervalle verwenden wir die folgenden Bezeichnungen:

$$\begin{aligned} [a, b] &= \{x \mid a \leq x \leq b\} \\ ]a, b[ &= \{x \mid a < x < b\} \\ [a, b[ &= \{x \mid a \leq x < b\} \\ ]a, b] &= \{x \mid a < x \leq b\} \\ [a, \infty[ &= \{x \mid a \leq x\} \\ ]a, \infty[ &= \{x \mid a < x\} \\ ]-\infty, b] &= \{x \mid x \leq b\} \\ ]-\infty, b[ &= \{x \mid x < b\} \\ ]-\infty, \infty[ &= \mathbb{R}. \end{aligned}$$

Ein Intervall der Form  $[a, b]$  heißt *abgeschlossenes* Intervall von  $a$  nach  $b$ , eines der Form  $]a, b[$  ein *offenes* Intervall von  $a$  nach  $b$ . Im letzteren Fall darf für  $b$  auch  $\infty$  und für  $a$  auch  $-\infty$  stehen. Der Fall  $a = b$  ist nicht besonders interessant und werde stillschweigend ausgeschlossen.

### 1.5. Weitere Funktionen

Es gibt natürlich auch ganz andere Funktionen als die oben beispielhaft genannten. So ist eine Folge  $(a_n)$  nichts anderes als eine Funktion  $a: \mathbb{N} \rightarrow \mathbb{R}, n \mapsto a_n$ ; hier wird der Funktionswert an der Stelle  $n$  eben anders, nämlich durch  $a_n$ , bezeichnet. Auch werden Funktionen nicht immer durch Rechenvorschriften (sogenannte Formeln) beschrieben. Die Zuordnungsvorschrift darf ja völlig beliebig sein. Zum Beispiel kann eine Funktion  $[0, 2] \rightarrow \mathbb{R}$  auf dem Teilstück  $]0, 1]$  durch *eine* Formel, auf dem Teilstück  $[1, 2]$  durch eine *andere* Formel erklärt werden.

Eine Funktion der Form  $f: \mathbb{R}^2 = \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (x, y) \mapsto f(x, y)$  bezeichnet man als eine Funktion in zwei Veränderlichen.

### 1.6. Graph einer Funktion

Um den Verlauf einer Funktion

$$f: A \rightarrow \mathbb{R}, \quad A \subset \mathbb{R}$$

zu veranschaulichen, wird der *Graph*  $G(f)$  der Funktion  $f$  verwendet. Er ist als die Teilmenge

$$G(f) = \{(x, f(x)) \mid x \in A\}$$

der Zahlenebene  $\mathbb{R}^2$  definiert.

**(1.12) Beispiele.** Der Graph einer konstanten Funktion ist eine Parallele zur  $x$ -Achse. Der Graph einer linearen Funktion ist eine Gerade. Jede Gerade, die nicht parallel zur  $y$ -Achse ist, ist Graph einer linearen Funktion  $f(x) = ax + b$ ; die Zahl  $a$  heißt in diesem Zusammenhang *Steigung* der Geraden. Der Graph einer quadratischen Funktion ist eine Parabel. Der Graph der Funktion  $f(x) = |x|$  besteht aus zwei Halbgeraden, die sich im Nullpunkt treffen und dort einen rechten Winkel bilden. Wenn man sich einer Polstelle einer rationalen Funktion von einer Seite her nähert, so läuft der Funktionswert ins positive oder negative „Unendliche“, etwa bei der Funktion  $f(x) = x^{-1}$ .  $\diamond$

Funktionen sind nicht immer durch Rechenvorschriften erklärt (oder erklärbar!). Ein typisches Beispiel dafür ist die Wertetabelle einer Meßreihe in einem Versuch, die manchmal in der Tabellenform

$$\begin{array}{c|c|c|c|c} t_1 & t_2 & t_3 & \dots & t_n \\ \hline a_1 & a_2 & a_3 & \dots & a_n \end{array}$$

notiert wird, wenn etwa  $a_k$  das Meßergebnis zur Zeit  $t_k$  ist. Diese Meßwerte werden dann graphisch aufgetragen, indem man in der Zahlenebene die Punkte  $(t_k, a_k)$  notiert. Die Vorstellung eines deterministischen Ablaufs eines Vorganges suggeriert dann den Wunsch, eine *Funktion*  $a$  zu finden, deren Graph durch die gegebenen Punkte  $(t_k, a_k)$  hindurchläuft, d. h. es soll gelten  $a_k = a(t_k)$ . Dieser Wunsch ist vielleicht nicht sehr zweckmäßig, weil man ja mit Meßfehlern rechnen muß. Wir wollen ihm aber trotzdem mathematisch etwas nachgehen. (Siehe dazu auch den Abschnitt III.3 über die Methode der kleinsten Quadrate.)

### 1.7. Interpolation

Eine Kurve durch gegebene Punkte hindurchzulegen, bezeichnet man als *Interpolation*. Wir werden uns den folgenden Satz überlegen, der die bekannte Tatsache verallgemeinert, daß durch zwei Punkte genau eine Gerade gelegt werden kann.

**(1.13) Satz.** *Sind  $n$  Punkte  $(t_1, a_1), \dots, (t_n, a_n)$  in der Ebene  $\mathbb{R}^2$  mit paarweise verschiedenen Zahlen  $t_k$  gegeben, so gibt es ein Polynom höchstens vom Grad  $n-1$ , dessen Graph durch diese Punkte läuft.*

BEWEIS. Wir bezeichnen mit

$$p_j(t) = \prod_{k, k \neq j} (t - t_k)$$

das Produkt der Faktoren  $t - t_1, t - t_2, \dots, t - t_n$  ohne den Faktor  $t - t_j$ . Da insgesamt  $n-1$  Faktoren ausmultipliziert werden, so ist  $p_j(t)$  ein Polynom in der Veränderlichen  $t$  vom Grad  $n-1$ . Nach Konstruktion und Voraussetzung über die  $t_k$  hat dieses Polynom die folgende Eigenschaft:

$$p_j(t_k) = 0 \quad \text{für} \quad k \neq j, \quad p_j(t_j) \neq 0.$$

Daraus folgt, daß die Summe

$$\sum_{j=1}^n a_j \frac{p_j(t)}{p_j(t_j)} = p(t)$$

ein Polynom höchstens vom Grad  $n-1$  ist, das an der Stelle  $t_k$  den Wert  $a_k$  annimmt.  $\square$

### 1.8. Die Umkehrfunktion

**(1.14) Definition.** Sei  $f: X \rightarrow Y$  eine Funktion. Eine *Umkehrfunktion* von  $f$  ist eine Funktion  $g: Y \rightarrow X$  (in der „umgekehrten Richtung“), so daß für alle  $x \in X$  und alle  $y \in Y$  die Gleichungen

$$(1.15) \quad g(f(x)) = x, \quad f(g(y)) = y$$

gelten.  $\diamond$

Bei der Untersuchung von Umkehrfunktionen ist es oft zweckmäßig, Definitionsbereich und Wertebereich geeignet einzugrenzen.

Ist  $g$  Umkehrfunktion von  $f$ , so ergibt sich der Graph  $G(g)$  aus  $G(f)$  durch Spiegelung an der Geraden  $D = \{(x, x) \mid x \in \mathbb{R}\}$ , also an der Winkelhalbierenden der positiven  $x$ -Achse und  $y$ -Achse.

Manche Funktionen werden dadurch erklärt, daß sie die Umkehrfunktion einer gegebenen Funktion  $f$  sein sollen. Deshalb ist nützlich zu wissen, wann denn  $f$  eine Umkehrfunktion hat.

**(1.16) Notiz.** *Eine Funktion  $f: X \rightarrow Y$  hat genau dann eine Umkehrfunktion, wenn sie die folgenden beiden Eigenschaften hat:*

- (1) *Aus  $x_1 \neq x_2$  folgt immer  $f(x_1) \neq f(x_2)$ . (Eine solche Funktion nennt man injektiv.)*

- (2) Zu jedem  $y \in Y$  gibt es ein  $x \in X$ , so daß  $y = f(x)$  ist, d. h. jedes  $y \in Y$  ist ein Funktionswert. (Eine solche Funktion nennt man surjektiv.)

Eine Funktion heißt bijektiv, wenn sie sowohl surjektiv als auch injektiv ist.  $\square$

In diesem Kontext benutzen wir die folgende Terminologie: Die Menge der Funktionswerte einer Funktion  $f$  heie das *Bild* von  $f$ . Allgemeiner: Ist  $f: X \rightarrow Y$  eine Funktion und  $A \subset X$ , so ist

$$f(A) = \{f(a) \mid a \in A\}$$

das *Bild* der Teilmenge  $A$  bei  $f$ ; demnach ist  $f(X)$  das Bild von  $f$ .

**(1.17) Beispiel.** Die Funktion

$$f: [0, \infty[ \rightarrow [0, \infty[, \quad x \mapsto x^2$$

hat die Umkehrfunktion

$$g: [0, \infty[ \rightarrow [0, \infty[, \quad x \mapsto \sqrt{x}.$$

Die *Quadratwurzel*  $\sqrt{x}$  von  $x$  wird in diesem Zusammenhang als Umkehrfunktion von  $f$  definiert. Es gibt offenbar hchstens eine Zahl  $b \geq 0$  mit der Eigenschaft  $b^2 = x$ , weil aus  $0 \leq b < c$  die Ungleichung  $b^2 < c^2$  folgt, das heit  $f$  ist injektiv. Wegen der Surjektivitt siehe (2.8). In derselben Weise definiert man fr jede natrliche Zahl  $n$  die Funktion *n-te Wurzel* als Umkehrfunktion von  $\mathbb{R}_+ \rightarrow \mathbb{R}_+$ ,  $x \mapsto x^n$ .  $\diamond$

**(1.18) Beispiel.** Die Funktion  $f: \mathbb{R} \rightarrow [1, \infty[$ ,  $x \mapsto (x+1)^4$  hat keine Umkehrfunktion. Sie ist surjektiv, wegen  $f(1) = f(-3)$  aber nicht injektiv.  $\diamond$

**(1.19) Beispiel.** Sei  $A$  die Menge der reellen Zahlen ohne die  $-1$ . Dann ist

$$f: A \rightarrow \mathbb{R}, \quad f(x) = \frac{1-x}{1+x}$$

definiert, weil der Nenner auf  $A$  ungleich Null ist. Sei  $f(x) = y$ . Wir multiplizieren mit  $1+x$  und erhalten

$$y(1+x) = 1-x.$$

Ist  $y = -1$ , so fhrt das zum Widerspruch  $-1 = 1$ ; also ist  $-1$  kein Funktionswert. Ist  $y \neq -1$ , so lsen wir die Gleichung nach  $x$  auf

$$x = \frac{1-y}{1+y},$$

das heit  $y$  ist dann ein Funktionswert. Diese Rechnung, etwas anders gewendet, liefert: Die Funktion

$$f: A \rightarrow A, \quad x \mapsto \frac{1-x}{1+x}$$

ist gleich ihrer Umkehrfunktion.  $\diamond$

## 2 Stetige Funktionen

Wir haben eine Funktion als eine Zuordnungsvorschrift definiert. Damit ist zwar das Wort Funktion erklrt, aber der Willkr sind Tr und Tor geffnet. Etwas theoretische und praktische Eingrenzung tut not. Wir sammeln in diesem Abschnitt grundstzliche theoretische Einsichten. Fr die Praxis mag man ihn getrost berspringen.

## 2.1. Stetigkeit

Eine stetige Funktion macht keine Sprünge, der Graph ist eine zusammenhängende Linie ohne Lücken. Dieser Spruch liefert eine brauchbare vorläufige Anschauung, ist aber keine mathematische Definition. Eine exakte Definition beruht auf dem Grenzwertbegriff.

**(2.1) Definition.** Eine Funktion  $f: A \rightarrow \mathbb{R}$  mit einem Definitionsbereich  $A \subset \mathbb{R}$  heißt *stetig im Punkt*  $a \in A$ , wenn zu jedem  $\varepsilon > 0$  ein  $\delta > 0$  existiert, so daß aus  $|x - a| < \delta, x \in A$  die Ungleichung  $|f(x) - f(a)| < \varepsilon$  folgt. Sie heißt *stetig*, wenn sie in jedem Punkt ihres Definitionsbereichs stetig ist.  $\diamond$

**(2.2) Beispiel.** Die Funktion  $f(x) = x$  ist stetig, denn mit  $\varepsilon = \delta$  ist die obige Definition in jedem Punkt  $a$  erfüllt. Aus demselben Grund, etwa, ist eine konstante Funktion stetig.  $\diamond$

In Worten umschrieben besagt die Definition (2.1) der Stetigkeit: Sei eine Fehler-schranke  $\varepsilon > 0$  vorgegeben. Dann kann dazu eine Grenze  $\delta > 0$  gefunden werden, so daß für alle  $x$ , die von  $a$  weniger als  $\delta$  abweichen, der Funktionswert  $f(x)$  höchstens um  $\varepsilon$  von  $f(a)$  abweicht.

Man kann die Stetigkeit auch mittels Konvergenz von Folgen definieren und zeigen, daß die folgende Definition zu (2.1) äquivalent ist. Siehe dazu den nächsten Abschnitt über Grenzwerte von Folgen.

**(2.3) Definition.** Sei  $f: J \rightarrow \mathbb{R}$  eine auf einem Intervall  $J$  definierte Funktion. Sie ist genau dann im Punkt  $a \in J$  stetig, wenn aus  $\lim_{n \rightarrow \infty} a_n = a$  immer  $\lim_{n \rightarrow \infty} f(a_n) = f(a)$  folgt. („Wenn  $x$  dem Wert  $a$  zustrebt, so auch  $f(x)$  dem Wert  $f(a)$ .“)  $\diamond$

Aus gegebenen stetigen Funktionen kann man sich leicht neue verschaffen. Mittels Definition (2.3) ist der folgende Satz eine direkte Folge aus den Grenzwertregeln (3.9).

**(2.4) Satz.** Seien  $f, g: A \rightarrow \mathbb{R}$  stetig. Dann sind auch die folgenden Funktionen stetig:  $f + g, f \cdot g, f/g$ ; im letzten Fall hat man, wie üblich,  $g(x) \neq 0$  vorauszusetzen.  $\square$

Aus diesem Satz und dem Beispiel der konstanten und linearen Funktionen folgt:

**(2.5) Satz.** Jede rationale Funktion ist stetig.  $\square$

**(2.6) Satz.** Die Verkettung stetiger Funktionen ist wieder stetig.  $\square$

## 2.2. Die Hauptsätze

Der nächste *Zwischenwertsatz* präzisiert, was mit „macht keine Sprünge“ gemeint ist.

**(2.7) Zwischenwertsatz.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  stetig. Dann nimmt  $f$  jeden Wert zwischen  $f(a)$  und  $f(b)$  an, das heißt ist etwa  $f(a) \leq c \leq f(b)$ , so gibt es ein  $x \in [a, b]$ , dessen Funktionswert  $f(x) = c$  ist.  $\square$

Ist insbesondere  $f(a) < 0$  und  $f(b) > 0$ , so gibt es zwischen  $a$  und  $b$  eine Nullstelle von  $f$ . Es ist leicht, eine Funktion auf einem Intervall anzugeben, die einen Wert überspringt; sie ist dann nicht stetig. Der Zwischenwertsatz wird offenbar falsch, wenn man nur auf eine einzige reelle Zahl verzichtet. Der Zwischenwertsatz ist ein theoretischer Satz, man weiß, es gibt eine Stelle, aber man kennt sie nicht ohne weiteres. Wir erläutern diesen Satz am Beispiel der Quadratwurzel.

**(2.8) Beispiel.** Die Funktion  $f: [0, \infty[ \rightarrow [0, \infty[$ ,  $x \mapsto x^2$  ist stetig. Es ist  $f(0) = 0$ . Sei  $c \geq 0$  gegeben. Es gibt sicherlich eine natürliche Zahl  $n$ , deren Quadrat  $n^2$  größer als  $c$  ist. Nach dem Zwischenwertsatz gibt es also eine Zahl  $b$  mit  $b^2 = c$ , das heißt es gibt die Quadratwurzel von  $c$ . Aber damit kennen wir sie natürlich noch nicht.  $\diamond$

**(2.9) Minimaxsatz.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  stetig. Dann gibt es eine Stelle  $x_m \in [a, b]$  mit minimalem und eine Stelle  $x_M \in [a, b]$  mit maximalem Funktionswert, mit anderen Worten, für alle  $x$  gilt  $f(x_m) \leq f(x) \leq f(x_M)$ . In diesem Fall heißt  $f(x_m)$  das Minimum und  $f(x_M)$  das Maximum der Funktion  $f$ .  $\square$

Ebenso wie der Zwischenwertsatz ist der Minimaxsatz ein rein theoretischer Satz und gibt für die Anwendungen der Mathematik wenig her. Aber solche Sätze beruhigen, und sie präzisieren, warum und wann gilt, was der Anschauung nach eigentlich gelten sollte. Die Funktion  $f: ]0, 1] \rightarrow \mathbb{R}$ ,  $f(x) = x^{-1}$  ist stetig; sie hat aber beliebig große Funktionswerte (bei der Annäherung  $x \rightarrow 0$ ), also kein Maximum. Im letzten Satz ist es deshalb wichtig, das Definitionsintervall als abgeschlossen vorauszusetzen. Benutzt man im Minimaxsatz noch den Zwischenwertsatz, so folgt, daß  $[f(x_m), f(x_M)]$  die Menge der Funktionswerte von  $f$  ist. Allgemein bildet eine stetige Funktion Intervalle wieder auf Intervalle ab.

Sei  $X \subset \mathbb{R}$ . Eine Funktion  $f: X \rightarrow \mathbb{R}$  heißt *streng monoton wachsend* (bzw. *fallend*), wenn aus  $a < b$  immer  $f(a) < f(b)$  (bzw.  $f(a) > f(b)$ ) folgt. Eine streng monoton wachsende oder fallende Funktion ist also injektiv. Wählen wir für eine solche Funktion  $Y$  als die Menge ihrer Funktionswerte, so hat  $f: X \rightarrow Y$  eine Umkehrfunktion. Hat eine streng monoton wachsende Funktion  $f$  eine Umkehrfunktion  $g$ , so ist auch  $g$  streng monoton wachsend.

**(2.10) Beispiel.** Die  $n$ -te Potenz  $f(x) = x^n$  ist auf  $[0, \infty[$  streng monoton wachsend, also auch ihre Umkehrfunktion, die  $n$ -te Wurzel.  $\diamond$

**(2.11) Satz.** Seien  $A$  und  $B$  Intervalle. Die Funktion  $f: A \rightarrow B$  sei stetig und habe eine Umkehrfunktion  $g: B \rightarrow A$ . Dann ist auch  $g$  stetig. Ferner sind  $f$  und  $g$  streng monoton.  $\square$

Die strenge Monotonie im letzten Satz ergibt sich mit Hilfe des Zwischenwertsatzes.

Stetige Funktionen können sehr kompliziert sein. Sie sind aber die einzigen Funktionen, die sich beliebig genau durch formelmäßig berechenbare Funktionen annähern lassen. Es gilt nämlich der folgende *Approximationssatz von Weierstraß*.

**(2.12) Satz.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  stetig. Zu jedem  $\varepsilon > 0$  gibt es ein Polynom  $p$ , so daß für alle  $x \in [a, b]$  die Ungleichung  $|f(x) - p(x)| < \varepsilon$  gilt.  $\square$

In der  $(\varepsilon, \delta)$ -Definition (2.1) einer stetigen Funktion kann man bei gegebenem  $\varepsilon$  im allgemeinen nicht ein festes  $\delta$  für alle Stellen  $a$  wählen. Ein Funktion  $f: A \rightarrow \mathbb{R}$  heißt *gleichmäßig stetig*, wenn zu jedem  $\varepsilon > 0$  ein  $\delta > 0$  so existiert, daß für  $|b - a| < \delta$  immer  $|f(b) - f(a)| < \varepsilon$  gilt.

**(2.13) Satz.** Eine stetige Funktion  $f: [a, b] \rightarrow \mathbb{R}$  auf einem abgeschlossenen Intervall ist gleichmäßig stetig.  $\square$



### 2.3. Grenzwerte

Es kommt häufig vor, daß man eine Funktion  $f: J \setminus \{a\} \rightarrow \mathbb{R}$  zu untersuchen hat, die auf dem Intervall  $J$  ohne den Punkt  $a \in J$  definiert ist (zum Beispiel bei der Untersuchung des Differenzenquotienten im vierten Abschnitt). Die Frage ist dann, ob die Funktionswerte  $f(x)$  einen Grenzwert haben, wenn  $x$  sich der Stelle  $a$  nähert. Wir sagen dann:

**(2.14) Definition.** Es gilt

$$\lim_{x \rightarrow a} f(x) = A,$$

gelesen: Limes  $f(x)$  für  $x$  gegen  $a$  ist  $A$ , wenn zu jedem  $\varepsilon > 0$  ein  $\delta > 0$  so existiert, daß für  $0 < |x - a| < \delta$  immer  $|f(x) - A| < \varepsilon$  gilt.  $\diamond$

Die Aussage der letzten Definition gilt, wenn die durch die Festsetzung  $f(a) = A$  auf ganz  $J$  definierte Funktion im Punkt  $a$  stetig ist.

Von dieser Definition gibt es Varianten für die Bewegung  $x \rightarrow \pm\infty$ .

**(2.15) Definition.** Sei  $f: [a, \infty[ \rightarrow \mathbb{R}$  gegeben. Es gilt

$$\lim_{x \rightarrow \infty} f(x) = A,$$

wenn zu jedem  $\varepsilon > 0$  ein  $M$  existiert, so daß für alle  $x > M$  die Ungleichung  $|f(x) - A| < \varepsilon$  gilt. Sei  $f: ]-\infty, a] \rightarrow \mathbb{R}$  gegeben. Es gilt

$$\lim_{x \rightarrow -\infty} f(x) = A,$$

wenn zu jedem  $\varepsilon > 0$  ein  $M$  existiert, so daß für alle  $x < -M$  die Ungleichung  $|f(x) - A| < \varepsilon$  gilt.  $\diamond$

Schließlich kann man auch noch Aussagen wie

$$\lim_{x \rightarrow a} f(x) = \pm\infty, \quad \lim_{x \rightarrow \pm\infty} f(x) = \pm\infty$$

erklären. Das geschieht in Analogie zum obigen, siehe dazu den Abschnitt über Folgen.

## 3 Folgen. Reihen. Grenzwerte

### 3.1. Folgen

**(3.1) Definition.** Eine *Zahlenfolge* ist eine Vorschrift, die jeder natürlichen Zahl  $n$  eine reelle Zahl  $a_n$  zuordnet. Wir schreiben eine solche Folge in der Form  $(a_n \mid n \in \mathbb{N})$  oder kurz  $(a_n)$  auf, oder auch  $a_1, a_2, a_3, \dots$ , wobei die drei Punkte andeuten, daß man weiß, wie es weitergeht. Wir nennen  $a_n$  das *n-te Folgenglied*. Gelegentlich ist es zweckmäßig, die Numerierung der Folgenglieder an einer anderen Stelle beginnen zu lassen, etwa  $a_0, a_1, a_2, \dots$  oder  $a_5, a_6, a_7, \dots$  (Formal definiert: Eine Zahlenfolge ist eine Abbildung  $\mathbb{N} \rightarrow \mathbb{R}$ ; so kann man dann Folgen  $\mathbb{N} \rightarrow M$  mit Werten in einer beliebigen Menge  $M$  definieren.)  $\diamond$

**(3.2) Beispiel.** Anfangsglieder und  $n$ -tes Glied von Folgen.

$$\begin{aligned} 1, \frac{1}{2}, \frac{1}{3}, \dots & ; a_n = \frac{1}{n} \\ 0, 0, 0, \dots & ; a_n = 0 \\ -1, 1, -1, 1, \dots & ; a_n = (-1)^n \\ 2, 4, 8, 16, \dots & ; a_n = 2^n \\ 1, 4, 7, 10, \dots & ; a_n = 3n - 2. \end{aligned}$$

Aus einigen Anfangswerten einer Folge das allgemeine Glied, das sogenannte Bildungsgesetz, zu erraten, ist eine beliebte Denksportaufgabe (Intelligenztest)<sup>5</sup>.  $\diamond$

Wir nennen vier Gründe, warum Folgen gebraucht werden.

- (1) Schrittweise Annäherung einer Zahl. Der unendliche periodische Dezimalbruch  $0,99999\dots$  ist eine symbolische Abkürzung für die Folge

$$0,9 \quad 0,99 \quad 0,999 \quad \dots$$

von Annäherungen an die Zahl 1.

- (2) Ergebniswerte von aufeinanderfolgenden Prozessen oder Messungen.  
 (3) Durchnummerierung der Elemente von geeigneten Mengen mit unendlich vielen Elementen.  
 (4) Studium des Langzeitverhaltens von Prozessen (Grenzwerte).

**(3.3) Geometrische Folge.** Seien  $a$  und  $q$  reelle Zahlen. Ein Folge der Form

$$a, aq, aq^2, aq^3, \dots$$

mit dem allgemeinen Glied  $a_n = aq^{n-1}$  heißt *geometrische Folge* mit *Anfangswert*  $a$  und dem *Quotienten*  $q$ . Stellen wir uns die Folge als Prozeß vor, so wird in jedem Schritt der augenblickliche Wert mit  $q$  multipliziert. Es gilt  $a_{n+1}/a_n = q$ , falls  $a$  und  $q$  nicht Null sind.  $\diamond$

### 3.2. Konvergenz. Grenzwert

Die schrittweise, immer genauere Annäherung einer Zahl ist ein Grenzprozeß. Ein solcher wird wie folgt definiert.

**(3.4) Definition.** Ein Folge  $(a_n \mid n \in \mathbb{N})$  *konvergiert* zum *Grenzwert*  $a$ , wenn zu jeder Zahl  $\varepsilon > 0$  eine Zahl  $N \in \mathbb{N}$  existiert, so daß für alle  $n \geq N$  die Ungleichung  $|a - a_n| < \varepsilon$  gilt. Wir sagen dann auch,  $a$  ist der *Limes* der Folge. Ist der Limes Null, so spricht man von einer *Nullfolge*. Eine Folge heißt *konvergent*, wenn sie einen Grenzwert hat, andernfalls *divergent*. Die symbolische Gleichung

$$\lim_{n \rightarrow \infty} a_n = a$$

(gelesen: Limes  $a_n$  für  $n$  gegen Unendlich ist gleich  $a$ ) ist eine Abkürzung für den Satz: Die Folge  $(a_n)$  konvergiert zum Grenzwert  $a$ . Ist das der Fall, so verwenden wir das Symbol  $\lim_{n \rightarrow \infty} a_n$  auch als Bezeichnung für die Zahl  $a$ . Wir lassen das Symbol  $n \rightarrow \infty$  am lim-Zeichen auch manchmal weg.

<sup>5</sup>Zum Beispiel: Wie geht die Buchstabenfolge O,T,T,F,F,S,S,E,N,T,E,... bis ins Unendliche weiter?

Wir führen ferner noch folgende Symbolik ein:  $\lim_{n \rightarrow \infty} a_n = \infty$  (bzw.  $\lim_{n \rightarrow \infty} a_n = -\infty$ ) bedeutet: Zu jedem  $A \in \mathbb{R}$  gibt es ein  $N$ , so daß für alle  $n \geq N$  die Ungleichung  $a_n > A$  (bzw.  $a_n < A$ ) gilt. Mit den Zeichen  $\pm\infty$  kann man allerdings nicht rechnen; es ist keine Zahl; insbesondere kann man nicht die Regeln (3.9) direkt anwenden.  $\diamond$

Wir erläutern und umschreiben die Struktur der Grenzwertdefinition. Es ist  $\varepsilon > 0$  eine Fehlerschranke, die man sich als sehr kleine Zahl vorstellen soll. Ist diese Schranke einmal vorgegeben, so muß es im Fall der Konvergenz möglich sein, eine Stelle  $N$  anzugeben, von der an die Folgenglieder von  $a$  niemals eine größere Abweichung als  $\varepsilon$  haben. Wird  $\varepsilon$  verkleinert, so muß im allgemeinen die zugehörige kritische Stelle  $N$  vergrößert werden.

**(3.5) Bemerkung.** Endlich viele Glieder am Anfang der Folge spielen für die Konvergenzuntersuchung keine Rolle, d. h.  $(a_n \mid n \in \mathbb{N})$  konvergiert genau dann, wenn  $a_k, a_{k+1}, a_{k+2}, \dots$  konvergiert, und der Grenzwert ist in beiden Fällen derselbe. Diese Bemerkung wird stillschweigend verwendet.  $\diamond$

Die logische Verneinung der Grenzwertdefinition lautet:  $a$  ist nicht Grenzwert der Folge  $(a_n)$ , wenn es eine Zahl  $\varepsilon > 0$  so gibt, daß zu jeder natürlichen Zahl  $N$  mindestens ein Folgenglied mit einer Nummer  $n \geq N$  die Ungleichung  $|a - a_n| \geq \varepsilon$  erfüllt. Eine Folge divergiert, wenn die Aussage des letzten Satzes für jede Zahl  $a$  richtig ist.

Gilt eine Aussage über die Glieder einer Folge für alle bis auf endlich viele, so sagt man auch, die Aussage gelte für *fast alle* Folgenglieder. Die Menge  $U(\varepsilon) = \{x \mid |x - a| < \varepsilon\}$  heißt  $\varepsilon$ -Umgebung von  $a$ . Mit dieser Sprechweise können wir anschaulich und salopp sagen, eine Folge konvergiert zum Grenzwert  $a$ , wenn jede  $\varepsilon$ -Umgebung von  $a$  fast alle Folgenglieder enthält.

**(3.6) Beispiel.** Die Folge  $a_n = \frac{1}{n}$  ist eine Nullfolge.  $\diamond$

**(3.7) Beispiel.** Die Folge  $a_n = (-1)^n$  divergiert.  $\diamond$

**(3.8) Beispiel.** Die konstante Folge  $a_n = a$  hat den Limes  $a$ .  $\diamond$

Die Grenzwertdefinition ist sprachlich schwerfällig und logisch kompliziert. Sie leidet überdies an dem Mangel, daß zu ihrer Anwendung der Grenzwert schon bekannt sein muß, was gerade meist nicht der Fall ist. Wir wenden uns deshalb Regeln zu, mit denen man Konvergenz feststellen und gegebenenfalls den Grenzwert berechnen kann.

**(3.9) Rechenregeln für Grenzwerte.** Seien  $(a_n)$  bzw.  $(b_n)$  konvergente Folgen mit den Grenzwerten  $a$  bzw.  $b$ . Dann gilt:

- (1)  $\lim_{n \rightarrow \infty} (a_n + b_n) = a + b$ .
- (2)  $\lim_{n \rightarrow \infty} (a_n b_n) = ab$ .
- (3) Gilt  $a_n \neq 0$  und  $a \neq 0$ , so ist  $\lim_{n \rightarrow \infty} \left(\frac{b_n}{a_n}\right) = \frac{b}{a}$ .
- (4) Gilt für alle  $n$  die Ungleichung  $a_n \leq b_n$ , so ist  $a \leq b$ . Ist speziell  $m \leq a_n \leq M$  für fast alle  $n$ , so gilt  $m \leq a \leq M$ .

In Teil (1) ist natürlich gemeint, daß die Folge mit dem  $n$ -ten Glied  $a_n + b_n$  betrachtet wird. Analog in (2) und (3). Aus (3.8) und (3.9.2) folgt speziell für eine reelle Zahl  $\lambda$  die Gleichheit  $\lim_{n \rightarrow \infty} (\lambda a_n) = \lambda a$ .  $\square$

### 3.3. Konvergenzkriterien für Folgen

Sei  $(a_n \mid n \in \mathbb{N})$  eine Zahlenfolge. Eine Zahl  $M$  heißt *obere Schranke* (bzw. *untere Schranke*) der Folge, wenn für alle  $n$  die Ungleichung  $a_n \leq M$  (bzw.  $M \leq a_n$ ) gilt. Gibt es eine obere (untere) Schranke, so heißt die Folge nach oben (unten) beschränkt. Eine *beschränkte* Folge ist eine nach oben und unten beschränkte Folge.

**(3.10) Notiz.** *Eine konvergente Folge ist beschränkt.* □

Eine Folge  $(a_n)$  heißt *monoton wachsend* (bzw. *fallend*), wenn für alle  $n$  die Ungleichung  $a_n \leq a_{n+1}$  (bzw.  $a_n \geq a_{n+1}$ ) gilt.

Das folgende Konvergenzkriterium ist eine fundamentale Eigenschaft der reellen Zahlen.

**(3.11) Satz.** *Sei  $(a_n)$  monoton wachsend mit oberer Schranke  $M$ . Dann konvergiert die Folge, und es gilt für alle  $k \in \mathbb{N}$  die Ungleichung  $a_k \leq \lim_{n \rightarrow \infty} a_n \leq M$ . Ist  $(a_n)$  monoton fallend mit unterer Schranke  $m$ , so konvergiert die Folge, und es gilt  $m \leq \lim_{n \rightarrow \infty} a_n \leq a_k$ .* □

Der Grenzwert einer monoton wachsenden konvergenten Folge ist selbst eine obere Schranke, und zwar die kleinste unter allen möglichen oberen Schranken.

**(3.12) Beispiel.** Für  $k \in \mathbb{N}$  ist die Folge  $a_n = n^{-k}$  monoton fallend, und Null ist eine untere Schranke. Also konvergiert die Folge. Wegen (3.6) und (3.9.4) ist  $\lim_{n \rightarrow \infty} a_n = 0$ . ◇

**(3.13) Beispiel.** Als einfaches Beispiel für das Arbeiten mit den Grenzwertregeln untersuchen wir die Folge

$$a_n = \frac{n^2 - 1}{n^2 + 1}$$

auf Konvergenz. Wir schreiben um:

$$a_n = \frac{1 - n^{-2}}{1 + n^{-2}} = \frac{b_n}{c_n}$$

mit  $b_n = 1 - n^{-2}$  und  $c_n = 1 + n^{-2}$ . Wegen (3.9.1) und (3.12) ist  $\lim_{n \rightarrow \infty} b_n = 1 = \lim_{n \rightarrow \infty} c_n$  und dann wegen (3.9.3)  $\lim_{n \rightarrow \infty} a_n = 1$ . ◇

**(3.14) Satz.** *Sei  $|q| < 1$ . Dann ist die geometrische Folge  $1, q, q^2, \dots$ , mit dem allgemeinen Glied  $a_n = q^{n-1}$ , eine Nullfolge.*

BEWEIS. Sei zunächst  $q \geq 0$ . Dann handelt es sich um eine monoton fallende Folge mit unterer Schranke 0. Es gibt also einen Grenzwert  $c$  mit  $1 > c \geq 0$ . Die Folge  $b_n = qa_n$  hat denselben Grenzwert, da es sich im wesentlichen um dieselbe Folge handelt. Nach (3.9) gilt

$$c = \lim_{n \rightarrow \infty} b_n = q \lim_{n \rightarrow \infty} a_n = qc.$$

Das ist wegen  $q \neq 1$  nur für  $c = 0$  möglich.

Ist  $q \leq 0$ , so betrachten wir die Konvergenzdefinition. Es ist

$$|a_n - 0| = ||a_n| - 0|.$$

Ist deshalb  $(|a_n|)$  eine Nullfolge so auch  $(a_n)$ . Damit haben wir alles auf den Fall  $q \geq 0$  zurückgeführt. □

Das folgende *Konvergenzkriterium von Cauchy* benutzt nicht die Kenntnis des Grenzwertes.

**(3.15) Satz.** Eine Folge  $(a_n)$  konvergiert genau dann, wenn es zu jedem  $\varepsilon > 0$  ein  $N$  gibt, so daß für alle  $m, n \geq N$  die Ungleichung  $|a_m - a_n| < \varepsilon$  gilt.  $\square$

### 3.4. Reihen

**(3.16) Definition.** Sei  $(a_n \mid n \in \mathbb{N})$  eine Folge. Die Summe

$$s_n = a_1 + a_2 + \cdots + a_n = \sum_{k=1}^n a_k$$

heißt  $n$ -te *Partialsomme* der Folge  $(a_n)$ . Die Folge  $(s_n \mid n \in \mathbb{N})$  der Partialsummen wird als *unendliche Reihe* mit den Gliedern  $a_n$  bezeichnet und durch das Symbol

$$\sum_{j=1}^{\infty} a_j = a_1 + a_2 + a_3 + \cdots$$

bezeichnet. Natürlich kann man nicht unendlich viele Zahlen summieren, so daß dieses Symbol zunächst keinen Zahlenwert repräsentiert. Die unendliche Reihe heißt *konvergent* zum Grenzwert  $s$ , wenn  $s = \lim_{n \rightarrow \infty} s_n$  ist. In diesem Fall schreiben wir

$$\sum_{j=1}^{\infty} a_n = s$$

und die linke Seite des Gleichheitszeichens steht dann auch für den Zahlenwert  $s$ . Eine entsprechende Symbolik wird verwendet, wenn die Folge von einer anderen Stelle an durchnummeriert wird. Der Beginn bei Null tritt häufig auf.  $\diamond$

**(3.17) Beispiel.**  $\sum_{j=1}^{\infty} j^2$  ist keine Zahl, da die Folge der Partialsummen nicht nach oben beschränkt ist, also nicht konvergiert.  $\diamond$

**(3.18) Bemerkung.** Endlich viele Glieder sind für die Konvergenz einer Reihe belanglos. Ist  $\sum_{j=1}^{\infty} a_j$  konvergent so auch  $\sum_{j=k}^{\infty} a_j$  für jedes  $k \geq 2$ , und es gilt

$$\sum_{j=1}^{\infty} a_j = \sum_{j=1}^{k-1} a_j + \sum_{j=k}^{\infty} a_j.$$

Diese Bemerkung wird meist stillschweigend verwendet.  $\diamond$

Die Reihe  $\sum_{j=0}^{\infty} q^j$  heißt *geometrische Reihe*.

**(3.19) Satz.** Für  $|q| < 1$  konvergiert die geometrische Reihe.

BEWEIS. Für die Partialsumme  $s_n = 1 + q + q^2 + \cdots + q^{n-1}$  gilt die Formel

$$s_n = \frac{1 - q^n}{1 - q}.$$

Das folgt, indem man  $s_n(1 - q)$  nach dem Distributivgesetz auflöst und erkennt, daß sich alle Summanden bis auf zwei paarweise aufheben. Wir haben schon gesehen, daß  $(q^n)$  eine Nullfolge ist. Nach den Grenzwertregeln (3.9) ist also

$$\lim s_n = \lim \frac{1 - q^n}{1 - q} = \frac{1 - \lim q^n}{1 - q} = \frac{1}{1 - q}.$$

□

**(3.20) Beispiel.** Die sogenannte *harmonische Reihe*  $\sum_{j=1}^{\infty} \frac{1}{j}$  divergiert. ◇

### 3.5. Konvergenzkriterien für Reihen

Das Cauchy-Kriterium für Folgen nimmt wegen

$$s_m - s_n = \sum_{j=n+1}^m a_j$$

für  $m > n$  die folgende Form an:

**(3.21) Konvergenzkriterium von Cauchy für Reihen.** *Eine unendliche Reihe konvergiert genau dann, wenn zu jedem  $\varepsilon > 0$  ein  $N$  existiert, so daß für  $m > n \geq N$  immer die Ungleichung  $|a_{n+1} + \dots + a_m| < \varepsilon$  besteht.* □

**(3.22) Majorantenkriterium** *Seien  $(a_n)$  und  $(b_n)$  Folgen und gelte immer  $|b_n| \leq a_n$ . Dann heißt die Reihe  $\sum a_j$  eine Majorante der Reihe  $\sum b_j$ . Konvergiert  $\sum a_j$  so auch  $\sum b_j$ .*

BEWEIS. Wegen der Dreiecksungleichung und der Voraussetzung gilt

$$\left| \sum_{j=n+1}^m b_j \right| \leq \sum_{j=n+1}^m |b_j| \leq \sum_{j=n+1}^m a_j.$$

Jetzt wende man das Cauchy-Kriterium (3.21) an. □

**(3.23) Quotientenkriterium.** *Sei  $0 \leq q < 1$ . Für fast alle  $n$  sei  $a_n \neq 0$  und gelte*

$$\left| \frac{a_{n+1}}{a_n} \right| \leq q.$$

*Dann ist  $\sum a_n$  konvergent.*

BEWEIS. Aus der Voraussetzung folgt induktiv

$$|a_n| \leq q|a_{n-1}| \leq \dots \leq q^{n-1}|a_1|.$$

Also ist die geometrische Reihe eine Majorante. □

Es sei betont, daß die Zahl  $q$  fest wählbar sein muß. Es reicht keineswegs aus, daß für fast alle  $n$  der Quotient  $|a_{n+1}/a_n|$  kleiner als 1 ist<sup>6</sup>, was die berühmte *harmonische Reihe*  $a_n = 1/n$  belegt, die so „langsam“ divergiert, daß man ihren Hang zum Unendlichen kaum bemerkt.

<sup>6</sup>Ein bekannter Anfängerfehler

**(3.24) Beispiel.** Sei  $x$  eine reelle Zahl. Wir untersuchen die Reihe  $\sum_{j=1}^{\infty} a_n$

$$x - \frac{x^2}{2} + \frac{x^3}{3} - \dots, \quad a_n = (-1)^{n+1} \frac{x^n}{n}.$$

Wegen

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{n}{n+1} x \right| \leq |x|$$

konvergiert die Reihe für  $|x| < 1$  nach dem Quotientenkriterium.  $\diamond$

**(3.25) Beispiel.** Sei  $x$  eine reelle Zahl. Die Reihe

$$\sum_{k=0}^{\infty} \frac{x^k}{k!}$$

heißt *Exponentialreihe* und wird bei der Untersuchung der Exponentialfunktion eine wichtige Rolle spielen. Nach dem Quotientenkriterium konvergiert sie für alle reellen Zahlen  $x$ .  $\diamond$

**(3.26) Beispiel.** Die Reihe  $\sum n^{-2}$  konvergiert. Das läßt sich aber *nicht* unmittelbar mit dem Quotientenkriterium feststellen! Zwar ist  $|a_{n+1}/a_n|$  immer kleiner als 1, es gibt aber kein  $q < 1$ , das die Bedingung aus (3.23) erfüllt.  $\diamond$

## 4 Differentialrechnung

### 4.1. Der Differenzenquotient

Die Fragestellung der Differentialrechnung läßt sich allgemein so umschreiben: Wie *ändert* sich eine Funktion und welche Rückschlüsse auf die Funktion gestattet die Kenntnis des Änderungsverhaltens? Die erste Definition gibt ein ungefähres Maß für die Änderung.

**(4.1) Definition.** Sei  $f: J \rightarrow \mathbb{R}$  eine auf einem Intervall  $J$  definierte Funktion und seien  $b, a$  zwei *verschiedene* Zahlen aus dem Intervall  $J$ . Der Quotient

$$(4.2) \quad \frac{f(b) - f(a)}{b - a}$$

heißt die *mittlere Änderungsrate der Funktion* im Intervall von  $b$  nach  $a$ . Ein Quotient der Form (4.2) wird *Differenzenquotient* von  $f$  genannt.

Der Graph der Funktion

$$(4.3) \quad x \mapsto f(a) + \frac{f(b) - f(a)}{b - a} (x - a)$$

ist eine Gerade durch die Punkte  $(a, f(a))$  und  $(b, f(b))$ , wie man durch Einsetzen von  $x = a, b$  sofort verifiziert. Der Differenzenquotient (4.2) ist die Steigung dieser Geraden. Da die Gerade den Graphen in zwei vorgegebenen Punkten schneidet, nennt man ihn auch eine *Sekante* des Graphen.  $\diamond$

Ein typisches Beispiel für eine Änderungsrate ist die mittlere (oder durchschnittliche) Geschwindigkeit. Wir betrachten einige Beispiele für Differenzenquotienten.

(4.4) **Beispiel.** Lineare Funktionen  $f(x) = \lambda x + c$ . Damit wird

$$\frac{f(b) - f(a)}{b - a} = \frac{\lambda b + c - (\lambda a + c)}{b - a} = \lambda.$$

Die durchschnittliche Änderungsrate ist also gleich der Steigung der Geraden.  $\diamond$

(4.5) **Beispiel.** Quadratische Funktion  $f(x) = x^2$ . Wir rechnen

$$\frac{f(b) - f(a)}{b - a} = \frac{b^2 - a^2}{b - a} = a + b.$$

Bei festgehaltenem  $a$  hängt die Änderungsrate also von der Länge des Intervalls ab.  $\diamond$

(4.6) **Beispiel.** Sei  $f(x) = x^{-1}$ . Damit wird

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

Die kleine Umformung zum letzten Gleichheitszeichen kann getrost dem Leser überlassen werden.  $\diamond$

## 4.2. Die Ableitung

Will man die „augenblickliche“ Änderungsrate an einer Stelle erklären, so wird man dazu geführt, den Punkt  $b$  immer näher an den Punkt  $a$  heranzurücken und zu untersuchen, ob es einen Grenzwert gibt. Die Grenzlage der Sekante wird dann eine *Tangente* an den Graphen.

(4.7) **Definition.** Eine Funktion  $f: J \rightarrow \mathbb{R}$  heißt *im Punkt*  $a \in J$  *differenzierbar*, wenn der Differenzenquotient

$$\frac{f(b) - f(a)}{b - a}$$

bei der Annäherung  $b \rightarrow a$  einen Grenzwert  $A$  hat, wenn also, symbolisch geschrieben gilt

$$\lim_{b \rightarrow a} \frac{f(b) - f(a)}{b - a} = A.$$

Die Zahl  $A$  heißt dann *Ableitung* oder *Differentialquotient* von  $f$  im Punkt  $a$  und wird mit

$$(4.8) \quad f'(a) \quad \text{oder} \quad \frac{df}{dx}(a)$$

bezeichnet. Wir nennen  $f$  *differenzierbar*, wenn  $f$  in jedem Punkt von  $J$  differenzierbar ist. Die dadurch gegebene Funktion

$$f' = \frac{df}{dx}: J \rightarrow \mathbb{R}, \quad x \mapsto f'(x)$$

heißt dann die *Ableitung* von  $f$ . Wir *differenzieren*  $f$ , indem wir die Ableitung  $f'$  bilden, und dieser Prozeß heißt *Differentiation*.  $\diamond$

Wir schreiben den Ableitungsprozeß auch manchmal in der Form

$$\lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h} = f'(a).$$



Die Gerade

$$x \mapsto f(a) + f'(a)(x - a)$$

durch den Punkt  $(a, f(a))$  heißt *Tangente* an den Graphen von  $f$  im Punkt  $a$ . Die Ableitung ist also deren Steigung.

Wir betrachten die früheren Beispiele.

**(4.9) Beispiel.** Sei  $f(x) = \lambda x + c$ . Da der Differenzenquotient konstant gleich  $\lambda$  ist, so ist auch der Grenzwert gleich  $\lambda$ . Also gilt  $f'(x) = \lambda$ . Insbesondere hat eine konstante Funktion die Ableitung Null.  $\diamond$

**(4.10) Beispiel.** Sei  $f(x) = x^2$ . Der Differenzenquotient hatte den Wert  $a + b$ , und der Grenzwert für  $b \rightarrow a$  ist offenbar  $2a$ . Also gilt  $f'(x) = 2x$ .  $\diamond$

**(4.11) Beispiel.** Für  $f(x) = x^{-1}$  hat der Differenzenquotient den Wert  $-(ab)^{-1}$ , und der Grenzwert für  $b \rightarrow a$  ist offenbar  $-a^{-2}$ . Also ist  $f'(x) = -x^{-2}$ .  $\diamond$

Aus beweistechnischen Gründen ist es zweckmäßig, die Differenzierbarkeit auch in der folgenden Weise zu formulieren:  $f$  ist im Punkt  $a$  differenzierbar, wenn es eine im Punkt  $a$  stetige Funktion  $\Psi$  gibt, mit der die Gleichung

$$f(x) = f(a) + (x - a)\Psi(x)$$

gilt; der Funktionswert  $\Psi(a)$  ist dann die Ableitung im Punkt  $a$ . Daraus folgt unmittelbar, daß eine im Punkt  $a$  differenzierbare Funktion dort auch stetig ist.

**(4.12) Beispiel.** Die Funktion  $f(x) = |x|$  ist im Punkt  $a = 0$  nicht differenzierbar, aber in allen anderen Punkten. Der Graph hat im Punkt  $(0, 0)$  eine „Ecke“. Für  $x > 0$  ist die Ableitung gleich 1, für  $x < 0$  gleich  $-1$ . Eine solche Funktion mit einem „Sprung“ an der Stelle Null kann durch keinen Funktionswert an der Stelle Null stetig gemacht werden.  $\diamond$

Die Ableitung einer Funktion ist wieder eine Funktion, die wir ebenfalls untersuchen können. Die Ableitung der Ableitung (wenn es sie gibt) heißt zweite Ableitung  $f''$ . So fortfahrend können wir gegebenenfalls die  $n$ -te Ableitung  $f^{(n)}$  von  $f$  bilden. Gibt es die  $n$ -te Ableitung von  $f$  und ist diese überdies stetig, so heißt  $f$   $n$ -mal *stetig differenzierbar*. Für die  $n$ -te Ableitung gibt es auch die Bezeichnung

$$\frac{d^n f}{dx^n}.$$

### 4.3. Ableitungsregeln

Die Definition (4.7) dient dazu, das theoretische Gemüt zu besänftigen. Arbeiten möchte man nicht damit. Deshalb benutzt man die Definition zunächst dazu, die im weiteren aufgeführten handlichen Rechenregeln herzuleiten.

**(4.13) Satz.** Seien  $f, g: J \rightarrow \mathbb{R}$  differenzierbare Funktionen und  $\lambda, \mu$  reelle Zahlen. Dann sind auch die Funktionen  $\lambda f + \mu g$ ,  $f \cdot g$  und (falls  $g(x) \neq 0$ )  $f/g$  differenzierbar und es gilt:

$$\begin{aligned} (\lambda f + \mu g)' &= \lambda f' + \mu g' \\ (f \cdot g)' &= f' \cdot g + f \cdot g' \\ \left(\frac{f}{g}\right)' &= \frac{f' \cdot g - f \cdot g'}{g^2}. \end{aligned}$$

Die erste Regel heißt die Linearität der Ableitung, die zweite Regel nennen wir die Produktregel und die dritte die Quotientenregel.

BEWEIS. Wir erläutern die Entstehung dieser Regeln. Die Produktregel ergibt sich durch Grenzübergang aus der Gleichung

$$\frac{f(b)g(b) - f(a)g(a)}{b - a} = \frac{f(b) - f(a)}{b - a}g(b) + f(a)\frac{g(b) - g(a)}{b - a}.$$

Die erste Regel ergibt sich, weil der Differenzenquotient der Summe gleich der Summe der Differenzenquotienten ist. Die Quotientenregel ergibt sich, wie weiter unten erläutert, aus anderen Regeln.  $\square$

Aus den Regeln folgt sofort durch wiederholte Anwendung, daß alle Polynome und rationalen Funktionen differenzierbar sind, denn diese lassen sich ja aus den schon als differenzierbar erkannten linearen Funktionen durch Summen-, Produkt-, und Quotientbildung herstellen. Außerdem sagen einem die Regeln, wie die Ableitungen zu berechnen sind. Es gilt nämlich:

**(4.14) Beispiel.** Sei  $n$  eine natürliche Zahl. Die Potenzfunktion  $f(x) = x^n$  hat die Ableitung  $f'(x) = nx^{n-1}$ . Das wird mit den Regeln durch Induktion nach  $n$  bewiesen, denn die Aussage ist uns für  $n = 1$  schon bekannt und für den Induktionsschritt wenden wir die Produktregel auf die Zerlegung  $x^{n+1} = x^n \cdot x$  an. Mit der Quotientenregel erhalten wir daraus für die Funktion  $f(x) = x^{-n}$  die Ableitung  $f'(x) = -nx^{-n-1}$ . Also gilt

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

für alle ganzen Zahlen  $n$ . Mit den Potenzfunktionen können wir dann wegen der Linearität und der Quotientenregel sofort alle rationalen Funktionen differenzieren. Die Polynomfunktion  $f(x) = \sum_{k=0}^n a_k x^k$  hat also die Ableitung  $f'(x) = \sum_{k=1}^n k a_k x^{k-1}$ .  $\diamond$

**(4.15) Kettenregel.** Seien  $f: A \rightarrow B$  und  $g: B \rightarrow C$  differenzierbare Funktionen. Dann gilt für die Ableitung der Verkettung

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

Diese letzte Formel wird Kettenregel genannt.  $\square$

In der Formulierung der Kettenregel ist zu beachten, daß die Ableitung der Funktion  $g$  an der Stelle  $f(x)$  zu nehmen ist. Der Beweis der Kettenregel folgt mit der Umformulierung der Differenzierbarkeit aus der Tatsache, daß die Verkettung stetiger Funktionen wieder stetig ist. Die Formel

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

verdeutlicht, wie es zur Kettenregel kommt.

**(4.16) Beispiel.** Sei  $f(x) \neq 0$ . Die Funktion  $h(x) = (f(x))^{-1}$  ist die Verkettung von  $g(x) = x^{-1}$  mit  $f$ . Die Kettenregel zusammen mit der Kenntnis der Ableitung von  $g$  liefert in diesem Fall das Ergebnis der Quotientenregel. Zusammen mit der Produktregel erhalten wir dann die allgemeine Quotientenregel.  $\diamond$

**(4.17) Beispiel.** Zwei einfache Spezialfälle der Kettenregel: Die Funktion  $x \mapsto f(x + c)$  hat an der Stelle  $x$  die Ableitung  $f'(x + c)$ ; und die Funktion  $x \mapsto f(\lambda x)$  hat an der Stelle  $x$  die Ableitung  $\lambda f'(\lambda x)$ .  $\diamond$

#### 4.4. Umkehrfunktionen

Seien  $f: A \rightarrow B$  und  $g: B \rightarrow A$  Umkehrfunktionen voneinander. Auf die Gleichung  $f(g(x)) = x$  wenden wir die Kettenregel an und erhalten

$$f'(g(x)) \cdot g'(x) = 1.$$

Mithin läßt sich die Ableitung von  $g$  aus derjenigen von  $f$  durch die Formel

$$(4.18) \quad g'(x) = \frac{1}{f'(g(x))}$$

berechnen (*Umkehrregel*). Wir sehen außerdem, daß die Ableitung der Funktion und ihrer Umkehrung nicht Null sind.

**(4.19) Beispiel.** Die differenzierbare Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto x^3$  hat eine Umkehrfunktion. Diese ist aber im Punkt Null nicht differenzierbar, weil  $f'(0) = 0$  ist. Geometrisch gesprochen: Die Umkehrfunktion von  $f$  hat an der Stelle Null eine senkrechte Tangente.  $\diamond$

**(4.20) Beispiel.** Sei  $f(x) = x^n$  und  $g(x) = \sqrt[n]{x}$ . Dann ergibt sich nach der letzten Formel

$$g'(x) = \frac{1}{n(\sqrt[n]{x})^{n-1}}.$$

Siehe dazu aber den späteren Abschnitt über Potenzfunktionen.  $\diamond$

**(4.21) Satz.** Die differenzierbare Funktion  $f: A \rightarrow B$  habe eine Umkehrfunktion  $g$ . Sei  $f'(a) \neq 0$ . Dann ist  $g$  im Punkt  $b = f(a)$  differenzierbar und für die Ableitung gilt die Formel (4.18).  $\square$

#### 4.5. Der Mittelwertsatz

Warum ist die Ableitung nützlich? Aus dem Verlauf der Ableitung kann man auf den Verlauf der Funktion zurückschließen. Das wichtigste theoretische Werkzeug dafür ist der folgende Hauptsatz über die Ableitung.

**(4.22) Der Mittelwertsatz.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  differenzierbar. Dann gibt es eine Stelle  $\xi \in ]a, b[$ , so daß die Gleichung

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

gilt.  $\square$

Dieser Satz hat die folgende anschauliche Bedeutung. Wir betrachten die Sekante durch die Punkte  $(a, f(a))$  und  $(b, f(b))$ . Wir verschieben sie sodann parallel solange nach oben oder unten, bis sie den Graphen nur noch „berührt“, das heißt zur Tangente degeneriert. Gibt es etwa unterhalb der Sekante Punkte des Graphen, so irgendwann bei Parallelverschiebung einen „letzten“  $(\xi, f(\xi))$ , der von der verschobenen Sekante getroffen wird.

Der folgende Spezialfall des Mittelwertsatzes wird *Satz von Rolle* genannt.

**(4.23) Satz.** Sei  $f: [a, b] \rightarrow \mathbb{R}$  differenzierbar und gelte  $f(a) = f(b)$ . Dann gibt es eine Stelle  $\xi \in ]a, b[$  mit  $f'(\xi) = 0$ .

BEWEIS. Der Satz ist sicherlich richtig, wenn  $f$  konstant ist. Andernfalls gibt es im Innern des Intervalls eine Stelle  $\xi$ , an der die stetige Funktion  $f$  ein Maximum oder ein Minimum hat. Der Satz von Rolle folgt deshalb aus dem folgenden Satz über lokale Extremwerte.  $\square$

Sei  $f: J \rightarrow \mathbb{R}$  eine auf dem offenen Intervall  $J$  erklärte Funktion. Wir sagen,  $f$  habe an der Stelle  $a \in J$  ein *lokales Maximum* (bzw. *lokales Minimum*), wenn es ein  $t > 0$  gibt, so daß für alle  $x$  im Intervall  $[a - t, a + t] \cap J$  die Ungleichung  $f(x) \leq f(a)$  (bzw.  $f(x) \geq f(a)$ ) gilt. Das lokale Maximum an der Stelle  $a$  heißt *isoliert*, wenn für alle  $c \neq a$ ,  $|c - a| < t$  gilt  $f(c) < f(a)$ . Analog für das Minimum.

**(4.24) Satz.** *Hat die differenzierbare Funktion  $f: J \rightarrow \mathbb{R}$  im Innern des Intervalles an der Stelle  $a$  ein lokales Maximum oder Minimum, so ist dort  $f'(a) = 0$ .*

BEWEIS. Liege bei  $a$  ein lokales Maximum, d. h. gelte für  $a - t \leq x \leq a + t$  immer  $f(x) \leq f(a)$ . Dann ist der Differenzenquotient

$$\frac{f(x) - f(a)}{x - a}$$

für  $a < x \leq a + t$  kleinergleich Null und für  $a - t \leq x < a$  größ ergleich Null. Für den Grenzwert bleibt dann nur die Möglichkeit, Null zu sein. Analog für das Minimum.  $\square$

*Beweis des Mittelwertsatzes.* Der Satz von Rolle wird auf die Funktion  $g$

$$g(x) = f(x) - \left( f(a) + (x - a) \frac{f(b) - f(a)}{b - a} \right)$$

angewendet. Dann ist nämlich  $g(a) = g(b) = 0$ , und  $g'(\xi) = 0$  liefert genau die Formel des Mittelwertsatzes.  $\square$

Aus dem Mittelwertsatz ergibt sich unmittelbar die folgende Anwendung, die zeigt, daß eine differenzierbare Funktion auf einem Intervall durch ihre Ableitung und einen einzigen Funktionswert festgelegt ist.

**(4.25) Folgerung.** *Die auf einem Intervall  $J$  definierte differenzierbare Funktion  $f$  habe die Ableitung Null. Dann ist  $f$  eine konstante Funktion. Haben  $f, g: J \rightarrow \mathbb{R}$  dieselbe Ableitung, so gilt mit einer Konstanten  $c$  die Gleichung  $f(x) = g(x) + c$  für alle  $x \in J$ .*  $\square$

Eine zweite wichtige Folgerung aus dem Mittelwertsatz ist:

**(4.26) Folgerung.** *Habe die auf einem Intervall erklärte Funktion eine überall positive (bzw. negative) Ableitung. Dann ist die Funktion streng monoton wachsend (bzw. fallend).*  $\square$

Eine differenzierbare Funktion  $F: J \rightarrow \mathbb{R}$  heißt *Stammfunktion* der Funktion  $f: J \rightarrow \mathbb{R}$ , wenn  $f$  die Ableitung von  $F$  ist. Der zweite Hauptsatz der Differentialrechnung ist:

**(4.27) Satz.** *Jede auf einem Intervall  $J$  definierte stetige Funktion besitzt eine Stammfunktion.*

Dieser Satz folgt aus dem später behandelten Hauptsatz der Differential- und Integralrechnung.

#### 4.6. Das lokale Verhalten

Wir geben Bedingungen an, mit denen man entscheiden kann, ob ein lokales Maximum oder Minimum vorliegt. Der folgende Satz ist der Ausgangspunkt.

**(4.28) Satz.** Sei  $f: J \rightarrow \mathbb{R}$  eine auf dem offenen Intervall  $J$  differenzierbare Funktion und sei  $f'(a) = 0$ .

- (1) Ist  $f'(a+h) \cdot h > 0$  für alle hinreichend kleinen  $h \neq 0$ , so ist  $a$  ein isoliertes lokales Minimum von  $f$ . Dies ist insbesondere der Fall, wenn  $f'$  um  $a$  streng monoton wächst, oder wenn  $f''(a)$  existiert und positiv ist.
- (2) Ist  $a$  ein isoliertes lokales Minimum von  $f'$ , so wächst  $f$  lokal um  $a$  streng monoton.

BEWEIS. (1) Für  $t \neq 0$  ist nach dem Mittelwertsatz  $f(a+t) = f(a) + f'(a+h)t$  für ein  $h \in ]0, t[$ . Aus der Voraussetzung folgt  $f'(a+h)t > 0$  für alle genügend kleinen  $t$ . Die Voraussetzung besagt, daß  $f'(a+h)$  dasselbe Vorzeichen wie  $h \neq 0$  hat; wegen  $f'(a) = 0$  folgt dies, wenn  $f'$  um  $a$  streng monoton wächst. Existiert  $f''(a)$ , so gibt es eine bei Null stetige Funktion  $\Phi$  mit  $f'(a+h) = f'(a) + \Phi(h)h = \Phi(h)h$  mit  $\Phi(0) = f''(a) > 0$ . Wegen der Stetigkeit von  $\Phi$  bei Null ist dann auch  $\Phi(h) > 0$  für alle hinreichend kleinen  $h$ .

(2) Wegen  $f'(a) = 0$  ist  $f'(a+h) > 0$  für kleine  $h \neq 0$ . Die Behauptung folgt jetzt aus (4.26).  $\square$

Der vorstehende Satz läßt sich auf höhere Ableitungen übertragen.

#### 4.7. Konvex. Konkav

Sei  $f: J \rightarrow \mathbb{R}$  auf dem offenen Intervall  $J$  differenzierbar. Die Funktion heißt *konvex* (bzw. *streng konvex*), wenn für alle  $a, b \in J$  gilt

$$f(b) \geq f(a) + f'(a)(b-a), \quad \text{bzw.} \quad f(b) > f(a) + f'(a)(b-a).$$

Sie heißt *konkav* (bzw. *streng konkav*), wenn für alle  $a, b \in J$  gilt

$$f(b) \leq f(a) + f'(a)(b-a), \quad \text{bzw.} \quad f(b) < f(a) + f'(a)(b-a).$$

Ist  $f$  streng konvex (konkav), so liegt der Graph von  $f$  mit Ausnahme des Punktes  $(a, f(a))$  oberhalb (unterhalb) der Tangente durch diesen Punkt.

**(4.29) Satz.** Eine in  $J$  differenzierbare Funktion  $f$  ist genau dann konvex (konkav), wenn  $f'$  monoton wachsend (fallend) ist. Analog mit den Zusätzen streng. Eine in  $J$  zweimal differenzierbare Funktion ist genau dann konvex (konkav), wenn für alle  $x \in J$  gilt  $f''(x) \geq 0$  ( $f''(x) \leq 0$ ).

#### 4.8. Verallgemeinerter Mittelwertsatz

**(4.30) Satz.** Seien  $f, g: [a, b] \rightarrow \mathbb{R}$  differenzierbar. Dann gibt es ein  $\xi \in ]a, b[$ , so daß

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

Verswindet  $g'$  nirgends auf  $]a, b[$ , so ist  $g(b) \neq g(a)$ , und es gilt

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

BEWEIS. Man wende den Satz von Rolle auf die Funktion

$$F(x) = (f(x) - f(a))(g(b) - g(a)) - (g(x) - g(a))(f(b) - f(a))$$

an. Die zweite Behauptung folgt, weil  $g$  streng monoton ist.  $\square$

#### 4.9. Grenzwertregeln

**(4.31) Satz.** Seien  $f, g: [a, b] \rightarrow \mathbb{R}$  differenzierbar, sei  $f(a) = g(a) = 0$  und  $g'(x) \neq 0$  auf  $]a, b[$ . Existiert dann

$$\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}, \quad \text{so auch} \quad \lim_{x \rightarrow a} \frac{f(x)}{g(x)},$$

und beide sind gleich.

**(4.32) Satz.** Seien  $f, g$  auf  $[b, \infty[$  differenzierbar, sei  $g' \neq 0$  und gelte  $\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} g(x) = 0$ . Dann gilt

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)},$$

falls der rechte Grenzwert existiert.

**(4.33) Satz.** Die Funktionen  $f, g$  seien auf  $[a, b[$  differenzierbar. Sei  $g' \neq 0$  und gelte  $\lim_{x \rightarrow b} g(x) = \infty$ . Dann ist

$$\lim_{x \rightarrow b} \frac{f(x)}{g(x)} = \lim_{x \rightarrow b} \frac{f'(x)}{g'(x)},$$

falls der rechte Grenzwert existiert. Hierbei darf auch  $b = \infty$  sein.

#### 4.10. Taylor-Polynome und Reihen

Sei  $f$  im Punkt  $a$   $n$ -mal differenzierbar. Dann heißt

$$j_a^n f(x) := f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

das *Taylor-Polynom*  $n$ -ten Grades von  $f$  an der Stelle  $a$ . Die Differenz  $f(x) - j_a^n f(x) = r_a^n(x)$  wird *Restglied* genannt. Das Taylor-Polynom ist die beste Approximation der Funktion in einer Umgebung von  $a$  durch ein Polynom vom Grad  $n$ ; der Fall  $n = 1$  war die Tangente an der Stelle  $a$ . Die Taylorsche Formel sagt etwas über die Größe des Restgliedes aus.

**(4.34) Satz.** Sei  $f$  im Intervall  $\{a+tx \mid 0 \leq t \leq 1\}$   $(n+1)$ -mal differenzierbar. Dann gibt es ein  $\delta \in ]0, 1[$  mit dem das Restglied die Form

$$r_a^n f(x) = \frac{f^{(n+1)}(a + \delta x)}{(n+1)!} x^{n+1}$$

hat.

Ist  $f$  beliebig oft differenzierbar, so heißt die Potenzreihe

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

die *Taylor-Reihe* von  $f$  an der Stelle  $a$ . Diese braucht nicht zu konvergieren, und wenn sie konvergiert, muß ihr Grenzwert nicht gleich  $f(x)$  sein. Falls ihr Grenzwert aber für alle  $x$  in einer Umgebung von  $a$  gleich  $f(x)$  ist, so nennen wir sie eine Darstellung der Funktion durch ihre Taylor-Reihe, eine *Potenzreihenentwicklung* der Funktion um  $a$ .

#### 4.11. Potenzreihen

Eine Reihe der Form

$$(4.35) \quad \sum_{k=0}^{\infty} c_k (x-a)^k$$

heißt *Potenzreihe* um  $a$ . Über das Konvergenzverhalten von Potenzreihen gilt:

**(4.36) Satz.** *Entweder konvergiert die Reihe (4.35) für alle  $x$  oder es gibt eine Zahl  $R$ , so daß die Reihe für  $|x-a| < R$  konvergiert und für  $|x-a| > R$  divergiert.*

Die Zahl  $R$  des letzten Satzes nennt man den *Konvergenzradius* der Reihe; wir setzen  $R = \infty$ , wenn der erste Fall des Satzes eintritt. Wir nennen  $\{x \mid |x-a| < R\}$  das *Konvergenzintervall* der Reihe.

**(4.37) Satz.** *Im Konvergenzintervall ist eine Potenzreihe differenzierbar und die Ableitung ist die Reihe mit gleichem Konvergenzintervall, die sich durch gliedweises Differenzieren ergibt. Im Konvergenzintervall ist die Reihe gleich der Taylor-Reihe der durch sie dargestellten Funktion.*

#### 4.12. Partielle Ableitungen

Sei  $\mathbb{R}^n$  die Menge aller  $n$ -Tupel  $(x_1, \dots, x_n)$  reeller Zahlen  $x_i$ , siehe III.2. Eine Menge  $D \subset \mathbb{R}^n$  heiße *offen*, wenn zu jedem  $x \in D$  ein  $\varepsilon$  existiert, so daß die sogenannte  $\varepsilon$ -Umgebung

$$U_\varepsilon(x) = \{y \in \mathbb{R}^n \mid \|y-x\| < \varepsilon\}$$

in  $D$  liegt. Offene Mengen sind geeignete Definitionsbereiche von Funktionen mehrerer Veränderlicher.

Sei  $f: D \rightarrow \mathbb{R}$  eine solche Funktion. Halten wir alle  $x_i$  außer  $x_k$  fest, so erhalten wir die Funktion

$$x_k \mapsto f(x_1, \dots, x_k, \dots, x_n).$$

Ist diese Funktion an der Stelle  $x_k$  differenzierbar, so bezeichnen wir die Ableitung mit

$$\frac{\partial f}{\partial x_k}(x_1, \dots, x_n)$$

und nennen sie die *partielle Ableitung* nach der Variablen  $x_k$ .

**(4.38) Beispiel.** Sei  $f(x, y) = x^3 + xy^2$ . Dann ist

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

Bei Funktionen von mehreren Veränderlichen nimmt die *Kettenregel* die folgende Form an:

**(4.39) Satz.** Seien  $\alpha_i(t)$  differenzierbare Funktionen und sei für  $t$  aus einem kleinen Intervall immer  $(\alpha_1(t), \dots, \alpha_n(t)) \in D$ . Dann ist die Ableitung von  $t \mapsto f(\alpha_1(t), \dots, \alpha_n(t))$  an der Stelle  $t$  gleich

$$\sum_{i=1}^n \frac{\partial f}{\partial x_i}(\alpha_1(t), \dots, \alpha_n(t)) \alpha_i'(t).$$

Der Vektor aus den partiellen Ableitungen heißt der *Gradient* der Funktion

$$\text{grad}f(x) = \left( \frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right).$$

Hat die Funktion an der Stelle  $x$  ein lokales Maximum oder Minimum, so ist  $\text{grad}f(x) = 0$ . Gibt es auch die zweiten partiellen Ableitungen, so kann man daraus die Matrix

$$H_f(x) \left( \frac{\partial^2 f}{\partial x_i \partial x_k} \right)$$

bilden, die auch *Hesse-Matrix* genannt wird.

**(4.40) Beispiel.** Die Matrix der zweiten partiellen Ableitungen von  $f(x, y) = x^3 + xy^2$  ist

$$\begin{pmatrix} 6x & 2y \\ 2y & 2x \end{pmatrix}.$$

**(4.41) Satz.** Habe  $f$  an der Stelle  $x$  den Gradienten Null. Ist für Vektoren  $v \neq 0$  immer das Skalarprodukt  $\langle H_f(x)v, v \rangle > 0$ , so hat  $f$  an der Stelle  $x$  ein lokales Minimum. Nimmt dieses Skalarprodukt sowohl positive als auch negative Werte an, so liegt kein lokales Extremum vor.  $\square$

Der Graph einer Funktion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  ist eine Fläche im  $\mathbb{R}^3$  und kann als ein „Gebirge“ veranschaulicht werden. Die Urbilder eines Punktes

$$f^{-1}(c) := \{y \in \mathbb{R}^2 \mid f(y) = c\}$$

sind die „Höhenlinien“ der Höhe  $c$ , auch Niveaukurven genannt. Ist  $w: J \rightarrow \mathbb{R}^2$  ein Weg, der ganz in einer Niveaukurve liegt, so ist der Geschwindigkeitsvektor  $w'(t)$  eine Tangente an die Niveaukurve. Die Funktion  $f \circ w$  ist konstant, hat also die Ableitung Null. Diese Ableitung ist nach der Kettenregel gleich dem Skalarprodukt

$$\langle \text{grad}f(w(t)), w'(t) \rangle = 0.$$

Also ist der Gradient orthogonal zur Niveaukurve; er zeigt in Richtung des steilsten Anstieges, weil  $t \mapsto f(x + t \text{grad}f)$  die Ableitung  $\langle \text{grad}f, \text{grad}f \rangle$  hat, also  $f$  in Richtung des Gradienten wächst.

**(4.42) Beispiel.** Die Niveaukurven von  $f(x, y) = x^2 + y^2$  sind Kreise. Der Gradient ist  $(2x, 2y)$ . Diese Vektoren sind senkrecht zu den Kreislinien.  $\diamond$



## 5 Exponentialfunktion und Logarithmus

Die Exponentialfunktion ist die wichtigste mathematische Funktion. Der Logarithmus ist ihre Umkehrfunktion. Beide Funktionen werden in diesem Abschnitt erklärt, und ihre wichtigsten Eigenschaften werden hergeleitet.

### 5.1. Potenzen mit reellen Exponenten

Für eine von Null verschiedene reelle Zahl  $a$  und eine ganze Zahl  $n$  haben wir die Potenz  $a^n$  früher erklärt und die Rechenregeln darüber in den Potenzgesetzen formuliert.

Es ist möglich, das Symbol  $a^x$  für *positive* reelle Zahlen  $a$  und *beliebige* reelle Zahlen  $x$  so zu erklären, daß es die im folgenden aufgezählten Eigenschaften hat.

#### (5.1) Eigenschaften der Potenz $a^x$ .

- (1) Es gilt immer  $a^x > 0$ .
- (2)  $a^{x+y} = a^x a^y$ .
- (3)  $(a^x)^y = a^{xy}$ .
- (4)  $a^x b^x = (ab)^x$ .
- (5) Für eine ganze Zahl  $n$  hat  $a^n$  die früher erklärte Bedeutung.
- (6)  $a^0 = 1, 1^x = 1$ .
- (7) Sei  $a > 1$ . Dann gilt  $x < y \Leftrightarrow a^x < a^y$ .
- (8) Sei  $a < 1$ . Dann gilt  $x < y \Leftrightarrow a^x > a^y$ .
- (9) Sei  $x > 0$ . Dann gilt  $a < b \Leftrightarrow a^x < b^x$ .
- (10) Sei  $x < 0$ . Dann gilt  $a < b \Leftrightarrow a^x > b^x$ .

In dem Symbol  $a^x$  können wir  $x$  oder  $a$  als Veränderliche betrachten. Wir erhalten dadurch zwei Sorten von Funktionen.

Die *Exponentialfunktion* zur *Basis*  $a > 0$  ist die Funktion

$$\exp_a: \mathbb{R} \rightarrow ]0, \infty[, \quad x \mapsto a^x.$$

Die *Potenzfunktion* zum *Exponenten*  $\alpha \in \mathbb{R}$  ist die Funktion

$$h_\alpha: ]0, \infty[ \rightarrow ]0, \infty[, \quad x \mapsto x^\alpha.$$

Über diese Funktionen notieren wir weitere Aussagen.

**(5.2) Satz.** *Die Potenzfunktion ist differenzierbar. Die Ableitung an der Stelle  $x$  ist  $(x^\alpha)' = \alpha x^{\alpha-1}$ .*

**(5.3) Satz.** *Die Exponentialfunktion zur Basis  $a$  ist differenzierbar. Die Ableitung ist proportional zur Ausgangsfunktion. Es gilt*

$$\exp'_a(x) = \exp'_a(0) \exp_a(x).$$

Die bislang genannten Eigenschaften sind nicht voneinander unabhängig. Die wichtigsten Eigenschaften sind das Potenzgesetz (5.1.2) und die Ableitungseigenschaft (5.3). Alles Weitere folgt aus einer dieser Eigenschaften.

**(5.4) Beispiel.** Aus den Potenzfunktionen erhält man speziell die  $n$ -ten Wurzeln. Die Funktion  $h_\alpha: x \mapsto x^\alpha$  hat nämlich die Umkehrfunktion  $h_{\frac{1}{\alpha}}$ , denn nach den genannten Regeln gilt

$$(x^\alpha)^{\frac{1}{\alpha}} = x = (x^{\frac{1}{\alpha}})^\alpha.$$

Ist  $\alpha = n$  eine natürliche Zahl, so ist deshalb  $x^{\frac{1}{n}}$  eine  $n$ -te Wurzel von  $x$ .  $\diamond$

## 5.2. Das Potenzgesetz

Wir untersuchen die Konsequenzen des Potenzgesetzes  $a^{x+y} = a^x a^y$ . Da wir die Funktion  $a^x$  noch nicht kennen, untersuchen wir differenzierbare Funktionen  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit der Eigenschaft

$$(5.5) \quad f(x+y) = f(x)f(y)$$

für alle Zahlen  $x$  und  $y$ . Eine Gleichung dieser Art, in der eine Eigenschaft der Funktion  $f$  zum Ausdruck kommt, nennt man eine *Funktionalgleichung*.

Ein Funktion  $f$ , die nur den Wert Null annimmt, erfüllt offenbar (5.5). Diese langweilige Funktion wollen wir ausschließen.

Wir kürzen ab:  $a = f(1)$ . Damit gilt  $f(2) = f(1+1) = f(1)f(1) = a^2$ . Durch vollständige Induktion ergibt sich allgemein für natürliche Zahlen  $n$  die Gleichung  $f(n) = a^n$ . Jede Funktion mit der Eigenschaft (5.5) verallgemeinert also die Potenzfunktion. Wir haben  $f(x) = f(x+0) = f(x)f(0)$ . Wäre  $f(0) = 0$ , so würde auch  $f(x) = 0$  für alle  $x$  gelten. Das hatten wir aber ausgeschlossen. Also ist  $f(0) \neq 0$ , und mit  $f(0) = f(0)f(0)$  folgt dann  $f(0) = 1$ . Wegen  $f(x)f(-x) = f(0) = 1$  erkennen wir, daß  $f(x)$  immer ungleich Null ist, und wegen  $f(x) = f(\frac{x}{2})f(\frac{x}{2}) = f(\frac{x}{2})^2$  sehen wir schließlich, daß immer  $f(x) > 0$  ist. Die beiden Funktionen  $x \mapsto f(a+x)$  und  $x \mapsto f(a)f(x)$  sind wegen (5.5) gleich und haben deshalb auch dieselbe Ableitung. Die Ableitung der ersten Funktion an der Stelle  $x$  ist  $f'(a+x)$ , die der zweiten  $f(a)f'(x)$ . Also gilt

$$f'(a+x) = f(a)f'(x).$$

Im Spezialfall  $x = 0$  erhalten wir

$$f'(a) = f'(0)f(a),$$

die Funktion ist proportional zu ihrer Ableitung. Ist  $f'(0) = 0$ , so folgt  $f'(a) = 0$ ; dann ist  $f$  konstant gleich 1. Diese konstante Funktion erfüllt natürlich auch (5.5).

Sei im folgenden  $f'(0) \neq 0$ . Mit einer reellen Zahl  $\lambda$  definieren wir eine neue Funktion  $g(x) = f(\lambda x)$ . Diese Funktion erfüllt offenbar auch die Funktionalgleichung (5.5). Für ihre Ableitung gilt

$$g'(x) = \lambda f'(\lambda x) = \lambda f'(0)f(\lambda x) = \lambda f'(0)g(x).$$

Wir wählen  $\lambda$  so, daß  $\lambda f'(0) = 1$  ist. Dann ist die Funktion  $g$  gleich ihrer Ableitung. Wenn es also überhaupt eine nicht-konstante Funktion gibt, die (5.5) erfüllt, so auch eine, die gleich ihrer Ableitung ist.

### 5.3. Die Exponentialfunktion

**(5.6) Definition.** Die *Exponentialfunktion* ist die differenzierbare Funktion

$$\exp: \mathbb{R} \rightarrow \mathbb{R}$$

mit den Eigenschaften

- (1)  $\exp'(x) = \exp(x)$
- (2)  $\exp(0) = 1$ .

Die erste Eigenschaft heißt die *Differentialgleichung* der Exponentialfunktion, die zweite die *Anfangsbedingung*.  $\diamond$

Die Exponentialfunktion ist nicht durch eine Formel erklärt, sondern durch eine *Eigenschaft*. Wir werden alsbald sehen, daß es nur eine Funktion mit dieser Eigenschaft gibt. Wir werden außerdem sehen, daß  $\exp$  die Funktionalgleichung

$$(5.7) \quad \exp(x + y) = \exp(x) \exp(y)$$

erfüllt. Es ist üblich,  $\exp(1) = e$  zu setzen und  $\exp(x)$  mit  $e^x$  zu bezeichnen. Man nennt deshalb die Exponentialfunktion salopp auch einfach die *e-Funktion*.

Für die weitere Untersuchung leiten wir aus den genannten Eigenschaften her, daß  $\exp(x)$  niemals gleich Null ist. Die Ableitung der Funktion  $x \mapsto \exp(x) \exp(-x)$  ergibt sich nämlich zu

$$\exp'(x) \exp(-x) + \exp(x) (-\exp'(-x)) = \exp(x) \exp(-x) - \exp(x) \exp(-x) = 0.$$

Also ist diese Funktion konstant, mit dem Wert 1 wegen der Anfangsbedingung.

### 5.4. Die Differentialgleichung $y' = \alpha y$

Sei  $\alpha$  eine reelle Zahl. Wir fragen nach differenzierbaren Funktionen  $f$ , für die immer

$$f'(x) = \alpha f(x)$$

gilt. Wir sagen dann, sie sind *Lösungen* der *Differentialgleichung*  $y' = \alpha y$ . Der folgenden Satz ist der Hauptsatz über diese Differentialgleichung. Im Beweis des Satzes gehen wir davon aus, daß eine Funktion  $\exp$  mit den genannten Eigenschaften vorgegeben ist; sie wird sich durch diesen Satz als eindeutig bestimmt erweisen.

**(5.8) Satz.** *Seien  $\alpha$  und  $\beta$  reelle Zahlen. Die Funktion*

$$u: \mathbb{R} \rightarrow \mathbb{R}, \quad u(x) = \beta \exp(\alpha x)$$

*ist die einzige differenzierbare Funktion  $u: \mathbb{R} \rightarrow \mathbb{R}$  mit den Eigenschaften*

- (1)  $u'(x) = \alpha u(x)$
- (2)  $u(0) = \beta$ .

**BEWEIS.** Nach der Kettenregel und den Eigenschaften (5.6) hat die Funktion  $x \mapsto \beta \exp(\alpha x)$  jedenfalls diese Eigenschaften. Sei  $u$  eine weitere derartige Funktion. Wir haben im letzten Abschnitt gesehen, daß  $\exp(\alpha x) \neq 0$  ist. Wir können deshalb die Funktion

$$v(x) = \frac{u(x)}{\exp(\alpha x)}$$

bilden. Deren Ableitung ergibt sich nach den Ableitungsregeln zu

$$v'(x) = \frac{\exp(\alpha x)u'(x) - u(x)\alpha \exp(\alpha x)}{(\exp(\alpha x))^2}.$$

Wegen der Voraussetzung  $u'(x) = \alpha u(x)$  ist der Zähler gleich Null. Also ist  $v$  eine konstante Funktion. Ihr Wert  $v(0)$  ergibt sich aus der Voraussetzung als 1.  $\square$

### 5.5. Die Funktionalgleichung der Exponentialfunktion

Wir weisen jetzt nach, daß die Exponentialfunktion die Funktionalgleichung (5.7) erfüllt. Dazu betrachten wir die Funktionen

$$x \mapsto \exp(a+x), \quad x \mapsto \exp(a)\exp(x).$$

Beide sind gleich ihrer Ableitung und haben bei Null denselben Wert. Nach (5.8) stimmen sie also überein.

Aus der Funktionalgleichung haben wir früher schon  $\exp(x) > 0$  gefolgert. Wegen  $\exp' = \exp$  sehen wir, daß  $\exp$  streng monoton wachsend ist. Wegen  $\exp(0) = 1$  ist also  $e = \exp(1) > 1$ . Aus

$$\lim_{n \rightarrow \infty} e^n = \infty$$

für eine Zahl  $e > 1$  sehen wir mit Hilfe des Zwischenwertsatzes, daß  $\exp$  alle positiven Zahlen als Funktionswert hat. Das verwenden wir im nächsten Abschnitt.

### 5.6. Der Logarithmus

Die Exponentialfunktion  $\exp: \mathbb{R} \rightarrow \mathbb{R}_+$  hat eine Umkehrfunktion. Diese heißt *natürlicher Logarithmus* und wird mit

$$\log: \mathbb{R}_+ \rightarrow \mathbb{R}, \quad x \mapsto \log(x)$$

bezeichnet. Die Regel über die Ableitung einer Umkehrfunktion liefert

$$\log'(x) = \frac{1}{\exp'(\log(x))} = \frac{1}{\exp(\log(x))} = \frac{1}{x}.$$

Wegen  $\exp(0) = 1$  gilt  $\log(1) = 0$ .

**(5.9) Satz.** *Der natürliche Logarithmus ist die differenzierbare Funktion*

$$\log: ]0, \infty[ \rightarrow \mathbb{R},$$

die durch die Bedingungen

$$\log'(x) = \frac{1}{x}, \quad \log(1) = 0$$

bestimmt ist<sup>7</sup>.  $\square$

Unabhängig von der Exponentialfunktion läßt sich deshalb der Logarithmus als Stammfunktion der Funktion  $\frac{1}{x}$  erklären. Der Hauptsatz der Differential- und Integralrechnung liefert die Existenz einer Stammfunktion für  $\frac{1}{x}$ .

<sup>7</sup>Diese Funktion wird auch mit  $\ln(x)$  bezeichnet

Durch Umkehrung der Funktionalgleichung (5.7) erhalten wir die *Funktionalgleichung des Logarithmus*

$$(5.10) \quad \log(ax) = \log(a) + \log(x).$$

Der Logarithmus verwandelt die Multiplikation in die Addition; das ist eine seiner historischen Bedeutungen (Logarithmentafeln).

Aus dem Satz (5.9) läßt sich (5.10) einsehen, denn beide Seiten haben dieselbe Ableitung und für  $x = 1$  denselben Wert. Mittels der Funktionalgleichung (5.10) überlegt man sich, daß der Logarithmus alle reellen Zahlen als Funktionswert hat. Man kann deshalb die Exponentialfunktion auch als Umkehrfunktion des Logarithmus definieren. Damit hat man eine Methode, die Existenz der Exponentialfunktion zu sichern, da der Logarithmus durch den Hauptsatz der Differential- und Integralrechnung hergestellt werden kann.

### 5.7. Die allgemeine Potenz

Wir *definieren* jetzt für  $a > 0$

$$(5.11) \quad a^x := e^{x \cdot \log(a)}.$$

Da  $\log$  die Umkehrfunktion von  $\exp$  ist, folgt

$$(5.12) \quad \log(a^x) = x \cdot \log(a).$$

Damit erhalten wir

$$(a^x)^y = e^{y \log(a^x)} = e^{yx \log(a)} = a^{xy}.$$

Ferner gilt

$$a^x b^x = e^{x \log a} e^{x \log b} = e^{x(\log a + \log b)} = e^{x \log(ab)} = (ab)^x.$$

Damit haben wir die Potenzgesetze hergeleitet. Es ist ferner

$$a^0 = e^{0 \cdot \log a} = e^0 = 1$$

$$1^x = e^{x \log 1} = e^{x \cdot 0} = e^0 = 1.$$

Für die Ableitung von  $x \mapsto a^x$  gilt

$$(a^x)' = \log(a) e^{x \cdot \log(a)} = \log(a) a^x.$$

Für die Ableitung von  $x \mapsto x^\alpha$  gilt

$$(x^\alpha)' = (e^{\alpha \log x})' = \alpha \frac{1}{x} e^{\alpha \log x} = \alpha x^{\alpha-1}.$$

Für  $a > 1$  ist  $\log(a) > 0$ . Also hat  $a^x$  eine positive Ableitung und ist deshalb streng monoton wachsend. Ist  $\alpha > 0$ , so hat  $x^\alpha$  eine positive Ableitung und ist deshalb streng monoton wachsend. Damit haben wir alle in 5.1 genannten Regeln nachgewiesen.

Sei  $a > 0$ ,  $a \neq 1$ . Die Funktion

$$\exp_a: \mathbb{R} \rightarrow ]0, \infty[, \quad x \mapsto a^x$$

hat eine Umkehrfunktion. Sie wird

$$\log_a: ]0, \infty[ \rightarrow \mathbb{R}, \quad x \mapsto \log_a(x)$$

bezeichnet und *Logarithmus zur Basis a* genannt.

### 5.8. Berechnung der Exponentialfunktion

Aus dem Quotientenkriterium für die Konvergenz von Reihen folgt, daß die unendliche Reihe

$$E(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

für jede reelle Zahl  $x$  konvergiert (siehe (3.25)). Somit erhalten wir eine Funktion  $E: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto E(x)$ . Man kann sich überlegen, daß eine solche Funktion differenzierbar ist und die Ableitung durch gliedweises Differenzieren gebildet werden kann (siehe (4.37)). Damit sieht man sofort, daß die Funktion  $E$  gleich ihrer Ableitung ist. Sie hat den Wert  $E(0) = 1$ . Nach dem Satz (5.8) ist  $E$  die Exponentialfunktion, und wir erhalten die Formel

$$(5.13) \quad \exp(x) = e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Setzen wir darin  $x = 1$ , so erhalten wir eine Reihe zur Berechnung der Eulerschen Zahl  $e$ . Wie groß ist der Fehler, wenn man nach  $n + 1$  Summanden aufhört? Eine Fehlerabschätzung kann durch die Taylorsche Formel gewonnen werden. Einfacher und direkter kann man in diesem Fall so argumentieren: Für  $|x| \leq 1$  ist

$$\begin{aligned} \left| \sum_{k=n+1}^{\infty} \frac{x^k}{k!} \right| &= \frac{|x|^{n+1}}{(n+1)!} \left( 1 + \frac{|x|}{n+2} + \frac{|x|^2}{(n+2)(n+3)} + \dots \right) \\ &\leq \frac{|x|^{n+1}}{(n+1)!} \left( 1 + \frac{1}{2} + \frac{1}{2^2} + \dots \right) \\ &= 2 \frac{|x|^{n+1}}{(n+1)!} \end{aligned}$$

Damit kann man relativ schnell die Zahl  $e$  annähern

$$e = 2,71828\dots$$

Noch eine Bemerkung zur Zahl  $e$ . Die Definition der Ableitung liefert die folgende Gleichungskette

$$\begin{aligned} 1 = \log'(1) &= \lim_{h \rightarrow 0} \frac{\log(1+h) - \log(1)}{h} \\ &= \lim_{h \rightarrow 0} \log(1+h)^{\frac{1}{h}} \\ &= \lim_{u \rightarrow \infty} \log \left( 1 + \frac{1}{u} \right)^u. \end{aligned}$$

Wenden wir auf diese Gleichung die stetige Funktion  $\exp$  an, so erhalten wir

$$e = \lim_{u \rightarrow \infty} \left( 1 + \frac{1}{u} \right)^u.$$

Der rechts stehende Grenzwert tritt bei der sogenannten „stetigen Verzinsung“ auf.

Aus der Exponentialreihe entnimmt man für positive  $x$

$$e^x > \frac{x^{n+1}}{(n+1)!}, \quad \frac{e^x}{x^n} > \frac{x}{(n+1)!}$$

und damit

$$\lim_{x \rightarrow \infty} \frac{e^x}{x^n} = \infty,$$

und dazu sagt man: die Exponentialfunktion wächst schneller als jede Potenz. Eine dazu gleichwertige Aussage ist

$$\lim_{x \rightarrow \infty} \frac{e^x}{x^{-n}} = \lim_{x \rightarrow \infty} \frac{x^n}{e^x} = 0.$$

Aus den eben genannten Wachstumseigenschaften der Exponentialfunktion erhält man die folgenden Grenzwerte:

$$\lim_{x \rightarrow \infty} x^a = \begin{cases} \infty & \text{für } a > 0 \\ 0 & \text{für } a < 0 \end{cases}$$

$$\lim_{x \rightarrow 0} x^a = \begin{cases} \infty & \text{für } a < 0 \\ 0 & \text{für } a > 0 \end{cases}$$

$$\lim_{x \rightarrow \infty} \frac{\log x}{x^a} = 0 \quad \text{für } a > 0$$

$$\lim_{x \rightarrow 0} x^a \log x = 0 \quad \text{für } a > 0$$

### 5.9. Logarithmusreihe und Binomialreihe

Die Reihe

$$L(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} x^n$$

hat nach dem Quotientenkriterium das Konvergenzintervall  $|x| < 1$ . Die gliedweise Ableitung führt auf die geometrische Reihe

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - \dots$$

Also ist  $L(x)$  eine Stammfunktion von  $(1+x)^{-1}$ . Durch Wertevergleich an der Stelle  $x = 1$  erkennt man:

$$(5.14) \quad \log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots$$

Man kann zeigen, daß die Reihe auch noch für  $x = 1$  konvergiert und die Formel

$$\log(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

liefert. Die Logarithmusreihe konvergiert allerdings langsam und ist nicht geeignet, die Logarithmen zu berechnen. Durch Subtraktion der Reihen für  $\log(1+x)$  und  $\log(1-x)$  erhält man

$$\log \frac{1+x}{1-x} = 2 \sum_{n=0}^{\infty} \frac{x^{2n+1}}{2n+1}.$$

Für  $x = \frac{1}{3}$  erhält man daraus zum Beispiel

$$\log(2) = 2 \left( \frac{1}{2} + \frac{1}{3 \cdot 3^3} + \frac{1}{5 \cdot 3^5} + \dots \right)$$

und damit recht schnell  $\log(2) = 0,69314\dots$

Wir definieren für eine reelle Zahl  $s$  und eine natürliche Zahl  $n$  verallgemeinerte Binomialkoeffizienten

$$\binom{s}{n} = \frac{s(s-1)(s-2)\cdots(s-n+1)}{n!}.$$

Ist  $m = s$  eine natürliche Zahl, so erhalten wir die früher definierten Binomialkoeffizienten. Wir setzen noch  $\binom{s}{0} = 1$  fest. Wie früher bestätigt man die Pascal-Formel

$$\binom{s-1}{k} + \binom{s-1}{k-1} = \binom{s}{k}.$$

Mit diesen Zahlen bilden wir die Reihe

$$B_s(x) = \sum_{n=0}^{\infty} \binom{s}{n} x^n.$$

Nach dem Quotientenkriterium konvergiert die Reihe für  $|x| < 1$ , denn für den Quotienten zweier aufeinanderfolgender Glieder gilt

$$\lim_{n \rightarrow \infty} \left| \frac{s-n}{n+1} x \right| = |x|.$$

Wie beim Beweis der binomischen Formel verifiziert man mittels der Pascal-Formel die Relation

$$(1+x)B_{s-1}(x) = B_s(x).$$

Durch gliedweise Differenzieren ergibt sich

$$B'_s(x) = sB_{s-1}(x).$$

Mit den letzten beiden Relationen folgt, daß  $B_s(x)(1+x)^{-s}$  die Ableitung Null hat. Das zeigt die Binomialreihenentwicklung

$$(5.15) \quad (1+x)^s = \sum_{n=0}^{\infty} \binom{s}{n} x^n,$$

die für  $|x| < 1$  besteht.

### 5.10. Wachstumsmodelle

Sei  $y(t)$  der Bestand einer Population zur Zeit  $t$ . Dann ist  $y'(t)/y(t)$  die *relative Wachstumsrate*. In Wachstumsmodellen nimmt man an, daß diese eine Funktion  $k(t, y(t))$  der Zeit und des augenblicklichen Bestandes ist. Im allgemeinen legt man noch einen Anfangswert  $y_0$  zur Zeit  $t_0$  fest.

Der einfachste Ansatz einer konstanten Funktion  $k$  führt auf die Funktion  $y(t) = y_0 e^{k(t-t_0)}$ . Wegen des explosionsartigen Wachstums der Exponentialfunktion kann keine Population auf die Dauer exponentiell zunehmen. Beschränktheit der Ressourcen bremsen das Wachstum, und man muß mathematische Modelle darüber entwickeln. Ein einfaches Modell ist

$$k(t, y) = a - by, \quad a, b > 0.$$



Dieses führt zu der sogenannten *logistischen Differentialgleichung*

$$(5.16) \quad y' = (a - by)y.$$

Dieser Differentialgleichung kann man schon einige *qualitative* Aussagen über eine mögliche positive Lösung entnehmen.

- (1) Eine Zunahme des Bestandes verringert die relative Wachstumsrate. Solange  $y(t) < a/b$  ist, ist  $k(y) > 0$ , also auch  $y'(t) > 0$ ; der Bestand wächst. Für  $y(t) > a/b$  nimmt der Bestand ab.
- (2) Ein beschleunigtes Wachstum ist für  $y' > 0$  und  $y'' > 0$  gegeben. Durch Ableitung der Differentialgleichung erhalten wir  $y'' = (a - 2by)y'$ , so daß also  $y'' > 0$  für  $y(t) < a/2b$  gilt.
- (3) Für sehr kleine Populationen kann man  $y^2$  gegenüber  $y$  vernachlässigen. Dann wird die Differentialgleichung angenähert zu  $y' = ay$ , und wir haben ein exponentielles Wachstum. Ebenso für kleine  $b$ .

Sei  $y(t)$  eine positive Lösung der logistischen Differentialgleichung. Wir betrachten die Funktion  $z(t) = y(t)^{-1}$ . Sie erfüllt die Differentialgleichung

$$(5.17) \quad z' = -az + b.$$

Eine Lösung davon ist die konstante Funktion  $z(t) = b/a$ . Sind  $z_1$  und  $z_2$  Lösungen, so ist  $u(t) = z_1(t) - z_2(t)$  eine Lösung von  $u'(t) = -au(t)$ . Sie hat, wie wir wissen, nur die Lösungen  $u(t) = ce^{-at}$ . Damit erhalten wir durch Zurückrechnen:

**(5.18) Satz.** Die für  $t \geq 0$  positiven Lösungen der logistischen Gleichung sind

$$y(t) = \frac{1}{b/a + ce^{-at}}, \quad c > -\frac{b}{a}.$$

Dieser Funktion sieht man an:

$$(5.19) \quad \lim_{t \rightarrow \infty} y(t) = \frac{a}{b}.$$

Das ist die „Gleichgewichtslage“. Es ist

$$y'(t) = \frac{ace^{-at}}{(b/a + ce^{-at})^2}.$$

Für  $c > 0$  wächst also die Population. Die Nullstellen  $t_0$  der zweiten Ableitung  $y''$  sind durch  $ac = be^{at_0}$  gegeben; es gibt also genau eine, und der Funktionswert an dieser Stelle ist  $y(t_0) = a/2b$ , im Einklang mit unseren qualitativen Überlegungen.

## 6 Trigonometrische Funktionen

### 6.1. Sinus und Cosinus

Die Funktionen Sinus und Cosinus werden, ähnlich wie die Exponentialfunktion, zunächst durch Eigenschaften definiert.

**(6.1) Definition.** Die Funktionen Sinus und Cosinus sind differenzierbare Funktionen

$$\sin: \mathbb{R} \rightarrow \mathbb{R}, \quad \cos: \mathbb{R} \rightarrow \mathbb{R}$$

mit den Eigenschaften

$$(6.2) \quad \sin'(x) = \cos(x), \quad \cos'(x) = -\sin(x)$$

$$(6.3) \quad \cos(0) = 1, \quad \sin(0) = 0.$$

Aus dieser Definition folgt unmittelbar, daß die Funktionen beliebig oft differenzierbar sind. Es gilt ferner

$$(6.4) \quad \cos^2(x) + \sin^2(x) = 1,$$

denn die Ableitung der linken Seite ist  $-2\cos(x)\sin(x) + 2\sin(x)\cos(x) = 0$ ; also handelt es sich um eine konstante Funktion, die wegen (6.3) den Wert 1 hat. Aus (6.4) folgt, daß die Funktionswerte beider Funktionen im Intervall  $[-1, 1]$  liegen und daß  $(\cos(x), \sin(x))$  ein Punkt der Ebene auf dem Kreis vom Radius 1 um den Nullpunkt ist. Geometrisch wird dabei  $x$  als Winkel, gemessen im Bogenmaß gegen den Uhrzeigersinn, zwischen der positiven  $x$ -Achse und dem Halbstrahl vom Nullpunkt durch  $(\cos(x), \sin(x))$  interpretiert. Die mathematische Begründung dafür werden wir nachliefern.

## 6.2. Die Differentialgleichung $y'' = -y$

Die Funktionen  $\sin$  und  $\cos$  haben beide die Eigenschaft, daß ihre zweite Ableitung gleich dem Negativen der Ausgangsfunktion ist. Man sagt, sie genügen der Differentialgleichung  $y'' = -y$ . Der folgende Satz bestimmt alle Funktionen mit dieser Eigenschaft.

**(6.5) Satz.** *Sei  $f: \mathbb{R} \rightarrow \mathbb{R}$  eine zweimal differenzierbare Funktion, die der Gleichung  $f''(x) = -f(x)$  für alle  $x \in \mathbb{R}$  genügt. Dann gibt es reelle Zahlen  $\alpha$  und  $\beta$ , so daß*

$$f(x) = \alpha \sin(x) + \beta \cos(x)$$

*ist. Darin ist  $\beta = f(0)$  und  $\alpha = f'(0)$ .*

BEWEIS. Seien  $\alpha$  und  $\beta$  wie im Satz angegeben durch  $f$  festgelegt. Wir betrachten die Funktion

$$g(x) = f(x) \sin(x) + f'(x) \cos(x).$$

Sie hat die Ableitung

$$f'(x) \sin(x) + f(x) \cos(x) + f''(x) \cos(x) - f'(x) \sin(x),$$

was wegen  $f'' = -f$  gleich Null ist. Also handelt es sich um eine Konstante, die durch Einsetzen von  $x = 0$  zu  $\alpha$  bestimmt wird. Ebenso ergibt sich, daß  $h(x) = f(x) \cos(x) - f'(x) \sin(x) = \beta$  ist. Wegen (6.4) errechnen wir daraus  $g(x) \sin(x) + h(x) \cos(x) = f(x)$ , was nach dem schon Gezeigten die Behauptung des Satzes liefert.  $\square$

Dieser Satz besagt insbesondere, daß durch die Eigenschaften (6.2) und (6.3) die Funktionen  $\sin$  und  $\cos$  eindeutig festgelegt sind.

Die vorstehenden Betrachtungen kann man noch etwas verallgemeinern. Sei  $\omega \neq 0$ . Die Funktionen  $u(t)$ , die der Differentialgleichung des sogenannten *harmonischen Oszillators*

$$u''(t) + \omega^2 u(t) = 0$$

genügen, sind genau die Funktionen der Form

$$u(t) = \alpha \cos(\omega t) + \beta \sin(\omega t), \quad \alpha, \beta \in \mathbb{R}.$$

Zum Beweis setzt man  $w(t) = u(t/\omega)$ . Dann gilt  $w'' = -w$ , und wir können den obigen Satz anwenden.

### 6.3. Additionstheoreme

Die *Additionstheoreme* für  $\sin(x)$  und  $\cos(x)$  sind die im nächsten Satz niedergelegten Relationen.

**(6.6) Satz.** *Für alle  $a, x$  gilt*

$$\cos(a+x) = \cos(a)\cos(x) - \sin(a)\sin(x)$$

$$\sin(a+x) = \sin(a)\cos(x) + \cos(a)\sin(x).$$

*Daraus ergeben sich für  $x = y$  speziell die Relationen*

$$\sin(2x) = 2\sin(x)\cos(x), \quad \cos(2x) = \cos^2(x) - \sin^2(x).$$

BEWEIS. Die Funktionen  $x \mapsto \cos(a+x)$ ,  $x \mapsto \sin(a+x)$  erfüllen die Voraussetzungen des Satzes (6.5). Dieser liefert die Formeln (6.6).  $\square$

Ebenfalls aus Satz (6.5) ergeben sich die folgenden Relationen

$$(6.7) \quad \cos(-x) = \cos(x), \quad \sin(-x) = -\sin(x),$$

denn die Funktionen  $x \mapsto \cos(-x)$  und  $x \mapsto \sin(-x)$  erfüllen auch die Differentialgleichung  $y'' = -y$ . Eine Funktion  $f$  mit  $f(x) = f(-x)$  nennt man *gerade*, eine mit  $f(-x) = -f(x)$  *ungerade*.

### 6.4. Nullstellen und Perioden

Die kleinste positive Nullstelle von  $\sin(x)$  wird mit  $\pi$  bezeichnet. Aus (6.6) sieht man, daß dann  $\frac{\pi}{2}$  die kleinste positive Nullstelle von  $\cos(x)$  ist. Wegen (6.2) ist dann  $\sin(x)$  im Intervall  $[0, \frac{\pi}{2}]$  streng monoton wachsend. Aus (6.4) folgt wegen  $\sin^2(\frac{\pi}{2}) = 1$  dann

$$\sin\left(\frac{\pi}{2}\right) = 1.$$

Aus (6.6) entnimmt man die Relation

$$\cos(\pi) = -1.$$

Indem man diese Ergebnisse in die Additionstheoreme einsetzt, erhält man nacheinander die folgenden Relationen.

$$\cos\left(x \pm \frac{\pi}{2}\right) = \mp \sin(x)$$

$$\sin\left(x \pm \frac{\pi}{2}\right) = \pm \cos(x)$$

$$\cos(x + \pi) = -\cos(x)$$

$$\sin(x + \pi) = -\sin(x)$$

$$\cos(x + 2\pi) = \cos(x)$$

$$\sin(x + 2\pi) = \sin(x)$$

Die Nullstellenmenge von  $\cos$  ist  $\{\frac{\pi}{2} + \pi n \mid n \in \mathbb{Z}\}$ , diejenige von  $\sin$  ist  $\{\pi n \mid n \in \mathbb{Z}\}$ .

Eine Funktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  heißt *periodisch* mit der Periode  $a$ , wenn für alle  $x$  die Gleichung  $f(x+a) = f(x)$  gilt. Die Funktionen  $\sin$  und  $\cos$  sind also periodisch mit der Periode  $2\pi$ . Periodische Funktionen sind das mathematische Werkzeug für die Beschreibung von Schwingungsvorgängen.

### 6.5. Reihenentwicklung

Die Existenz von Funktionen mit den in (6.1) niedergelegten Eigenschaften kann man durch Reihenentwicklungen beweisen. Es gilt nämlich

$$(6.8) \quad \cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

$$(6.9) \quad \sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

Nach dem Quotientenkriterium konvergieren die Reihen für alle  $x$ . Durch gliedweises Differenzieren verifiziert man sodann für die dadurch definierten Funktionen die Relationen (6.2).

Damit die vorstehenden Überlegungen vollständig sind muß man noch beweisen, daß  $\cos(x)$  überhaupt positive Nullstellen hat.

### 6.6. Tangens und Cotangens

Die Funktionen *Tangens*  $\tan(x)$  und *Cotangens*  $\cot(x)$  werden durch

$$\tan(x) = \frac{\sin(x)}{\cos(x)}, \quad \cot(x) = \frac{\cos(x)}{\sin(x)}$$

definiert. Ihr Definitionsbereich sind die reellen Zahlen ohne die Nullstellen des jeweiligen Nenners. Aus den Eigenschaften von  $\sin$  und  $\cos$  leitet man die folgenden Aussagen über diese Funktionen her.

Beide Funktionen haben die Periode  $\pi$  und sind ungerade. Ihre Ableitungen sind

$$(6.10) \quad \tan'(x) = 1 + \tan^2(x), \quad \cot'(x) = -1 - \cot^2(x).$$

Deshalb ist der Tangens im Intervall  $]-\frac{\pi}{2}, \frac{\pi}{2}[$  streng monoton wachsend und hat als Wertebereich die Menge der reellen Zahlen. Der Cotangens ist im Intervall  $]0, \pi[$  streng monoton fallend und der Wertebereich ist ebenfalls die Menge der reellen Zahlen. Es gelten die *Additionstheoreme*

$$(6.11) \quad \tan(x+y) = \frac{\tan(x) + \tan(y)}{1 - \tan(x)\tan(y)}, \quad \cot(x+y) = \frac{\cot(x)\cot(y) - 1}{\cot(x) - \cot(y)}.$$

### 6.7. Umkehrfunktionen

Die Umkehrfunktionen der trigonometrischen Funktionen werden *Arcus-Funktionen* genannt. Die Umkehrfunktion von  $\tan: ]-\frac{\pi}{2}, \frac{\pi}{2}[ \rightarrow \mathbb{R}$  heißt *arcus tangens*

$$\arctan: \mathbb{R} \rightarrow ]-\frac{\pi}{2}, \frac{\pi}{2}[.$$

Die Umkehrfunktion von  $\cot: ]0, \pi[ \rightarrow \mathbb{R}$  heißt *arcus cotangens*

$$\operatorname{arccot}: \mathbb{R} \rightarrow ]0, \pi[.$$

Der Sinus ist im Intervall  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  streng monoton wachsend und hat die Umkehrfunktion *arcus sinus*

$$\arcsin: [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}],$$

und der Cosinus ist im Intervall  $[0, \pi]$  streng monoton fallend und hat die Umkehrfunktion *arcus cosinus*

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

Für die Ableitungen dieser Funktionen erhält man aus der Umkehrregel

$$\begin{aligned} \arcsin'(x) &= \frac{1}{\sqrt{1-x^2}} \\ \arccos'(x) &= -\frac{1}{\sqrt{1-x^2}} \\ \arctan'(x) &= \frac{1}{1+x^2} \\ \operatorname{arccot}'(x) &= -\frac{1}{1+x^2}. \end{aligned}$$

Aus dem Zusammenhang zwischen Sinus und Cosinus und den Periodizitäten erhält man Umkehrfunktionen zu anderen Definitionsintervallen.

## 7 Integralrechnung

### 7.1. Das Integral

Der Flächeninhalt eines Rechtecks mit den Kantenlängen  $a$  und  $b$  ist das Produkt  $ab$ , wenn man dem Quadrat mit der Seitenlänge 1 den Inhalt 1 zubilligt.

Ein anschauliches Motiv für die Integralrechnung ist die Frage nach dem Flächeninhalt krummlinig begrenzter Figuren, etwa eines Kreises. Folgende Situation ist dabei für die rechnerische Behandlung typisch. Sei  $f: [a, b] \rightarrow \mathbb{R}$  eine Funktion mit positiven Werten. Die Fläche *unter dem Graphen* ist die Punktmenge

$$\{(x, y) \mid a \leq x \leq b, 0 \leq y \leq f(x)\}.$$

Der Flächeninhalt dieser Menge soll aus der Kenntnis der Funktion bestimmt werden. Die Integralrechnung löst diese Aufgabe.

Um einfache Rechenregeln zu gewinnen, hat sich herausgestellt, daß für die Fläche *über dem Graphen* einer Funktion mit negativen Werten

$$\{(x, y) \mid a \leq x \leq b, 0 \geq y \geq f(x)\}$$

besser das Negative des Flächeninhalts betrachtet wird.

Da die Zuordnungsvorschrift einer Funktion völlig beliebig ist, kann man sich „wilde“ Funktionen mit vielen Sprüngen ausdenken, für die die Menge unter dem Graphen schwerlich einen sinnvollen Flächeninhalt hat. Es ist deshalb für die Integralrechnung nötig, der Kreis der Funktionen geeignet einzuschränken; die zugelassenen werden dann

als integrierbar bezeichnet. Dazu sollen die stetigen Funktionen gehören und die so gleich zu definierenden Treppenfunktionen. Man möchte gewisse Sprünge zulassen, weil die Fläche unter dem Graphen der Funktion  $f: [0, 2] \rightarrow \mathbb{R}$ , die auf  $[0, 1]$  konstant gleich 1 und auf  $]1, 2]$  konstant gleich zwei ist, aus zwei Rechtecken besteht, und dafür ist der Flächeninhalt ja gerade leicht zu erklären.

Die Grundidee zur Definition des Flächeninhalts besteht darin, krumme Flächen möglichst genau durch Rechtecke anzunähern. Sie führt zu der folgenden Begriffsbildung.

Eine *Zerlegung*  $\mathfrak{Z}$  des Intervalls  $[a, b]$  ist ein System  $(x_0, x_1, \dots, x_m)$  von Teilungspunkten

$$a = x_0 < x_1 < x_2 < \dots < x_m = b.$$

Eine *Treppenfunktion*  $t$  zu einer solchen Zerlegung  $\mathfrak{Z}$  ist eine Funktion, die auf den offenen Intervallen  $]x_{j-1}, x_j[$ ,  $1 \leq j \leq m$  jeweils einen konstanten Wert  $c_j$  hat. An den Stellen  $x_0, \dots, x_m$  kann sie irgendwelche Werte haben, das ist uns (und dem Integral) egal. Das *Integral* der Treppenfunktion wird als die Summe

$$\sum_{j=1}^m c_j (x_j - x_{j-1})$$

erklärt. Es handelt sich offenbar um eine Summe von Rechtecksinhalten, und zwar positiv gezählt, wenn  $c_j > 0$  ist, und sonst negativ. Das so definierte Integral einer Treppenfunktion wird mit

$$\int_a^b t(x) dx$$

bezeichnet. Seiner Definition nach hängt es zunächst nicht nur von der Funktion sondern scheinbar auch von der Zerlegung ab; wir unterdrücken jedoch die Zerlegung in der Notation.

Sind  $f, g: A \rightarrow \mathbb{R}$  Funktionen, so schreiben wir  $f \leq g$ , wenn für alle  $a \in A$  die Ungleichung  $f(a) \leq g(a)$  gilt.

Wir nähern nun Funktionen durch Treppenfunktionen an und definieren damit Integrierbarkeit und Integral.

**(7.1) Definition.** Eine Funktion  $f: [a, b] \rightarrow \mathbb{R}$  heißt *integrierbar*, wenn es zu jedem  $\varepsilon > 0$  eine *obere Treppenfunktion*  $t_o$  zu  $f$  und eine *untere Treppenfunktion*  $t_u$  zu  $f$  so gibt, daß  $t_u \leq f \leq t_o$  ist und

$$\left| \int_a^b t_o(x) dx - \int_a^b t_u(x) dx \right| < \varepsilon$$

gilt. Ist  $f$  integrierbar, so hat die Menge aller Zahlen der Form

$$\int_a^b t_o(x) dx, \quad t_o \leq f \quad \text{Treppenfunktion}$$

eine obere Schranke und die kleinste darunter wird das *Integral*

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

von  $f: [a, b] \rightarrow \mathbb{R}$  über das Intervall  $[a, b]$  genannt. Wir nennen dabei  $a$  die *untere* und  $b$  die *obere Integrationsgrenze*.  $\diamond$

Ähnlich wie beim Laufindex des Summenzeichens darf auch im Integralsymbol der Buchstabe  $x$  durch irgendeinen anderen, gerade zweckmäßigen, ersetzt werden, etwa

$$\int_a^b f(x) dx = \int_a^b f(t) dt.$$

(Nicht zweckmäßig in dieser Formel sind offenbar die Buchstaben  $d, f, a, b$  anstelle von  $x$ .) Die Integraldefinition ist theoretisch. Aus ihr lassen sich aber nicht allzu schwer Rechenregeln herleiten, die dann alsbald in vielen Fällen zu praktischen und konkreten Resultaten führen. Treppenfunktionen sind übrigens im Sinne der Definition (7.1) integrierbar, und das Integral stimmt mit der früheren Festlegung durch eine Rechtecksumme überein.

Es ist üblich, im Fall  $b < a$  zu definieren

$$\int_a^b f = - \int_b^a f.$$

## 7.2. Das Rechnen mit Integralen

Ist  $f: [a, b] \rightarrow \mathbb{R}$  eine Funktion, so liefert dieselbe Zuordnungsvorschrift  $f$  auch auf jeder Teilmenge  $A \subset [a, b]$  eine neue Funktion. Im allgemeinen sagt man dann, daß die neue Funktion durch *Einschränkung* des Definitionsbereichs entsteht. Wir wollen die eingeschränkten Funktionen aber im folgenden mit demselben Symbol  $f$  bezeichnen.

Wir sammeln zunächst Aussagen darüber, welche Funktionen integrierbar sind. Dazu vorher noch einige Bezeichnungen. Mit  $|f|$  bezeichnen wir die Funktion  $x \mapsto |f(x)|$ ; mit  $\max(f, g)$  die Funktion  $x \mapsto \max(f(x), g(x))$ ; analog  $\min(f, g)$ ; schließlich sei  $f_+ : x \mapsto \max(f(x), 0)$  und  $f_- : x \mapsto \min(f(x), 0)$ .

**(7.2) Satz.** *Stetige und monotone Funktionen sind integrierbar. Treppenfunktionen sind integrierbar. Eine Funktion  $f: [a, b] \rightarrow \mathbb{R}$  ist genau dann integrierbar, wenn die Einschränkungen auf  $[a, c]$  und  $[c, b]$  integrierbar sind. Mit  $f$  und  $g$  sind auch für beliebige reelle Zahlen  $\lambda$  und  $\mu$  die Funktion  $\lambda f + \mu g$  integrierbar, sowie  $f_+$ ,  $f_-$ ,  $|f|$ ,  $\max(f, g)$  und  $\min(f, g)$ .  $\square$*

### (7.3) Eigenschaften des Integrals.

- (1)  $\int_a^b f = \int_a^c f + \int_c^b f$  für  $a \leq c \leq b$ .
- (2) Aus  $f \leq g$  folgt  $\int_a^b f \leq \int_a^b g$ .
- (3) Sei  $f = c$  konstant. Dann gilt  $\int_a^b f = c \cdot (b - a)$ .
- (4)  $\int_a^b (\lambda f + \mu g) = \lambda \int_a^b f + \mu \int_a^b g$ .
- (5)  $|\int_a^b f| \leq \int_a^b |f|$ .

Wir geben diesen Eigenschaften Namen. (1) Intervall-Additivität. (2) Monotonie. (3) Normierung. (4) Linearität. (5) Dreiecksungleichung.

### 7.3. Mittelwertsatz der Integralrechnung.

Der in der Überschrift genannte Satz ist:

**(7.4) Satz.** Sei  $f$  eine stetige und  $p$  eine integrierbare Funktion auf  $[a, b]$  und gelte  $p \geq 0$ . Dann existiert ein  $\xi \in [a, b]$ , so daß

$$\int_a^b f(x)p(x) dx = f(\xi) \int_a^b p(x) dx.$$

Insbesondere ist

$$\int_a^b f(x) dx = f(\xi)(b - a)$$

für ein geeignetes  $\xi \in [a, b]$ .

BEWEIS. Sei  $m$  das Minimum und  $M$  das Maximum von  $f$  auf  $[a, b]$ . Dann ist  $mp \leq fp \leq Mp$  und deshalb gilt  $m \int_a^b p \leq \int_a^b fp \leq M \int_a^b p$ . Wir wenden nun den Zwischenwertsatz auf die Funktion  $x \mapsto f(x) \cdot \int_a^b p$  an.  $\square$

### 7.4. Der Hauptsatz

Der Hauptsatz der Differential- und Integralrechnung verbindet Differentiation und Integration und besagt, grob gesprochen, daß diese beiden Prozesse zueinander invers sind. Wir verwenden immer ein Intervall  $J$  mit mehr als einem Punkt.

**(7.5) Definition.** Eine Funktion  $F: J \rightarrow \mathbb{R}$  heißt *Stammfunktion* von  $f: J \rightarrow \mathbb{R}$ , wenn  $F$  die Ableitung  $f$  hat,  $F' = f$ . Wir wissen schon: Zwei Stammfunktionen unterscheiden sich um eine Konstante.  $\diamond$

**(7.6) Hauptsatz.** Sei  $f: J \rightarrow \mathbb{R}$  eine auf dem Intervall  $J$  stetige Funktion. Sei  $a \in J$  gewählt. Dann ist die durch

$$F(x) = \int_a^x f(t) dt$$

definierte Funktion  $F: J \rightarrow \mathbb{R}$  eine Stammfunktion von  $f$ . Ist  $F$  irgendeine Stammfunktion von  $f$ , so gilt

$$\int_a^x f(t) dt = F(x) - F(a) =: [F]_a^x.$$

BEWEIS. Sei  $g(x) = \int_a^x f(t) dt$ . Dann ist nach dem Mittelwertsatz der Integralrechnung

$$g(x+h) - g(x) = \int_x^{x+h} f(t) dt = f(x + \delta_h \cdot h) \cdot h$$

mit einer Zahl  $\delta_h \in [0, 1]$ . Wegen der Stetigkeit von  $f$  ist  $\lim_{h \rightarrow 0} f(x + \delta_h \cdot h) = f(x)$ . Das zeigt  $g'(x) = f(x)$ , das heißt  $g(x)$  ist eine Stammfunktion von  $f$ .

Ist  $F$  eine weitere Stammfunktion, so sehen wir, daß  $g$  und  $x \mapsto F(x) - F(a)$  zwei Stammfunktionen sind, die bei  $x = a$  denselben Wert haben, also überhaupt gleich sind.  $\square$



Wegen des Hauptsatzes nennt man eine Stammfunktion von  $f$  auch ein *unbestimmtes Integral* und bezeichnet sie durch

$$\int f = \int f(t) dt.$$

So hat zum Beispiel ein Polynom  $f(x) = \sum_{k=0}^n a_k x^k$  das unbestimmte Integral

$$\int f(t) dt = a + \sum_{k=0}^n \frac{a_k}{k+1} x^{k+1}.$$

**(7.7) Beispiel.** Sei  $u(x)$  auf dem Intervall  $[a, b]$  positiv. Dann ist  $u'(x)/u(x)$  nach der Kettenregel die Ableitung von  $\log u(x)$ . Folglich ist

$$\int_a^b \frac{u'(x)}{u(x)} dx = \log u(b) - \log u(a) = \log \frac{u(b)}{u(a)}.$$

Hierbei haben wir benutzt, daß immer  $f$  eine Stammfunktion von  $f'$  ist. Hiermit kann man zum Beispiel

$$\int_a^b \frac{x^3}{x^4 + 1} dx$$

berechnen, weil im Zähler bis auf den Faktor 4 die Ableitung des Nenners steht.  $\diamond$

## 7.5. Anwendungen des Hauptsatzes

Der Hauptsatz erlaubt, Differentiationsregeln in Integrationsregeln zu verwandeln.

**(7.8) Transformationsformel.** Ist  $f: J \rightarrow \mathbb{R}$  eine stetige Funktion und  $\varphi: [a, b] \rightarrow J$  stetig differenzierbar, so gilt

$$\int_a^b f(\varphi(t)) \cdot \varphi'(t) dt = \int_{\varphi(a)}^{\varphi(b)} f(u) du.$$

BEWEIS. Wir ersetzen in beiden Seiten der behaupteten Gleichung  $b$  durch  $x$  und betrachten sie als Funktionen von  $x$ . Für  $x = a$  haben sie denselben Wert. Beide Seiten haben auch dieselbe Ableitung; links folgt das direkt aus dem Hauptsatz; rechts wende man bitte noch die Kettenregel an.  $\square$

**(7.9) Beispiel.** Leichte Anwendungen der Transformationsformel sind

$$\begin{aligned} \int_a^b f(t+c) dt &= \int_{a+c}^{b+c} f(x) dx \\ c \int_a^b f(ct) dt &= \int_{ac}^{bc} f(x) dx. \end{aligned}$$

**(7.10) Beispiel.** Manchmal ist es möglich, einer Funktion  $g(t)$  die Form  $f(\varphi(t))\varphi'(t)$  zu geben. Dann ist  $\int_a^b g(t) dt = \int_{\varphi(a)}^{\varphi(b)} f(u) du$ . Für  $f(y) = e^{-y}$  und  $u(x) = \frac{1}{2}x^2$  ist  $x e^{-\frac{1}{2}x^2} = f(u(x)) \cdot u'(x)$  und deshalb gilt

$$\int_a^b x e^{-\frac{1}{2}x^2} dx = \int_{u(a)}^{u(b)} e^{-y} dy = e^{-\frac{1}{2}a^2} - e^{-\frac{1}{2}b^2},$$

unter Benutzung des Integral der Exponentialfunktion nach dem Hauptsatz.  $\diamond$

**(7.11) Beispiel.** Sei  $u: [\alpha, \beta] \rightarrow [a, b]$  eine bijektive stetig differenzierbare Funktion. Dann liefert die Substitutionsregel

$$\int_a^b f(x) dx = \int_\alpha^\beta f(u(t)) \cdot u'(t) dt.$$

Hierbei darf eventuell  $\beta < \alpha$  sein. Man merkt sich das so:  $x$  wird durch  $x = u(t)$  ersetzt und  $dx$  durch  $dx = u'(t) dt$ , denn letzteres stimmt ja optisch gut mit  $\frac{dx}{dt} = u'(t)$  überein.

Im Integral

$$\int_{-1}^1 \sqrt{1-x^2} dx$$

setzen wir  $x = \sin t$  und erhalten nach der eben erläuterten Vorschrift

$$\int_{-\pi/2}^{\pi/2} \sqrt{1-\sin^2 t} \cos t dt = \int_{-\pi/2}^{\pi/2} \cos^2 t dt.$$

Dieses Integral ist gleich  $\pi/2$ , nach dem Beispiel über partielle Integration. Nach der geometrischen Interpretation des Integral als Flächeninhalt hat also der Halbkreis vom Radius 1 den Flächeninhalt  $\pi/2$ .

**(7.12) Partielle Integration.** Sind  $f$  und  $g$  auf dem Intervall  $[a, b]$  stetig differenzierbare Funktionen, so gilt:

$$\int_a^b f(t)g'(t) dt = [f \cdot g]_a^b - \int_a^b f'(t)g(t) dt.$$

BEWEIS. Wir ersetzen  $b$  durch  $x$  und differenzieren nach  $x$ . Die Behauptung ergibt sich aus der Produktregel.  $\square$

**(7.13) Beispiel.** Wir setzen  $f(x) = \cos(x)$  und  $g(x) = \sin(x)$ . Mittels  $\sin^2 = 1 - \cos^2$  erhalten wir aus der partiellen Integration

$$\int_a^b \cos^2(x) dx = [\cos x \cdot \sin x]_a^b + \int_a^b \sin^2 x dx$$

die Gleichung

$$2 \int_a^b \cos^2 x dx = b - a + [\sin x \cdot \cos x]_a^b.$$

## 7.6. Kurvenlänge

Eine Funktion  $w: [a, b] \rightarrow \mathbb{R}^n$  heißt *Kurve* oder *Weg* im  $\mathbb{R}^n$  mit Parameterintervall  $[a, b]$ . Wir schreiben

$$w(t) = (w_1(t), \dots, w_n(t)).$$

Sind alle  $w_i(t)$  stetig differenzierbar, so heißt der Weg stetig differenzierbar. Den Vektor der Ableitungen

$$w'(t) := (w'_1(t), \dots, w'_n(t))$$

bezeichnen wir als die Ableitung von  $w$  und auch als den *Geschwindigkeitsvektor* von  $w$  zur Zeit  $t$ . Das Integral

$$\int_a^b \|w'(t)\| dt$$

bezeichnen wir als die *Länge* des Weges oder der Kurve.

**(7.14) Beispiel.** Die Kurve

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

durchläuft genau einmal den Einheitskreis. Als Länge ergibt sich

$$\int_0^{2\pi} \cos^2 t + \sin^2 t dt = \int_0^{2\pi} 1 dt = 2\pi.$$

Dieselbe Rechnung zeigt auch, daß die Zahl  $t$  in  $\cos t$  wirklich das Bogenmaß ist.  $\diamond$

### 3 Analytische Geometrie. Lineare Algebra

#### 1 Lineare Gleichungen

##### 1.1. Gleichungssysteme

Wir behandeln Gleichungen in mehreren Unbekannten. Die Anzahl der Unbekannten wollen wir nicht festlegen. Deshalb brauchen wir eine systematische Notation.

Der *n-dimensionalen Zahlenraum*  $\mathbb{R}^n$  ist die Menge aller geordneten Systeme  $(a_1, a_2, \dots, a_n)$  von reellen Zahlen  $a_j$ . Ein solches System von reellen Zahlen bezeichnet man mit dem Kunstwort *n-Tupel* (in einer glücklichen Stunde aus Tripel, Quadrupel, Quintupel abstrahiert).

Entsprechend bezeichnen wir Unbekannte mit  $x_1, x_2, \dots, x_n$ . Ein Ausdruck der Form

$$(1.1) \quad a_1x_1 + a_2x_2 + \dots + a_nx_n = b$$

mit reellen Zahlen  $a_j$  und Unbekannten  $x_j$  heißt *lineare Gleichung*. Wir wollen mehrere solcher Gleichungen gleichzeitig betrachten und brauchen auch dafür eine systematische Notation.

**(1.2) Definition.** Ein *lineares Gleichungssystem* aus  $k$  Gleichungen in  $n$  Unbekannten  $x_1, \dots, x_n$  besteht aus  $k$  Gleichungen der Form (1.1)

$$(1.3) \quad \begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + \dots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \dots + & a_{2n}x_n & = & b_2 \\ \dots & & \dots & & \dots & & \dots \\ a_{k1}x_1 & + & a_{k2}x_2 & + \dots + & a_{kn}x_n & = & b_k. \end{array}$$

Darin sind die  $a_{ij}$  und die  $b_i$  reelle Zahlen. An den  $a$ -Zahlen hängt ein *Doppelindex*; zur Verdeutlichung könnten wir auch  $a_{i,j}$  schreiben. Wir nennen  $i$  den *Zeilenindex* und  $j$  den *Spaltenindex* von  $a_{ij}$ . Der Zeilenindex gibt an, in welcher Gleichung wir uns befinden; der Spaltenindex gibt an, vor welcher Unbekannten die Zahl steht. Die  $a_{ij}$  dürfen auch Null sein. Ein Gleichungssystem (1.3) heißt *homogen*, wenn alle  $b_j$  gleich Null sind.  $\diamond$

In einem konkret gegebenen System fallen die Indices natürlich weg, da sie nur der abstrakten theoretischen Notation dienen.

**(1.4) Beispiel.** Ein System von 5 Gleichungen ( $G_1$  bis  $G_5$ ) in 4 Unbekannten kann etwa so aussehen:

$$\begin{array}{lclclclcl} G_1 : & x_1 & + & 2x_2 & - & x_3 & - & x_4 & = & 0 \\ G_2 : & x_1 & & & + & 2x_3 & & & = & 4 \\ G_3 : & x_1 & & & + & 2x_3 & & & = & 4 \\ G_4 : & x_1 & & & + & 2x_3 & & & = & 5 \\ G_5 : & & & & & x_3 & + & 2x_4 & = & 2. \end{array}$$

Wegen  $0 \cdot x_j = 0$  haben wir die Stellen mit  $a_{ij} = 0$  weggelassen. Es ist zunächst nicht verboten, daß gewisse Gleichungen überflüssig sind (wie  $G_3$ ) oder sich widersprechen ( $G_3$  und  $G_4$ ).  $\diamond$

**Hauptproblem.** Bestimmung der Lösungsgesamtheit (der Lösungsmenge) des Systems (1.3). Die Lösungsmenge besteht aus den  $n$ -Tupeln  $(x_1, \dots, x_n)$  reeller Zahlen,

für die alle  $k$  Gleichungen erfüllt sind. Die Lösungsmenge ist also eine Teilmenge des Zahlenraumes  $\mathbb{R}^n$ .  $\diamond$

Wie wir schon bei der Betrachtung von Geraden in der Ebene in I.3 gesehen haben, kann die Lösungsmenge leer sein, oder aus einem, oder aus unendlich vielen Punkten bestehen.

## 1.2. Matrizen

Ein System (1.3) ist durch die Zahlen  $a_{ij}$  und  $b_j$  bestimmt. In vielen Bereichen der Mathematik treten ähnliche Zahlenschemata auf. Wir betrachten sie deshalb für sich genommen.

**(1.5) Definition.** Eine  $k \times n$ -Matrix ist eine Vorschrift, die jedem Paar  $(i, j)$  natürlicher Zahlen im Bereich  $1 \leq i \leq k$ ,  $1 \leq j \leq n$  eine Zahl  $a_{ij}$  zuordnet. Wir notieren eine solche Matrix als rechteckiges Schema

$$(1.6) \quad A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & a_{k3} & \dots & a_{kn} \end{pmatrix}.$$

Darin heißt  $a_{ij}$  der *Eintrag* oder das *Element* der Matrix an der Stelle  $(i, j)$  mit Zeilenindex  $i$  und Spaltenindex  $j$ . Wir nennen

$$(a_{i1}, a_{i2}, \dots, a_{in})$$

die  $i$ -te *Zeile* der Matrix und

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{kj} \end{pmatrix}$$

die  $j$ -te *Spalte* der Matrix. Die Zeile können wir also als  $1 \times n$ -Matrix auffassen und die Spalte als  $k \times 1$ -Matrix. In Kurzform schreiben wir die Matrix auch

$$(a_{ij}), \quad (a_{ij} \mid 1 \leq i \leq k, 1 \leq j \leq n)$$

oder in ähnlicher Weise.  $\diamond$

In Bezug auf das Gleichungssystem (1.3) heißt  $A = (a_{ij})$  die *Koeffizientenmatrix* des Systems. Fügen wir die  $b_j$  als weitere Spalte hinzu

$$\left( \begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & a_{k3} & \dots & a_{kn} & b_k \end{array} \right),$$

so sprechen wir von der *erweiterten Matrix des Gleichungssystems*.

**(1.7) Beispiel.** Das System (1.4) hat die erweiterte Matrix

$$\left( \begin{array}{cccc|c} 1 & 2 & -1 & -1 & 0 \\ 1 & 0 & 2 & 0 & 4 \\ 1 & 0 & 2 & 0 & 4 \\ 1 & 0 & 2 & 0 & 5 \\ 0 & 0 & 1 & 2 & 2 \end{array} \right).$$

Wir unterscheiden und bezeichnen Matrizen nach ihrer äußeren Form. Eine  $n \times n$ -Matrix heißt *quadratisch*. In einer quadratischen  $n \times n$ -Matrix  $(a_{ij})$  bilden die Einträge  $a_{11}, a_{22}, \dots, a_{nn}$  die *Hauptdiagonale* der Matrix. Eine  $n \times n$ -Matrix heißt *Diagonalmatrix*, wenn außerhalb dieser Diagonalen die Einträge alle Null sind. Die Diagonalmatrix mit lauter Einsen auf der Diagonale werde *Einheitsmatrix* genannt und mit  $E_n$  bezeichnet. Eine Matrix, die unterhalb der Diagonalen nur Nullen hat, heißt *obere Dreiecksmatrix*. Analog wird eine *untere Dreiecksmatrix* erklärt. Beispielsweise ist

$$\begin{pmatrix} 1 & 5 & 3 \\ 0 & 4 & 2 \\ 0 & 0 & 6 \end{pmatrix}$$

eine obere  $3 \times 3$ -Dreiecksmatrix. Werden in einer Matrix  $A$  Zeilen und Spalten vertauscht, so entsteht die *transponierte Matrix*  $A^t$  von  $A$ ; ist  $A = (a_{ij})$  eine  $m \times n$ -Matrix, so ist  $A^t$  eine  $n \times m$ -Matrix mit  $a_{ji}$  an der Stelle  $(i, j)$ .

Wir wenden diese Begriffsbildung in dem folgenden Satz an.

**(1.8) Satz.** *Ein Gleichungssystem, dessen Koeffizientenmatrix eine obere Dreiecksmatrix mit von Null verschiedenen Einträgen auf der Diagonale ist, hat immer eine eindeutig bestimmte Lösung.*

BEWEIS. Es genügt, das Prinzip im Fall  $n = 3$  zu erläutern. Dann sieht nämlich das System so aus:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{33}x_3 &= b_3. \end{aligned}$$

Nach Voraussetzung ist  $a_{33} \neq 0$ . Also läßt sich aus der letzten Gleichung  $x_3$  ausrechnen. Den Wert von  $x_3$  setzen wir in die vorletzte Gleichung ein und rechnen dann  $x_2$  aus. Die Werte für  $x_2$  und  $x_3$  setzen wir in die erste Gleichung ein und rechnen  $x_1$  aus. Im allgemeinen Fall handelt man sich so von unten nach oben durch.  $\square$

### 1.3. Zeilenstufenmatrizen

Das einfache Verfahren des letzten Satzes läßt sich auch noch bei etwas allgemeineren Matrizen durchführen. Das besprechen wir jetzt.

**(1.9) Definition.** Eine Matrix der folgenden Form heißt *Zeilenstufenmatrix*.

$$\begin{array}{cccc|cccc|cccc}
 0 & \dots & 0 & 1 & ??? & 0 & ??? & 0 & ??? & 0 & ??? & 0 & ??? \\
 & & & 0 & & 1 & ??? & 0 & ??? & 0 & ??? & 0 & ??? \\
 & & & & & 0 & & 1 & ??? & 0 & ??? & 0 & ??? \\
 & & & & & & & 0 & & \cdot & \dots & \cdot & \dots \\
 & & & & & & & & & 0 & ??? & 0 & ??? \\
 & & & & & & & & & 1 & ??? & 0 & 0 \\
 & & & & & & & & & 0 & & \cdot & \\
 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{array}$$

$j_1$                    $j_2$                                    $j_k$

An den mit ? bezeichneten Stellen stehen irgendwelche Einträge. Alle freigelassenen Stellen unterhalb der „Treppe“ sind mit Nullen gefüllt. Diese Stellen müssen aber nicht wirklich vorhanden sein, wie die folgenden Beispiele von Zeilenstufenmatrizen verdeutlichen. In den Spalten  $j_1, \dots, j_k$  sind die „Stufen“ der Matrix.  $\diamond$

Zwei Beispiele von Zeilenstufenmatrizen.

$$\begin{pmatrix} 1 & 0 & 3 & 0 & 0 & 6 & 3 \\ 0 & 1 & 2 & 7 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 & -1 & 8 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 1 & 2 & 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 0 & 1 & 0 & 4 & 6 \\ 0 & 0 & 0 & 0 & 0 & 1 & 3 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Die erste hat ihre Stufen in den Spalten 1,2,5; die zweite in den Spalten 3,5,6. Für ein Gleichungssystem, dessen Koeffizientenmatrix Zeilenstufenform hat, kann die Lösungsmenge ebenfalls leicht bestimmt werden. Da die Lösungsmenge im allgemeinen nicht nur aus einem Element besteht, sondern unendlich ist, müssen wir uns darüber verständigen, unter welchen Umständen wir die Lösungsmenge „kennen“.

**(1.10) Satz.** Sei ein Gleichungssystem gegeben, dessen Koeffizientenmatrix Zeilenstufenform hat. Hat die Matrix eine Zeile, die nur aus Nullen besteht, so ist das System unlösbar, wenn die erweiterte Matrix in dieser Zeile einen von Null verschiedenen Eintrag hat. Andernfalls kann man die Werte der Unbekannten an den Positionen, die keine Stufen sind, beliebig vorgeben (diese heißen dann freie Parameter), während die Unbekannte, die zur Stufe  $j$  gehört, sich durch die Zeile  $j$  dann durch die freien Parameter ausrechnen läßt.  $\square$

**(1.11) Beispiel.** Die zweite der obigen Beispielmatrizen hat Stufen an den Stellen 3,5,6. Die erweiterte Matrix muß im Falle der Lösbarkeit Nullzeilen 4,5 haben. Die

Unbekannten  $x_1, x_2, x_4, x_7, x_8$  sind freie Parameter. Die dritte Zeile etwa liefert

$$x_6 = -3x_7 - 7x_8 + b_3$$

als Festlegung von  $x_6$ .  $\diamond$

**(1.12) Beispiel.** Die Einheitsmatrix ist eine quadratische Zeilenstufenmatrix.  $\diamond$

#### 1.4. Elementare Zeilenumformungen

Unser Ziel ist ein Rechenverfahren zur Lösung linearer Gleichungen. Es beruht auf einem Verfahren, eine Matrix in Zeilenstufenform zu überführen. Das Verfahren setzt sich aus den folgenden Schritten zusammen.

**(1.13) Definition.** Die folgenden Veränderungen einer Matrix werden als *elementare Zeilenumformungen* bezeichnet.

- (1) Die Einträge einer Zeile werden alle mit einer festen Zahl  $\lambda \neq 0$  multipliziert.
- (2) Eine Zeile wird zu einer anderen addiert. Das soll heißen: wird die  $r$ -te Zeile zur  $s$ -ten addiert ( $r \neq s$ ), so hat die neue Matrix als  $s$ -te Zeile

$$a_{s1} + a_{r1}, \dots, a_{sn} + a_{rn}$$

während die anderen Zeilen ungeändert bleiben.

- (3) Vertauschen zweier Zeilen.

**(1.14) Beispiel.** Hier ist ein Beispiel einer Sequenz von solchen Umformungen.

$$\begin{pmatrix} 2 & -3 \\ 1 & -2 \end{pmatrix} \mapsto \begin{pmatrix} 2 & -3 \\ 2 & -4 \end{pmatrix} \mapsto \begin{pmatrix} 2 & -3 \\ 0 & -1 \end{pmatrix} \mapsto \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Nacheinander werden angewendet: (1), (2), (2), (1)(1).  $\diamond$

**(1.15) Satz.** Wird die erweiterte Matrix eines Gleichungssystems einer elementaren Zeilenumformung unterworfen, so hat das zur neuen Matrix gehörende System dieselbe Lösungsmenge wie das ursprüngliche.

BEWEIS. Die Aussage ist offensichtlich für (3). Zu (1): Löst das  $n$ -Tupel  $(x_j)$  die Gleichung (1.1), so auch die Gleichung, die daraus durch Multiplikation mit  $\lambda \neq 0$  entsteht. Also ist die Lösungsmenge des alten Systems in der des neuen enthalten. Durch Division mit  $\lambda$  ergibt sich die umgekehrte Inklusion. Zu (2): Löst ein  $n$ -Tupel zwei lineare Gleichungen, so auch deren Summe. Durch Subtraktion einer Gleichung von der Summengleichung können wir wieder zurückrechnen.  $\square$

Wir können die Prozesse (1) und (2) natürlich auch gleichzeitig durchführen: Addition des  $\lambda$ -fachen einer Zeile zu einer anderen. Den Prozeß (3) kann man übrigens aus den Prozessen (1) und (2) aufbauen, so daß man ihn für die allgemeinen Überlegungen weglassen könnte. Ein Beweis dafür wird durch die folgende Sequenz angedeutet, in der  $a_1$  und  $a_2$  Zeilen einer Matrix sind.

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \begin{pmatrix} a_1 + a_2 \\ a_2 \end{pmatrix}, \quad \begin{pmatrix} a_1 + a_2 \\ -a_1 \end{pmatrix}, \quad \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}, \quad \begin{pmatrix} a_2 \\ a_1 \end{pmatrix}$$



### 1.5. Der Gaußsche Algorithmus

Ein *Algorithmus* ist ein systematisches (schematisches) Rechenverfahren. Schematisch bedeutet, daß es selbst ein Computer kann. Der *Gaußsche Algorithmus* ist das im Beweis des nächsten Satzes geschilderte Verfahren.

**(1.16) Satz.** *Eine Matrix läßt sich durch elementare Zeilenumformungen in Zeilenstufenform bringen.*

BEWEIS. Wir betrachten nacheinander die Spalten der Matrix. Bestehen die Spalten 1 bis  $j - 1$  nur aus Nullen so sind sie schon erledigt. Wir erreichen durch Zeilenumformungen, daß an der Stelle  $(1, j)$  ein von Null verschiedenes Element steht. Durch (1) machen wir es zu 1. Indem wir nun geeignete Vielfache der ersten Zeile von den anderen subtrahieren, machen wir alle anderen Einträge der  $j$ -ten Spalte zu Null. Die Stelle  $(1, j)$  ist die erste „Stufe“ der gesuchten Matrix.

Nun sehen wir auf die Matrix, die durch Streichung der ersten Zeile entsteht, und behandeln sie genau so. Und so weiter.

Das Resultat hat noch nicht ganz Zeilenstufenform. Wir betrachten die Stufen von links nach rechts und machen durch Prozesse (2) die Einträge oberhalb jeder Stufe nacheinander zu Null.  $\square$

## 2 Vektorrechnung

### 2.1. Der Standardvektorraum

Mit den Elementen des Zahlenraumes  $\mathbb{R}^n$  kann man zwei Rechenoperationen durchführen. In diesem Kontext werden die Elemente des Zahlenraumes nicht mehr Punkte genannt sondern *Vektoren*.

**(2.1) Definition.** Die *Addition von Vektoren* wird durch die Formel

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$$

erklärt. Die *Multiplikation eines Vektors mit einem Skalar* wird durch die Formel

$$\lambda(x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n)$$

erklärt. In diesem Zusammenhang wird eine reelle Zahl  $\lambda$  *Skalar* genannt. Wir nennen  $x_k$  die  $k$ -te *Komponente* des Vektors  $(x_1, \dots, x_n)$ . Bei der Addition von Vektoren werden also jeweils dieselben Komponenten addiert. Bei der Skalarmultiplikation mit  $\lambda$  wird jede Komponente mit  $\lambda$  multipliziert. Wir setzen

$$-(x_1, \dots, x_n) = (-x_1, \dots, -x_n).$$

Der Vektor  $(0, \dots, 0)$  heißt *Nullvektor*.  $\diamond$

Wir bezeichnen Vektoren durch kleine Buchstaben

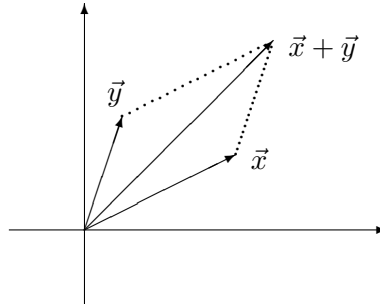
$$x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Der Nullvektor würde dann einfach auch wieder durch eine Null bezeichnet. Am Anfang und zur Unterscheidung verwenden wir zur Bezeichnung von Vektoren auch deutsche Buchstaben oder lateinische mit einem Pfeil darüber

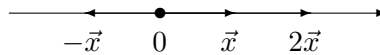
$$\mathbf{x} = \vec{x} = (x_1, \dots, x_n).$$

Die Bezeichnung mit dem Pfeil hat den Grund, daß wir bei Veranschaulichung in der Ebene oder im dreidimensionalen Raum den Vektor  $\vec{x}$  durch einen Pfeil vom Nullpunkt zum Punkt  $x$  zeichnen.

Geometrisch wird die Addition von Vektoren durch die *Parallelogrammregel* veranschaulicht:



Die Skalarmultiplikation wird durch eine *Dehnung* oder *Stauchung* eventuell mit Umkehrung der Richtung veranschaulicht: Ist  $\vec{x} \neq 0$ , so bilden die Punkte der Form  $\lambda\vec{x}$  eine Gerade durch den Nullpunkt und durch  $\vec{x}$ .



**(2.2) Definition.** Die Menge  $\mathbb{R}^n$  zusammen mit den beiden Rechenoperationen „Addition von Vektoren“ und „Skalarmultiplikation“ heißt *n-dimensionaler Standardvektorraum*.  $\diamond$

Sei  $e_i$  (oder  $\vec{e}_i$ ) der Vektor, der an der  $i$ -ten Stelle eine 1 hat und sonst nur Nullen. Er heißt  $i$ -ter (*Standard-*)*Einheitsvektor*. In der Zahlenebene zeigt  $e_1$  in die Richtung der positiven  $x$ -Achse und  $e_2$  in die Richtung der positiven  $y$ -Achse. Allgemein bezeichnet man die Menge der  $n$ -Tupel, die höchstens an der  $i$ -ten Stelle einen von Null verschiedenen Eintrag haben, als die  $i$ -te *Koordinatenachse* des  $\mathbb{R}^n$ .

**(2.3) Beispiel.** Die Rechenoperationen des Vektorraums liefern die Gleichung

$$x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 = (x_1, x_2, x_3).$$

Analog im  $\mathbb{R}^n$ .  $\diamond$

Indem man von dem Punkt  $P \in \mathbb{R}^n$  zu dem Punkt  $P + \vec{y}$  übergeht, sagt man auch, man habe den Vektor  $\vec{y} \in \mathbb{R}^n$  vom Punkt  $P$  aus *abgetragen*. Wir haben an dieser Stelle absichtlich einmal in der Notation zwischen Punkten und Vektoren unterschieden, obgleich beides Elemente des  $\mathbb{R}^n$  sind. Ist  $Q$  das Resultat dieses Abtragens, so schreibt man auch  $\vec{y} = \overrightarrow{PQ}$ .

## 2.2. Skalarprodukt. Länge. Winkel

Vektoren haben eine Länge. Je zwei bilden einen Winkel. Diese beiden Begriffe werden aus dem sogenannten Skalarprodukt hergeleitet.

**(2.4) Definition.** Das (*Standard-*)*Skalarprodukt* auf dem  $\mathbb{R}^n$  ist die Vorschrift, die je zwei Vektoren

$$\vec{x} = (x_1, \dots, x_n), \quad \vec{y} = (y_1, \dots, y_n)$$

die Zahl

$$\langle \vec{x}, \vec{y} \rangle = x_1 y_1 + \dots + x_n y_n$$

zuordnet, die *Skalarprodukt* von  $\vec{x}$  und  $\vec{y}$  genannt wird.  $\diamond$

Es gibt hier zwei sprachlich sehr ähnliche Begriffe, die etwas völlig anderes bedeuten, und deshalb sorgsam unterschieden werden müssen:

- (1) Das Produkt eines Vektors mit einem Skalar. Das Resultat ist wieder ein Vektor.
  - (2) Das Skalarprodukt zweier Vektoren. Das Resultat ist eine Zahl (ein Skalar).
- Aus der Definition des Skalarproduktes liest man unmittelbar die folgenden Rechenregeln ab:

**(2.5) Eigenschaften des Skalarprodukts.**

$$\begin{aligned} \langle \vec{a} + \vec{b}, \vec{c} \rangle &= \langle \vec{a}, \vec{c} \rangle + \langle \vec{b}, \vec{c} \rangle \\ \langle \lambda \vec{a}, \vec{b} \rangle &= \lambda \langle \vec{a}, \vec{b} \rangle \\ \langle \vec{a}, \vec{b} \rangle &= \langle \vec{b}, \vec{a} \rangle \end{aligned}$$

Die dritte Eigenschaft bezeichnet man als *Symmetrie* des Skalarprodukts, die ersten beiden als *Linearität* in der ersten Variablen. Wegen der Symmetrie (oder direkt aus der Definition ersichtlich) gelten analoge Linearitätseigenschaften natürlich auch für die zweite Variable.

Das Skalarprodukt ist anschaulich nicht unmittelbar einsichtig. Wir verarbeiten es aber sogleich zu nützlichen Begriffen.

Zunächst einmal ist  $\langle \vec{x}, \vec{x} \rangle$  für jeden Vektor eine nichtnegative Zahl. Wir können deshalb daraus die (nichtnegative) Quadratwurzel ziehen.

**(2.6) Definition.** Als *Norm* des Vektors  $\vec{x} = (x_1, \dots, x_n)$  bezeichnen wir die nichtnegative Zahl

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Die Norm bezeichnen wir auch als *Länge* des Vektors. Mit Hilfe der Norm wird der *Abstand zweier Vektoren*  $\vec{x}, \vec{y}$  (bzw., in anderer Sprechweise, zweier Punkte) als Norm der Differenz  $\|\vec{x} - \vec{y}\|$  definiert.  $\diamond$

In der Ebene oder im dreidimensionalen Raum steht diese Definition im Einklang mit dem Satz des Pythagoras aus der Elementargeometrie.

**(2.7) Eigenschaften der Norm.**

$$\begin{aligned} \|\vec{x} + \vec{y}\| &\leq \|\vec{x}\| + \|\vec{y}\| \\ \|\lambda \vec{x}\| &= |\lambda| \cdot \|\vec{x}\| \\ |\langle \vec{x}, \vec{y} \rangle| &\leq \|\vec{x}\| \|\vec{y}\| \end{aligned}$$

Die erste Ungleichung wird als *Dreiecksungleichung* bezeichnet; in der geometrischen Interpretation bedeutet sie nämlich, daß in einem Dreieck die Länge einer Seite

höchstens gleich der Längensumme der beiden anderen ist. Die Ungleichung in der dritten Zeile heißt *Ungleichung von Cauchy-Schwarz*. In der zweiten Zeile ist  $|\lambda|$  der absolute Betrag der reellen Zahl  $\lambda$ . Wegen  $|\lambda| = \sqrt{\lambda^2}$  folgt die zweite Zeile unmittelbar aus der Definition der Norm; die anderen beweisen wir alsbald.

Mit der Norm und dem Abstand haben wir *eine* Verwendung des Skalarproduktes kennengelernt. Eine zweite betrifft die Winkel.

**(2.8) Definition.** Zwei Vektoren  $\vec{x}, \vec{y}$  heißen *orthogonal* oder *senkrecht zueinander*, wenn ihr Skalarprodukt  $\langle \vec{x}, \vec{y} \rangle = 0$  ist. Wir verwenden dafür das Symbol  $\vec{x} \perp \vec{y}$ .  $\diamond$

**(2.9) Beispiel.** Für  $i \neq j$  sind offenbar die Standardeinheitsvektoren  $e_i$  und  $e_j$  orthogonal. Das entspricht der Anschauung.  $\diamond$

Wir können aus dem Skalarprodukt den allgemeinen *Satz des Pythagoras* in der folgenden Form herleiten:

**(2.10) Satz.** Seien  $\vec{u}$  und  $\vec{v}$  orthogonale Vektoren. Dann gilt

$$\|\vec{u}\|^2 + \|\vec{v}\|^2 = \|\vec{u} + \vec{v}\|^2.$$

BEWEIS. Wir rechnen mit den Regeln (2.5) des Skalarproduktes

$$\langle \vec{u} + \vec{v}, \vec{u} + \vec{v} \rangle = \langle \vec{u}, \vec{u} \rangle + \langle \vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{u} \rangle + \langle \vec{v}, \vec{v} \rangle.$$

Nach Voraussetzung sind die beiden mittleren Summanden Null. Nach Definition der Norm verbleibt die Behauptung.  $\square$

**(2.11) Notiz.** Seien  $\vec{x}, \vec{y} \in \mathbb{R}^n$  gegeben und sei  $\vec{x} \neq 0$ . Dann gibt es eine sogenannte orthogonale Zerlegung der Form

$$\vec{y} = \lambda \vec{x} + \vec{z}, \quad \lambda \in \mathbb{R}, \quad \langle \vec{x}, \vec{z} \rangle = 0.$$

Sie ist eindeutig bestimmt, und es ist  $\lambda = \|\vec{x}\|^{-2} \langle \vec{x}, \vec{y} \rangle$ .

BEWEIS. Aus  $\vec{z} = \vec{y} - \lambda \vec{x}$  und  $\langle \vec{x}, \vec{z} \rangle = 0$  folgt nämlich

$$0 = \langle \vec{x}, \vec{z} \rangle = \langle \vec{x}, \vec{y} - \lambda \vec{x} \rangle = \langle \vec{x}, \vec{y} \rangle - \lambda \langle \vec{x}, \vec{x} \rangle,$$

so daß  $\lambda$  die behauptete Gestalt hat. Umgekehrt rechnet man in ähnlicher Weise nach, daß mit diesem  $\lambda$  der Vektor  $\vec{y} - \lambda \vec{x}$  orthogonal zu  $\vec{x}$  ist.  $\square$

*Beweis der Cauchy-Schwarz'schen Ungleichung.* Wir verwenden Resultat und Bezeichnung der voranstehenden Notiz. Wir wollen  $|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \|\vec{y}\|$  zeigen. Für  $\vec{x} = 0$  ist die Behauptung offenbar richtig. Andernfalls folgt aus der letzten Notiz und dem Satz des Pythagoras

$$\|\vec{y}\|^2 = \|\lambda \vec{x}\|^2 + \|\vec{z}\|^2 \geq |\lambda|^2 \|\vec{x}\|^2 = \frac{\langle \vec{x}, \vec{y} \rangle^2}{\|\vec{x}\|^4} \|\vec{x}\|^2,$$

und das ist zu der behaupteten Ungleichung gleichwertig.  $\square$

**(2.12) Zusatz zur Cauchy-Schwarz'schen Ungleichung.** Aus dem Beweis des letzten Satzes erkennen wir, daß die Gleichheit  $|\langle \vec{x}, \vec{y} \rangle| = \|\vec{x}\| \|\vec{y}\|$  genau dann besteht,

wenn  $\vec{y}$  ein skalares Vielfaches von  $\vec{x}$  oder  $\vec{x}$  gleich Null ist. Das entspricht nämlich dem Fall  $z = 0$  im letzten Beweis.  $\square$

*Beweis der Dreiecksungleichung.* Wir verwenden wieder die Bezeichnungen des letzten Satzes und wollen  $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$  zeigen. Das wird durch die folgende Rechnung belegt:

$$\begin{aligned} \|\vec{x} + \vec{y}\|^2 &= \|\vec{x}\|^2 + \|\vec{y}\|^2 + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle \\ &\leq \|\vec{x}\|^2 + \|\vec{y}\|^2 + 2|\langle \vec{x}, \vec{y} \rangle| \\ &\leq \|\vec{x}\|^2 + \|\vec{y}\|^2 + 2\|\vec{x}\|\|\vec{y}\| \\ &= (\|\vec{x}\| + \|\vec{y}\|)^2. \end{aligned}$$

Darin haben wir die Cauchy-Schwarz'sche Ungleichung verwendet.  $\square$

Sei  $\vec{x} \neq \vec{0}$  und  $\vec{y} \neq \vec{0}$ . Wir können dann die Cauchy-Schwarz'sche Ungleichung in der Form

$$-1 \leq \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\|\|\vec{y}\|} \leq 1$$

schreiben. Wir verwenden nun die trigonometrische Funktion Cosinus (siehe dazu II.6). Wir brauchen folgende Aussage über ihren Verlauf: Zu jeder Zahl  $t$  mit der Eigenschaft  $-1 \leq t \leq 1$  gibt es genau eine Zahl  $\alpha$  mit  $0 \leq \alpha \leq \pi$ , so daß  $\cos \alpha = t$  ist. Es ist  $\cos 0 = 1$ ,  $\cos \pi/2 = 0$ ,  $\cos \pi = -1$ . Wenn  $\alpha$  von Null nach  $\pi$  läuft, so fällt  $\cos \alpha$  streng monoton von 1 auf  $-1$ .

Wir benutzen diese Aussagen in der Cauchy-Schwarz'schen Ungleichung und schreiben

$$(2.13) \quad \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\|\|\vec{y}\|} = \cos \alpha.$$

(2.14) **Definition.** Die durch (2.13) eindeutig bestimmte Zahl  $\alpha$  zwischen 0 und  $\pi$  bezeichnen wir als den von  $\vec{x}$  und  $\vec{y}$  aufgespannten *Winkel*.  $\diamond$

Diese Definition steht im Einklang mit der Definition von orthogonal, weil  $\cos \pi/2 = 0$  ist. Beispielsweise ist der Winkel zwischen  $\vec{x} = (1, 0)$  und  $\vec{y} = (1, 1)$  wegen  $\cos(\pi/4) = 1/\sqrt{2}$  gleich  $\pi/4$  (entspricht  $45^\circ$ ).

### 2.3. Hyperebenen

Die lineare Gleichung

$$a_1x_1 + \cdots + a_nx_n = \lambda$$

können wir mit den Vektoren

$$\vec{a} = (a_1, \dots, a_n) \quad (x_1, \dots, x_n)$$

in der Kurzform

$$(2.15) \quad \langle \vec{a}, \vec{x} \rangle = \lambda$$

schreiben. Die Lösungsmenge dieser Gleichung wird im Fall  $\vec{a} \neq \vec{0}$  als *Hyperebene* des  $\mathbb{R}^n$  bezeichnet und (2.15) heißt *Hyperebenengleichung*. Für  $n = 2$  ist das eine Gerade in der Ebene, für  $n = 3$  eine *Ebene* im dreidimensionalen Raum.

Bekanntlich können wir eine Hyperebenengleichung mit einer von Null verschiedenen reellen Zahl multiplizieren, ohne die Lösungsgesamtheit zu verändern. Das nutzen wir jetzt aus, um die Gleichung in eine Normalform zu überführen.

Einen Vektor der Länge 1 nennen wir allgemein *Einheitsvektor*. Ist  $\vec{a} \neq 0$  ein beliebiger Vektor, so ist  $\vec{a}/\|\vec{a}\|^{-1}$  ein Einheitsvektor. Indem wir also (2.15) gegebenenfalls mit  $\|\vec{a}\|^{-1}$  multiplizieren, können wir zur Beschreibung der Hyperebene  $\vec{a}$  als Einheitsvektor voraussetzen. Falls  $\lambda < 0$  ist, können wir die Gleichung noch mit  $-1$  multiplizieren. Damit haben wir das folgende Ergebnis gewonnen:

**(2.16) Definition.** Die *Hesse'sche Normalform* einer *Hyperebenengleichung* ist eine Gleichung der Form (2.15) mit einem Einheitsvektor  $\vec{a}$  und einer Zahl  $\lambda \geq 0$ . Jede Hyperebene läßt sich durch eine derartige Gleichung beschreiben.  $\diamond$

Die Lösungsmenge  $L$  einer Gleichung  $\langle \vec{a}, \vec{x} \rangle = \lambda$  in Hessescher Normalform ist die Menge

$$L = \{\vec{y} + \lambda\vec{a} \mid \vec{y} \perp \vec{a}, \vec{y} \in \mathbb{R}^n\},$$

denn aus  $\vec{y} \perp \vec{a}$  folgt  $\langle \vec{a}, \vec{y} + \lambda\vec{a} \rangle = \langle \vec{a}, \vec{y} \rangle + \langle \vec{a}, \lambda\vec{a} \rangle = \lambda$ , und ist  $\vec{x} \in L$ , so rechnet man ebenso  $(\vec{x} - \lambda\vec{a}) \perp \vec{a}$  nach. Nach dem Satz des Pythagoras ist  $\|\vec{y} + \lambda\vec{a}\| = \|\vec{y}\| + |\lambda|\|\vec{a}\|$ . Unter den Vektoren  $\vec{y} + \lambda\vec{a}$ ,  $\vec{y} \perp \vec{a}$  hat deshalb  $\lambda\vec{a}$  die kleinste Norm  $|\lambda|$ . Deshalb bezeichnet man  $\lambda$  als *orientierten Abstand der Hyperebene vom Nullpunkt* und  $\lambda\vec{a}$  als *Lotvektor* vom Nullpunkt auf  $L$ .

Die Punkte, die nicht auf der Hyperebene liegen, zerfallen in zwei Typen, je nachdem ob  $\langle \vec{a}, \vec{x} \rangle - \lambda$  positiv oder negativ ist. Wir nennen

$$H(+)=\{\vec{x} \mid \langle \vec{a}, \vec{x} \rangle - \lambda > 0\}$$

den durch  $L$  bestimmten *positiven Halbraum* und

$$H(-)=\{\vec{x} \mid \langle \vec{a}, \vec{x} \rangle - \lambda < 0\}$$

den *negativen Halbraum*. Wir nennen  $\langle \vec{a}, \vec{x} \rangle - \lambda$  den *orientierten Abstand* von  $\vec{x}$  von der Hyperebene  $L$ . Wir sagen „orientiert“, weil der Abstand hier mit einem Vorzeichen versehen ist; normalerweise ist ein Abstand ja nichtnegativ.

**(2.17) Notiz.** Ist  $\vec{x} = \mu\vec{a} + \vec{w}$ ,  $\vec{a} \perp \vec{w}$  eine orthogonale Zerlegung eines Vektors  $\vec{x} \in \mathbb{R}^n$ , so ist  $|\mu - \lambda|$  der minimale Abstand eines Punktes aus  $L$  von  $\vec{x}$ .

BEWEIS. Nach dem Satz des Pythagoras ist

$$\|\vec{x} - (\vec{y} + \lambda\vec{a})\|^2 = \|(\vec{w} - \vec{y}) + (\mu - \lambda)\vec{a}\|^2 = \|\vec{w} - \vec{y}\|^2 + |\mu - \lambda|^2.$$

Der minimale Wert ergibt sich, wenn der Summand  $\|\vec{w} - \vec{y}\|^2 = 0$  ist.  $\square$

Sind  $\langle \vec{a}_1, \vec{x} \rangle = \lambda_1$  und  $\langle \vec{a}_2, \vec{x} \rangle = \lambda_2$  Hyperebenengleichungen, so wird der Winkel zwischen  $\vec{a}_1$  und  $\vec{a}_2$  auch *Winkel zwischen den beiden Hyperebenen* genannt.

## 2.4. Geraden

Sei  $\vec{x} \in \mathbb{R}^n$  ein von Null verschiedener Vektor. Die Menge aller Vielfachen  $\{\lambda\vec{x} \mid \lambda \in \mathbb{R}\}$  nennen wir *Richtung* von  $\vec{x}$  und verwenden dafür das Symbol  $[\vec{x}]$ . Zwei von Null verschiedene Vektoren  $\vec{x}$  und  $\vec{y}$  haben also genau dann dieselbe Richtung, wenn es eine von Null verschiedene reelle Zahl  $\mu$  gibt, mit der  $\vec{x} = \mu\vec{y}$  gilt.

**(2.18) Definition.** Seien  $\vec{x}, \vec{b} \in \mathbb{R}^n$  und sei  $\vec{x} \neq \vec{0}$ . Die Menge aller Punkte der Form

$$(2.19) \quad \lambda \vec{x} + \vec{b}, \quad \lambda \in \mathbb{R}$$

heißt *Gerade* im  $\mathbb{R}^n$  der *Richtung*  $[\vec{x}]$  mit dem *Fußpunkt*  $\vec{b}$ . Diese Darstellung einer Geraden heißt *Parameterdarstellung*.  $\diamond$

Dieselbe Gerade hat verschiedene Parameterdarstellungen. Welche Freiheiten haben wir?

Hat  $\vec{y}$  dieselbe Richtung wie  $\vec{x}$ , so gibt es eine Relation  $\vec{x} = \mu \vec{y}$ ,  $\mu \neq 0$ . Es gilt

$$\lambda \vec{x} + \vec{b} = \lambda \mu \vec{y} + \vec{b}.$$

Mit  $\lambda$  durchläuft auch  $\lambda \mu$  alle reellen Zahlen. Wir können die Gerade also auch durch  $\vec{y}$  darstellen.

Ist  $\vec{c} = \lambda_0 \vec{x} + \vec{b}$  ein Punkt der Geraden, so gilt

$$\lambda \vec{x} + \vec{b} = (\lambda - \lambda_0) \vec{x} + \vec{c}.$$

Wir können auch  $\vec{c}$  als Fußpunkt wählen.

Trotz der verschiedenen Darstellungen einer Geraden ist die zugehörige Richtung eindeutig bestimmt. Es gilt nämlich:

**(2.20) Notiz.** Sei  $L$  eine Gerade. Die Menge aller Differenzen

$$\{\vec{u} - \vec{v} \mid \vec{u}, \vec{v} \in L\}$$

ist die Richtung von  $L$ .

BEWEIS. Das folgt unmittelbar aus der Beschreibung (2.19): Die Differenz zweier Vektoren dieser Form liegt in  $[\vec{x}]$  und alle Vektoren in  $[\vec{x}]$  treten auch wirklich als eine solche Differenz auf.  $\square$

In (2.19) steht eigentlich die Funktion

$$(2.21) \quad \mathbb{R} \rightarrow \mathbb{R}^n, \quad \lambda \mapsto \lambda \vec{x} + \vec{b},$$

und die Gerade ist die Menge der Funktionswerte. Da diese Menge durch eine freie Variable  $\lambda$  oder, wie man sagt, durch den *Parameter*  $\lambda$  beschrieben wird, heißt (2.19) eine *Parameterdarstellung* der Geraden.

Sind  $\vec{u}, \vec{v} \in \mathbb{R}^n$ , so liefert

$$(2.22) \quad \lambda(\vec{u} - \vec{v}) + \vec{v} = \lambda \vec{u} + (1 - \lambda) \vec{v}$$

eine Gerade durch  $\vec{u}$  (für  $\lambda = 1$ ) und  $\vec{v}$  (für  $\lambda = 0$ ). Durchläuft  $\lambda$  in (2.22) nur das Intervall  $[0, 1]$ , so ist die resultierende Punktmenge die *Strecke* von  $\vec{u}$  nach  $\vec{v}$ . Ihr Mittelpunkt ist durch  $\lambda = 1/2$  gegeben.

Wir nennen zwei Geraden *windschief*, wenn sie sich nicht schneiden und verschiedene Richtung haben.

**(2.23) Satz.** Seien  $G$  und  $H$  windschiefe Geraden im  $\mathbb{R}^n$ . Unter den Punktepaaren  $(\mathbf{a}, \mathbf{b})$  mit  $\mathbf{a} \in G$  und  $\mathbf{b} \in H$  gibt es genau eines  $(\vec{u}, \vec{v})$  mit minimalem Abstand. Der Vektor  $\vec{u} - \vec{v}$  ist orthogonal zur Richtung von  $G$  und  $H$ . Die Gerade durch  $\vec{u}$  und  $\vec{v}$  heißt Lotgerade von  $G, H$ .

BEWEIS. Sei  $G$  durch  $\lambda\vec{x} + \vec{b}$  und  $H$  durch  $\mu\vec{y} + \vec{c}$  beschrieben. Wir fragen zunächst nach Vektoren der Form

$$(\lambda\vec{x} + \vec{b}) - (\mu\vec{y} + \vec{c}),$$

die auf  $\vec{x}$  und  $\vec{y}$  senkrecht stehen. Das führt durch Einsetzen in das Skalarprodukt mit den Regeln (2.5) zu den Bedingungen

$$\begin{aligned}\lambda\langle\vec{x}, \vec{x}\rangle - \mu\langle\vec{y}, \vec{x}\rangle &= \langle\vec{c} - \vec{b}, \vec{x}\rangle \\ \lambda\langle\vec{x}, \vec{y}\rangle - \mu\langle\vec{y}, \vec{y}\rangle &= \langle\vec{c} - \vec{b}, \vec{y}\rangle.\end{aligned}$$

Das ist ein Gleichungssystem für die Unbekannten  $\lambda$  und  $\mu$ . Nach I(3.5) hat diese Gleichung eine eindeutig bestimmte Lösung, wenn

$$\langle\vec{x}, \vec{x}\rangle\langle\vec{y}, \vec{y}\rangle - \langle\vec{x}, \vec{y}\rangle^2 \neq 0$$

ist. Nach dem Zusatz (2.12) ist das der Fall, wenn  $\vec{x}$  und  $\vec{y}$  verschiedene Richtung haben. Sei mithin  $\lambda, \mu$  die eindeutig bestimmte Lösung und

$$\vec{u} = \lambda\vec{x} + \vec{b}, \quad \vec{v} = \mu\vec{y} + \vec{c}.$$

Ein beliebiger Punkt von  $G$  bzw.  $H$  hat dann die Form  $\alpha\vec{x} + \vec{u}$  bzw.  $\beta\vec{y} + \vec{c}$ . Das Abstandsquadrat dieser Punkte errechnet sich nach dem Satz des Pythagoras zu

$$\begin{aligned}\|(\alpha\vec{x} + \vec{u}) - (\beta\vec{y} + \vec{c})\|^2 &= \|(\alpha\vec{x} - \beta\vec{y}) + (\vec{u} - \vec{y})\|^2 \\ &= \|\alpha\vec{x} - \beta\vec{y}\|^2 + \|\vec{u} - \vec{v}\|^2,\end{aligned}$$

da  $\vec{u} - \vec{v}$  nach Konstruktion zu  $\vec{x}$  und  $\vec{y}$ , also auch zu  $\alpha\vec{x} - \beta\vec{y}$  orthogonal ist. Der minimale Abstand liegt vor, wenn  $\|\alpha\vec{x} - \beta\vec{y}\| = 0$ , also  $\alpha\vec{x} - \beta\vec{y} = \vec{0}$  ist. Da  $\vec{x}, \vec{y}$  aber verschiedene Richtung haben, ist die letzte Gleichung nur für  $\alpha = \beta = 0$  möglich.  $\square$

### 3 Die Methode der kleinsten Quadrate

Nicht immer werden  $n$ -Tupel von Zahlen als Vektoren interpretiert. Sie treten auch einfach als Datenmengen auf. Dafür liefert dieser Abschnitt ein Beispiel.

Die quadratische Funktion

$$px^2 + 2qx + r = p\left(x + \frac{q}{p}\right)^2 + r - \frac{q^2}{p}$$

hat für  $p > 0$  an der Stelle  $x = -\frac{q}{p}$  einen minimalen und für  $p < 0$  an dieser Stelle einen maximalen Funktionswert.

Seien  $(x_1, y_1), \dots, (x_n, y_n)$  eine Reihe von Meßwerten. Es wird eine Gerade  $y = ax + b$  gesucht, die am besten mit diesen Werten übereinstimmt. Der folgende *Ansatz* hat sich bewährt: Die Gerade soll so bestimmt werden, daß die Summe der Fehlerquadrate

$$(3.1) \quad S(a, b) = \sum_{i=1}^n (ax_i + b - y_i)^2$$

möglichst klein ist. Ist diese Summe Null, so liegen alle Punkte genau auf der Geraden. Wir rechnen das Quadrat aus. Da im folgenden nur  $\sum_{i=1}^n$  vorkommt, schreiben wir dafür lediglich  $\sum$ . Es ergibt sich



$$(3.2) \quad S(a, b) = a^2 \sum x_i^2 + nb^2 + 2ab \sum x_i - 2a \sum x_i y_i - 2b \sum y_i + \sum y_i^2.$$

Wir benutzen nun die Anfangsbemerkung über quadratische Funktionen. Wir nehmen an,  $a$  und  $b$  seien so bestimmt, daß  $S(a, b)$  der minimale Funktionswert der Funktion in zwei Veränderlichen  $(u, v) \mapsto S(u, v)$  ist. Wir werden sehen, daß sich  $a$  und  $b$  unter dieser Annahme ausrechnen lassen. Man kann zeigen, daß dann auch tatsächlich ein Minimum dieser Funktion vorliegt (siehe dazu den Abschnitt über lokale Minima bei Funktionen mehrerer Veränderlicher in II.4). Wir halten zunächst  $b$  fest. Dann liegt auch ein minimaler Funktionswert der quadratischen Funktion  $u \mapsto S(u, b)$  vor. Dieser ergibt sich an der Stelle

$$(3.3) \quad a = \frac{\sum x_i y_i - b \sum x_i}{\sum x_i^2}.$$

Nun halten wir  $a$  fest und bestimmen die Stelle, für die die quadratische Funktion  $v \mapsto S(a, v)$  einen minimalen Funktionswert hat. Es ergibt sich

$$(3.4) \quad b = \frac{\sum y_i - a \sum x_i}{n}.$$

Durch (3.3) und (3.4) haben wir zwei Gleichungen für die unbekanntenen Daten  $a$  und  $b$  erhalten. Wir setzen den Wert (3.4) für  $b$  in (3.3) ein und lösen nach  $a$  auf. Wir erhalten

$$(3.5) \quad a = \frac{\sum x_i y_i - \frac{1}{n} \sum x_i \cdot \sum y_i}{\sum x_i^2 - \frac{1}{n} (\sum x_i)^2}.$$

Diese Rechnung ist natürlich nur erlaubt, wenn der Nenner von Null verschieden ist. Um das zu untersuchen, rechnen wir den Nenner um.

Es ist nützlich, in diesem Kontext den *Mittelwert* (auch *arithmetisches Mittel* genannt) der  $x_i$  und  $y_i$  zu verwenden. Dieser ist durch

$$(3.6) \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

definiert. Der *mittlere quadratische Fehler* für die Abweichung einer Zahl  $z$  von den  $x_i$  ist

$$\frac{1}{n} \sum_{i=1}^n (x_i - z)^2.$$

Dieser ist minimal für  $z = \bar{x}$ . Mit dem durch (3.5) bestimmten Wert von  $a$  erhalten wir  $b$  aus der Gleichung (3.4)

$$(3.7) \quad b = \bar{y} - a\bar{x}.$$

Diese Gleichung besagt, daß die zu bestimmende Gerade jedenfalls durch die Mittelwerte  $(\bar{x}, \bar{y})$  läuft.

Eine kleine Rechnung, die als Aufgabe gelten mag, zeigt:

$$(3.8) \quad \frac{1}{n} \sum (x_i - \bar{x})^2 = \frac{1}{n} \sum x_i^2 - \bar{x}^2.$$

Dieser Wert, der im wesentlichen der Nenner von (3.5) ist, ist genau dann Null, wenn alle  $x_i = \bar{x}$  sind. Dann liegen aber alle Punkte auf einer Parallelen zur  $y$ -Achse. Diesen Fall können wir ausschließen.

Wir führen einige Bezeichnungen für die gewonnenen Größen ein, die sich auf das Datensystem  $(x_1, y_1), \dots, (x_n, y_n)$  beziehen.

$$(3.9) \quad s_x^2 = v_x = \frac{1}{n} \sum_{i=1}^n (\bar{x} - x_i)^2$$

heißt *mittlere quadratische Abweichung* vom Mittelwert  $\bar{x}$  oder *Varianz* und  $s_x = \sqrt{v_x}$  die *Standardabweichung*. Dieser Wert ist ein Maß dafür, wie stark die  $x_i$  verstreut sind. Wir nennen  $s_{xy} = \frac{1}{n} \sum x_i y_i - \bar{x} \bar{y}$  die *Kovarianz* und (siehe (3.5))

$$(3.10) \quad a = \frac{s_{xy}}{s_x^2}$$

*Regressionskoeffizient*. Die erhaltene Gerade  $y = ax + b$  heiße *Regressionsgerade*.

Wenn man zwei Datenreihen  $x_1, \dots, x_n$  und  $y_1, \dots, y_n$  mißt und nach einem eventuellen Zusammenhang sucht, hat man die Wahl, ob man die  $x_i$  oder die  $y_i$  als unabhängige Variable betrachtet. Durch Vertauschen der Variablen wird man auf das Problem geführt, eine Regressionsgerade  $x = a'y + b'$  zu suchen. Das geschieht wie oben. Es ergibt sich

$$(3.11) \quad a' = \frac{\frac{1}{n} \sum x_i y_i - \bar{x} \bar{y}}{\frac{1}{n} \sum y_i^2 - \bar{y}^2}, \quad b' = \bar{x} - a' \bar{y}.$$

Das ist im allgemeinen nicht dieselbe Gerade wie  $y = ax + b$ . Dieselbe Gerade liegt genau dann vor, wenn  $aa' = 1$ . Als Maß für die Abweichung führt man den *Korrelationskoeffizienten*

$$(3.12) \quad r(x, y) = \frac{s_{xy}}{s_x s_y}$$

ein. Die Geraden, die beide durch den Punkt  $(\bar{x}, \bar{y})$  laufen, stimmen für  $r(x, y) = \pm 1$  wegen  $r(X, Y)^2 = aa'$  überein. Es gilt immer

$$(3.13) \quad -1 \leq \frac{s_{xy}}{s_x s_y} \leq 1$$

(sofern wir  $s_x = 0$  oder  $s_y = 0$  wie oben ausschließen).

*Beweis von (3.13)*. Die Ungleichung ist äquivalent zu

$$\left( \sum x_i y_i - n \bar{x} \bar{y} \right)^2 \leq \left( \sum (x_i - \bar{x})^2 \right) \left( \sum (y_i - \bar{y})^2 \right).$$

Es ist

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \sum x_i \bar{y} - \sum y_i \bar{x} + n \bar{x} \bar{y} = \sum x_i y_i - n \bar{x} \bar{y}.$$

Sie ist demnach eine Anwendung der Cauchy-Schwarzschen Ungleichung auf die Vektoren  $(x_1 - \bar{x}, \dots, x_n - \bar{x})$  und  $(y_1 - \bar{y}, \dots, y_n - \bar{y})$ .  $\square$

## 4 Koordinatensysteme

### 4.1. Linearkombinationen

Die nächste Definition und die daraus abgeleiteten weiteren sind für die gesamte lineare Algebra von grundsätzlicher Bedeutung.

**(4.1) Definition.** Seien  $\vec{x}_1, \dots, \vec{x}_r$  Vektoren des  $\mathbb{R}^n$ . Ein Vektor der Form

$$\vec{x} = \lambda_1 \vec{x}_1 + \dots + \lambda_r \vec{x}_r$$

mit reellen Zahlen  $\lambda_j$  heißt *Linearkombination* der Familie  $(\vec{x}_1, \dots, \vec{x}_r)$ . Die Linearkombination heißt *echt*, wenn nicht alle  $\lambda_j$  Null sind. Eine Darstellung des Nullvektors als Linearkombination nennen wir eine *lineare Relation* zwischen den Mitgliedern der Familie.  $\diamond$

**(4.2) Beispiel.** Mit den Standardeinheitsvektoren  $\vec{e}_j \in \mathbb{R}^n$  gilt immer

$$(\lambda_1, \dots, \lambda_n) = \sum_{i=1}^n \lambda_i \vec{e}_i.$$

Jeder Vektor des  $\mathbb{R}^n$  ist eine Linearkombination der Familie der  $\vec{e}_i$ .  $\diamond$

**(4.3) Beispiel.** Für die Vektoren

$$\vec{x}_1 = (1, 1, 1), \quad \vec{x}_2 = (1, 2, 0), \quad \vec{x}_3 = (0, -1, 1)$$

gilt

$$-\vec{x}_1 + \vec{x}_2 + \vec{x}_3 = \vec{0}.$$

Der Nullvektor ist eine echte Linearkombination.  $\diamond$

**(4.4) Definition.** Die Familie  $(\vec{x}_1, \dots, \vec{x}_r)$  heißt *linear unabhängig*, wenn der Nullvektor keine echte Linearkombination dieser Familie ist; andernfalls heißt die Familie *linear abhängig*. Im Falle der linearen Unabhängigkeit folgt also aus einer Relation  $\sum_{j=1}^r \lambda_j \vec{x}_j = \vec{0}$ , daß alle  $\lambda_j = 0$  sind.  $\diamond$

Besteht die Familie aus einem einzigen Vektor  $\vec{x}$ , so ist sie genau dann linear unabhängig, wenn  $\vec{x} \neq \vec{0}$  ist. Enthält eine Familie den Nullvektor, so ist sie linear abhängig. Seien  $\vec{x}_1, \vec{x}_2$  zwei von Null verschiedene Vektoren. Ist  $\lambda_1 \vec{x}_1 + \lambda_2 \vec{x}_2 = \vec{0}$  und ist die Linearkombination eine echte, so läßt sich ein Vektor durch den anderen ausrechnen

$$\vec{x}_1 = -\frac{\lambda_2}{\lambda_1} \vec{x}_2.$$

Diese beiden Vektoren sind also genau dann linear abhängig, wenn sie auf derselben Geraden durch den Nullpunkt liegen. Sind drei von Null verschiedene Vektoren  $\vec{x}_j$  gegeben, so läßt sich ebenso aus einer echten Linearkombination  $\sum \lambda_j \vec{x}_j = \vec{0}$  einer aus den restlichen ausrechnen; sie sind also genau dann linear abhängig, wenn sie alle in einer Ebene oder auf einer Geraden durch den Nullpunkt liegen.

**(4.5) Beispiel.** Sei  $\vec{x}_1 = (1, 2)$  und  $\vec{x}_2 = (3, 1)$ . Sei der Vektor  $\vec{a} = (a_1, a_2)$  gegeben. Die Gleichung  $\lambda \vec{x}_1 + \mu \vec{x}_2 = \vec{a}$  ist äquivalent zu den beiden Gleichungen für die erste und zweite Komponente

$$\lambda + 3\mu = a_1, \quad 2\lambda + \mu = a_2.$$

Dieses Gleichungssystem für die Unbekannten  $\lambda, \mu$  hat immer eine eindeutig bestimmte Lösung. Also ist jeder Vektor des  $\mathbb{R}^2$  eine eindeutig bestimmte Linearkombination von  $\vec{x}_1, \vec{x}_2$ . Insbesondere entsteht der Nullvektor nur für  $\lambda = \mu = 0$ , d. h. die Familie ist linear unabhängig. Wir sagen,  $(\vec{x}_1, \vec{x}_2)$  ist ein Koordinatensystem für  $\mathbb{R}^2$ .  $\diamond$

**(4.6) Definition.** Mit  $L = L(\vec{x}_1, \dots, \vec{x}_r)$  bezeichnen wir die Menge aller Linearkombinationen von  $\vec{x}_1, \dots, \vec{x}_r$ . Wir sagen, die Menge  $L$  wird von den genannten Vektoren *aufgespannt* oder *erzeugt*, und  $(\vec{x}_1, \dots, \vec{x}_r)$  heißt dann ein *Erzeugendensystem* von  $L$ .  $\diamond$

Die nächste Definition faßt die formalen Eigenschaften solcher Mengen ins Auge.

**(4.7) Definition.** Eine Teilmenge  $L \subset \mathbb{R}^n$  heißt ein (linearer) *Unterraum* des Vektorraumes  $\mathbb{R}^n$ , wenn sie folgende Eigenschaften hat:

- (1) Sind  $x, y \in L$ , so ist auch  $x + y \in L$ .
- (2) Ist  $x \in L$  und  $\lambda \in \mathbb{R}$ , so ist  $\lambda x \in L$ .

Die in der vorigen Definition genannten Mengen sind lineare Unterräume.  $\diamond$

**(4.8) Beispiel.** Die Unterräume des  $\mathbb{R}^2$  sind: Der Nullraum  $\{\vec{0}\}$ , der gesamte Raum  $\mathbb{R}^2$ , sowie die Geraden durch den Nullpunkt. Die Unterräume des  $\mathbb{R}^3$  sind:  $\{\vec{0}\}$ ,  $\mathbb{R}^3$ , die Geraden durch den Nullpunkt, die Ebenen durch den Nullpunkt. Die genaue Begründung dafür ergibt sich aus dem alsbald genannten Basissatz.  $\diamond$

## 4.2. Basis. Koordinatensystem

**(4.9) Definition.** Eine Familie  $B = (\vec{x}_1, \dots, \vec{x}_r)$  aus Vektoren des Unterraums  $U$  heißt *Basis* oder *Koordinatensystem* von  $U$ , wenn jeder Vektor von  $U$  eine Linearkombination von  $B$  ist und wenn  $B$  linear unabhängig ist. Besteht die Basis aus paarweise orthogonalen Einheitsvektoren, so heißt sie eine *Orthonormalbasis*.

Ist  $B$  eine Basis von  $U$  und  $\vec{x} \in U$ , so gibt es also reelle Zahlen  $\lambda_1, \dots, \lambda_r$ , mit denen eine Darstellung  $\vec{x} = \sum_{j=1}^r \lambda_j \vec{x}_j$  gilt. Die Zahlen  $\lambda_j$  sind durch  $\vec{x}$  eindeutig bestimmt, denn wäre auch  $\vec{x} = \sum_{j=1}^r \mu_j \vec{x}_j$ , so würde folgen

$$\vec{0} = \vec{x} - \vec{x} = \sum \lambda_j \vec{x}_j - \sum \mu_j \vec{x}_j = \sum (\lambda_j - \mu_j) \vec{x}_j,$$

und weil die  $\vec{x}_j$  linear unabhängig sind, folgt  $\lambda_j - \mu_j = 0$  für alle  $j$ . Wir nennen  $\lambda_1, \dots, \lambda_r$  die *Koordinaten* von  $\vec{x}$  bezüglich des Koordinatensystems  $B$ .  $\diamond$

Die Familie der  $\vec{e}_j \in \mathbb{R}^n$  ist eine Orthonormalbasis des  $\mathbb{R}^n$ , wie man mittels (4.2) und (4.4) sofort erkennt. Die Koordinaten von  $(\lambda_1, \dots, \lambda_n)$  bezüglich dieser Basis sind die  $\lambda_j$ . Für das theoretische Verständnis der Basen ist der folgende Satz grundlegend.

**(4.10) Basissatz.** *Jeder Unterraum  $U$  des  $\mathbb{R}^n$  hat eine Basis. Die Anzahl der Elemente einer Basis ist durch  $U$  eindeutig bestimmt und heißt die Dimension von  $U$ . Ist  $S \subset U$  eine linear unabhängige Teilmenge und  $T \subset U$  ein Erzeugendensystem von  $U$ , so kann  $S$  durch Vektoren aus  $T$  zu einer Basis von  $U$  ergänzt werden.*

Ein eindimensionaler Unterraum heißt *Gerade*, ein zweidimensionaler *Ebene*. Der  $\mathbb{R}^n$  hat die Dimension  $n$ .

**(4.11) Beispiel.** Die Lösungsmenge eines *homogenen* linearen Gleichungssystems in  $n$  Unbekannten ist ein Unterraum des  $\mathbb{R}^n$ , wie man aus den Definitionen sogleich verifiziert. Ist  $B = (\vec{x}_1, \dots, \vec{x}_r)$  eine Basis dieses Lösungsraumes, so ist die Lösungsmenge also die Gesamtheit der Linearkombinationen  $\sum_{j=1}^r \lambda_j \vec{x}_j$ . Wie berechnet man eine Basis des Lösungsraumes? Siehe dazu (5.5).  $\diamond$

Hier noch eine sprachliche Vereinbarung. Der nur aus dem Nullvektor bestehende Nullraum ist ein Unterraum. Wir sagen, die leere Menge sei eine Basis des Nullraumes. Seine Dimension ist Null.

**(4.12) Basen des  $\mathbb{R}^n$ .** Seien  $\vec{a}_j = (a_{1j}, \dots, a_{jn})$  für  $j = 1, \dots, n$  Vektoren des  $\mathbb{R}^n$ . Eine Relation  $x_1 \vec{a}_1 + \dots + x_n \vec{a}_n = \vec{0}$  ist äquivalent zu den  $n$  Gleichungen des Systems

$$\sum_{i=1}^n a_{ij} x_j = 0, \quad j = 1, \dots, n.$$

In der Spalte  $j$  der Koeffizientenmatrix stehen die Komponenten von  $\vec{a}_j$ . Die lineare Unabhängigkeit besagt: Das Gleichungssystem hat nur die triviale Lösung  $(0, \dots, 0)$ . Das ist nach den Überlegungen des ersten Abschnittes genau dann der Fall, wenn sich die Koeffizientenmatrix durch elementare Zeilenumformungen in die Einheitsmatrix verwandeln läßt. In diesem Fall ist jedes System mit dieser Koeffizientenmatrix eindeutig lösbar. Die Vektoren bilden dann also sogar ein Koordinatensystem, wenn sie linear unabhängig sind.  $\diamond$

Sei  $A$  eine  $m \times n$ -Matrix. Dazu bilden wir den von den  $m$  Zeilenvektoren aufgespannten *Zeilenraum*  $ZR(A) \subset \mathbb{R}^n$  und den von den  $n$  Spaltenvektoren aufgespannten *Spaltenraum*  $SR(A) \subset \mathbb{R}^m$ . Beides sind lineare Unterräume. Die Dimension von  $ZR(A)$  heißt *Zeilenrang* von  $A$ , die Dimension von  $SR(A)$  *Spaltenrang* von  $A$ .

### 4.3. Basiswechsel

Koordinatensysteme sind nicht eindeutig bestimmt. In der Natur gibt es keine Koordinatensysteme. Die Wahl eines Koordinatensystems hat immer etwas Willkürliches. Seien

$$A = (\vec{a}_1, \dots, \vec{a}_n), \quad B = (\vec{b}_1, \dots, \vec{b}_n)$$

Basen des  $\mathbb{R}^n$ . Es gibt also reelle Zahlen  $\lambda_{ij}$ , so daß

$$\vec{a}_j = \sum_{i=1}^n \lambda_{ij} \vec{b}_i, \quad j = 1, \dots, n$$

ist, denn jeder Vektor, also auch  $\vec{a}_j$ , ist eine Linearkombination von  $B$ . Dadurch erhalten wir eine  $n \times n$ -Matrix

$$M_A^B := (\lambda_{kl}).$$

Sie heie *Basiswechselmatrix* für den Wechsel von  $A$  (alte Basis) nach  $B$  (neue Basis).

**(4.13) Merkregel.** In der  $k$ -ten Spalte der Basiswechselmatrix  $M_A^B$  stehen die Koordinaten des alten Vektors  $a_k$  bezüglich der neuen Basis  $B$ . (Man muß bei dieser Vereinbarung auf die Reihenfolge von  $A$  und  $B$  achten; man könnte es auch umgekehrt festsetzen.)  $\diamond$

**(4.14) Beispiel.** Wir betrachten die Basen  $A = (\vec{e}_1, \vec{e}_2)$  und  $B = (\vec{x}_1, \vec{x}_2)$  aus Beispiel (4.5). Dann gilt

$$\begin{aligned} \vec{x}_1 &= 1 \cdot \vec{e}_1 + 2 \cdot \vec{e}_2, & \vec{x}_2 &= 3 \cdot \vec{e}_1 + 1 \cdot \vec{e}_2 \\ \vec{e}_1 &= -\frac{1}{5}(1, 2) + \frac{2}{5}(3, 1), & \vec{e}_2 &= \frac{3}{5}(1, 2) - \frac{1}{5}(3, 1). \end{aligned}$$

Daraus lesen wir die Basiswechselmatrizen ab:

$$M_B^A = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}, \quad M_A^B = \begin{pmatrix} -\frac{1}{5} & \frac{3}{5} \\ \frac{3}{5} & -\frac{1}{5} \end{pmatrix}.$$

◇

## 5 Matrizenrechnung

### 5.1. Das Matrizenprodukt

Mit Matrizen gleicher Form kann man ähnlich wie mit Vektoren rechnen: Addition und Multiplikation mit einem Skalar. Außerdem gibt es als wesentliche neue Rechenart die Multiplikation von Matrizen.

Sei  $M(m \times n)$  die Menge aller  $m \times n$ -Matrizen. Sind  $A = (a_{ij})$  und  $B = (b_{ij})$  zwei solche und ist  $\lambda$  eine reelle Zahl, so definieren wir neue Matrizen  $A + B$  und  $\lambda A$  aus  $M(m \times n)$  durch die Vorschrift

$$(a_{ij}) + (b_{ij}) := (a_{ij} + b_{ij}), \quad \lambda(a_{ij}) := (\lambda a_{ij}).$$

Hier sind zwei Beispiele.

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix}$$

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

Die  $m \times 1$ -Matrizen nennen wir auch *Spaltenvektoren* und die  $1 \times n$ -Matrizen *Zeilenvektoren*.

**(5.1) Definition.** Das *Matrizenprodukt* ordnet einer  $m \times n$ -Matrix  $A = (a_{ij})$  und einer  $n \times p$ -Matrix  $B = (b_{kl})$  eine  $m \times p$ -Matrix zu, die mit  $AB$  oder  $A \cdot B$  bezeichnet wird, *Produkt* von  $A$  und  $B$  genannt wird, und deren Eintrag an der Stelle  $(k, j)$  die Summe

$$\sum_{l=1}^n a_{kl} b_{lj}$$

ist. Mit anderen Worten: An der Stelle  $(k, j)$  steht das Skalarprodukt aus dem  $k$ -ten Zeilenvektor des ersten Faktors  $A$  mit dem  $j$ -ten Spaltenvektor des zweiten Faktors  $B$ . ◇

Wir notieren einige Beispiele von Matrizenprodukten.

$$\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} -\frac{1}{5} & \frac{3}{5} \\ \frac{3}{5} & -\frac{1}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = (ac + bd)$$

$$\begin{pmatrix} c \\ d \end{pmatrix} \begin{pmatrix} a & b \end{pmatrix} = \begin{pmatrix} ca & cb \\ da & db \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = ?? \quad (\text{nicht definiert})$$

Wir kommentieren nun die Beispiele.

Im ersten treten die Matrizen aus Beispiel (5.14) auf. Das Resultat ist kein Zufall. Für Basiswechsel gilt immer, daß  $M_A^B \cdot M_B^A$  die Einheitsmatrix ist, wie wir gleich sehen werden.

Die zweite und dritte Gleichung zeigen, daß es bei der Matrizenmultiplikation auf die Reihenfolge der Faktoren ankommt. Im allgemeinen ist  $BA \neq AB$ . Für die Matrizenmultiplikation gilt nicht das Kommutativgesetz.

Das Resultat der vierten Multiplikation ist eine  $1 \times 1$ -Matrix. Eine solche ist dasselbe wie eine Zahl.

Die sechste Gleichung ist in anderer Bezeichnungweise noch einmal die Vorschrift für die Multiplikation zweier  $2 \times 2$ -Matrizen.

Das letzte Beispiel ist Unsinn. Die Matrizen passen der „Größe“ nach nicht zusammen; sie können nicht miteinander multipliziert werden. Die folgenden Regeln bestätigt man durch geduldiges Nachrechnen aus den Definitionen.

### (5.2) Rechenregeln für das Matrizenprodukt.

- (1)  $A(BC) = (AB)C$  Assoziativgesetz
- (2)  $A(B + C) = AB + AC$
- (3)  $(A + B)C = AC + BC$
- (4)  $(\lambda A)B = \lambda(AB) = A(\lambda B)$
- (5)  $(AB)^t = B^t A^t$

Hierbei wird unterstellt, daß man die Produkte jeweils bilden kann. □

**(5.3) Beispiel.** Bezeichne  $E_n$  die  $n \times n$ -Einheitsmatrix. Sei  $A$  eine  $p \times q$ -Matrix. Dann gilt

$$E_p A = A E_q = A,$$

wie man aus der Definition des Matrizenprodukts fast unmittelbar entnimmt. ◇

Wozu dient das Matrizenprodukt? Wir beschreiben einige typische Anwendungen: Basiswechsel, Koordinatenwechsel, inverse Matrizen, Gleichungssysteme, Zeilenumformungen.

### 5.2. Basiswechsel

Seien  $A, B, C$  drei Basen des  $\mathbb{R}^n$ . Für die  $n \times n$ -Matrizen der Basiswechsel gilt immer

$$M_B^C M_A^B = M_A^C.$$

Zum Nachweis setzen wir gemäß (4.13) an

$$\vec{b}_k = \sum_l \mu_{lk} \vec{c}_l, \quad \vec{a}_j = \sum_i \lambda_{ij} \vec{b}_i.$$

Wir setzen ein und erhalten

$$\vec{a}_k = \sum_i \left( \sum_l \mu_{il} \lambda_{lk} \right) \vec{c}_i.$$

Daraus entnimmt man die Behauptung. Da  $M_A^A$  immer die Einheitsmatrix  $E_n$  ist, so gilt insbesondere

$$M_A^B M_B^A = E_n.$$

### 5.3. Koordinatenwechsel

Die Basiswechselmatrizen sagen auch, wie sich die Koordinaten eines Vektors bezüglich zweier verschiedener Basen ineinander umrechnen lassen. Sei nämlich

$$\vec{x} = \sum_j \alpha_j \vec{a}_j = \sum_j \beta_j \vec{b}_j,$$

so ergibt sich durch Einsetzen des Basiswechsels  $\vec{a}_j = \sum_i \lambda_{ij} \vec{b}_i$  und Vergleich der Koordinaten die Regel

$$\beta_i = \sum_j \lambda_{ij} \alpha_j.$$

Bilden wir aus den  $\alpha_j$  und den  $\beta_j$  jeweils einen Spaltenvektor  $\vec{\alpha}$  und  $\vec{\beta}$ , so fassen wir die letzte Gleichung in der Matrixgleichung

$$\vec{\beta} = M_A^B \vec{\alpha}$$

zusammen. Man muß hier etwas aufpassen: Die Matrix  $M_A^B$  hat bezüglich Basen und Koordinaten eine andere Funktion. Wir stellen sie einmal gegenüber

$$\vec{a}_j = \sum_i \lambda_{ij} \vec{b}_i, \quad \beta_i = \sum_j \lambda_{ij} \alpha_j.$$

Will man es andersherum, so braucht man die inversen Matrizen, die wir jetzt erklären.



### 5.4. Inverse Matrizen

Wir nennen die  $n \times n$ -Matrix  $B$  *invers* zur  $n \times n$ -Matrix  $A$ , wenn

$$AB = BA = E_n$$

ist. Wir schreiben dann auch  $B = A^{-1}$ . Eine  $n \times n$ -Matrix  $A$  heißt *invertierbar*, wenn sie ein Inverses hat. Ein Inverses von  $A$  ist eindeutig bestimmt, denn aus  $AB = BA = E_n$  und  $AC = CA = E_n$  folgt

$$C = CE_n = C(AB) = (CA)B = E_n B = B.$$

Sind  $A$  und  $B$  invertierbare  $n \times n$ -Matrizen, so gilt

$$(AB)(B^{-1}A^{-1}) = ((AB)B^{-1})A^{-1} = (A(BB^{-1}))A^{-1} = (AE_n)A^{-1} = E_n.$$

Also ist  $(AB)^{-1} = B^{-1}A^{-1}$ . Die Faktoren muß man beim Invertieren vertauschen<sup>8</sup>. Insbesondere ist das Produkt invertierbarer Matrizen wieder invertierbar.

Basiswechselformen sind immer invertierbar, wie in (5.3) festgestellt wurde.

**(5.4) Beispiel.** Sei  $\text{Dia}(\lambda_1, \dots, \lambda_n)$  die Diagonalmatrix mit der Diagonale  $(\lambda_1, \dots, \lambda_n)$ . Sind alle  $\lambda_j \neq 0$ , so ist das Inverse  $\text{Dia}(\lambda_1^{-1}, \dots, \lambda_n^{-1})$ .  $\diamond$

**(5.5) Beispiel.** Die Matrix  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  ist genau dann invertierbar, wenn  $D := ad - bc \neq 0$  ist. Die inverse Matrix lautet

$$\frac{1}{D} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

### 5.5. Zeilenumformungen

Die früher betrachteten elementaren Zeilenumformungen einer Matrix  $A$  lassen sich dadurch bewirken, daß die Matrix von links mit elementaren invertierbaren Matrizen (sogenannten *Elementarmatrizen*) multipliziert wird. Soll die  $k$ -te Zeile mit  $\lambda \neq 0$  multipliziert werden, so verwenden wir in diesem Sinne eine Diagonalmatrix, die auf der Diagonale nur Einsen hat, bis auf die Stelle  $k$ , wo ein  $\lambda$  steht. Soll das  $\lambda$ -fache der  $i$ -te Zeile zur  $k$ -ten addiert werden, so verwenden wir eine Matrix, die auf der Diagonale Einsen hat, sonst aber nur Nullen, bis auf die Stelle  $(k, i)$ , wo ein  $\lambda$  steht. Die dadurch definierten elementaren Matrizen sind invertierbar, und die Inversen sind wieder elementare Matrizen.

**(5.6) Satz.** *Eine quadratische Matrix  $A$  ist genau dann invertierbar, wenn sie Produkt von Elementarmatrizen ist.*

BEWEIS. Sie läßt sich durch elementare Umformungen in die Einheitsmatrix verwandeln. Also gibt es elementare Matrizen  $S_1, \dots, S_k$ , so daß  $S_k \dots S_1 \cdot A$  die Einheitsmatrix ist. Dann ist aber  $A = S_1^{-1} \dots S_k^{-1}$ .  $\square$

Wir haben damit auch eine Methode, um daß Inverse einer Matrix zu berechnen.

<sup>8</sup>Wenn man erst das Hemd anzieht und dann die Jacke, so muß man sie in umgekehrter Reihenfolge wieder ausziehen.

**(5.7) Verfahren zur Berechnung der inversen Matrix.** Man schreibe nebeneinander  $A \mid E_m$ , bringe  $A$  durch Zeilenumformungen auf Zeilenstufenform und wende dieselben Zeilenumformungen auf  $E_m$  an; das letztere liefert  $A^{-1}$ .  $\diamond$

Analog zu den elementaren Zeilenumformungen definiert man elementare Spaltenumformungen (Vertauschen der Rolle von Zeilen und Spalten). Diese werden entsprechend durch Multiplikation von rechts mit elementaren Matrizen bewirkt.

Ist  $A$  eine  $m \times n$ -Matrix und  $B$  eine  $m \times m$ -Matrix, so sind die Zeilen von  $BA$  Linearkombinationen der Zeilen von  $A$ . Deshalb gilt für die Zeilenräume  $ZR(BA) \subset ZR(A)$ . Ist  $B$  invertierbar, so gilt  $ZR(A) = ZR(B^{-1}BA) \subset ZR(BA)$ . Für eine invertierbare Matrix gilt deshalb  $ZR(BA) = ZR(A)$ . Ebenso folgt für invertierbare Matrizen  $C$  die Gleichheit  $SR(CA) = SR(A)$ .

### 5.6. Gleichungssysteme

Sei  $A = (a_{ij})$  eine  $m \times n$ -Matrix. Mit den Spaltenvektoren, alias  $n \times 1$ -Matrizen,

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

können wir das Gleichungssystem aus den linearen Gleichungen

$$a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j$$

in der Kurzform

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

schreiben. Ist  $m = n$  und  $A$  eine invertierbare  $n \times n$ -Matrix, so folgt aus  $A\vec{x} = \vec{b}$  die Gleichung

$$A^{-1}(A\vec{x}) = (A^{-1}A)\vec{x} = E_n\vec{x} = \vec{x} = A^{-1}\vec{b},$$

das heißt  $A^{-1}\vec{b}$  ist die Lösung. Kenntnis der inversen Matrix liefert also sofort für jedes  $\vec{b}$  eine Lösung.

**(5.8) Satz.** *Ein Gleichungssystem mit quadratischer Koeffizientenmatrix hat genau dann eine eindeutig bestimmte Lösung, wenn die Koeffizientenmatrix invertierbar ist. Dieses ist genau dann der Fall, wenn die Spaltenvektoren linear unabhängig sind.*

BEWEIS. Das System ist genau dann eindeutig lösbar, wenn sich die Matrix durch Zeilenumformungen in die Einheitsmatrix überführen läßt, und das ist genau dann der Fall, wenn die Koeffizientenmatrix ein Produkt von Elementarmatrizen ist. Die Aussage über die lineare Unabhängigkeit haben wir auch schon gezeigt.  $\square$

**(5.9) Beispiel.** Die Lösung von

$$\begin{aligned} ax + by &= u \\ cx + dy &= v \end{aligned}$$

ist gegeben durch

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{D} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} D^{-1}(du - bv) \\ D^{-1}(av - cu) \end{pmatrix}.$$

**(5.10) Berechnung einer Basis des Lösungsraumes.** Sei  $A$  gegeben und  $S$  Zeilenstufenform von  $A$  mit den Stufen  $(j_1, \dots, j_k)$ . Dann ist  $Ax = 0$  äquivalent zu  $Sx = 0$ . Sei  $B$  die Matrix, die aus  $S$  durch Weglassen der Spalten zum Index  $j_1, \dots, j_k$  und der letzten  $n - k$  Zeilen entsteht. Die verbleibenden Spalten mögen die Nummern  $r_1, \dots, r_{n-k}$  tragen, d. h. es ist  $\{j_1, \dots, j_k\} \cup \{r_1, \dots, r_{n-k}\} = \{1, \dots, n\}$ . Dann ist  $Sx = 0$  äquivalent zu

$$\begin{pmatrix} x_{j_1} \\ \vdots \\ x_{j_k} \end{pmatrix} = -B \begin{pmatrix} x_{r_1} \\ \vdots \\ x_{r_{n-k}} \end{pmatrix}.$$

Setzt man rechts die Standardbasis des  $\mathbb{R}^{n-k}$  ein, so erhält man eine Basis des Lösungsraums.  $\diamond$

## 6 Lineare Abbildungen

### 6.1. Der Begriff einer linearen Abbildung

Wir schreiben jetzt Vektoren aus dem  $\mathbb{R}^n$  als Spaltenvektoren. Sei  $A = (a_{ij})$  eine  $m \times n$ -Matrix. Dann ist  $A\vec{x}$  für jedes  $\vec{x} \in \mathbb{R}^n$  ein Spaltenvektor aus  $\mathbb{R}^m$ . Wir erhalten somit eine Funktion (= Abbildung)

$$l_A: \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad \vec{x} \mapsto A\vec{x}.$$

Die Eigenschaften einer Funktion dieser Art nehmen wir zum Anlaß für die folgende wichtige Definition.

**(6.1) Definition.** Eine *lineare Abbildung* von  $\mathbb{R}^n$  nach  $\mathbb{R}^m$  ist eine Abbildung  $l: \mathbb{R}^n \rightarrow \mathbb{R}^m$  mit den Eigenschaften: Für alle  $\vec{x}, \vec{y} \in \mathbb{R}^n$  und alle  $\lambda \in \mathbb{R}$  gilt

- (1)  $l(\vec{x} + \vec{y}) = l(\vec{x}) + l(\vec{y})$
- (2)  $l(\lambda\vec{x}) = \lambda l(\vec{x})$ .

Aus den Eigenschaften (1) und (2) einer linearen Abbildung folgt (durch vollständige Induktion über  $r$ ):

$$(6.2) \quad l\left(\sum_{k=1}^r \lambda_k \vec{x}_k\right) = \sum_{k=1}^r \lambda_k l(\vec{x}_k),$$

die man so in Worten umreißt: Eine lineare Abbildung ist mit Linearkombinationen *verträglich* oder *vertauschbar*.

**(6.3) Satz.** Sei  $B = (\vec{x}_1, \dots, \vec{x}_n)$  eine Basis des  $\mathbb{R}^n$ . Eine lineare Abbildung  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  ist durch die Werte  $f(\vec{x}_j)$ ,  $1 \leq j \leq n$ , bestimmt, und diese Werte können auch beliebig vorgegeben werden.

BEWEIS. Die Eindeutigkeit ist aus (6.2) klar, und diese Formel sagt auch, daß man für gegebene  $\vec{b}_j \in \mathbb{R}^m$  durch

$$f\left(\sum_{j=1}^n \lambda_j \vec{x}_j\right) = \sum_{j=1}^n \lambda_j \vec{b}_j$$

eine lineare Abbildung definieren kann.  $\square$

**(6.4) Satz.** Ist  $l: \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine lineare Abbildung, so gibt es eine eindeutig bestimmte  $m \times n$ -Matrix  $A$ , mit der  $l = l_A$  ist.

BEWEIS. Der Bild  $l(\vec{e}_i)$  des  $i$ -ten Einheitsvektors  $\vec{e}_i \in \mathbb{R}^n$  ist eine Linearkombination  $l(\vec{e}_i) = \sum_{j=1}^m a_{ji} \vec{e}_j$  der Einheitsvektoren<sup>9</sup> des  $\mathbb{R}^m$ . Die damit gewonnene Matrix  $A = (a_{kl})$  ist die Gewünschte. Sie hat als  $i$ -ten Spaltenvektor den Vektor  $l(\vec{e}_i)$ .  $\square$

Wegen der Regel  $A(B\vec{x}) = (AB)\vec{x}$  gilt für die Verkettung linearer Abbildungen die Aussage

$$l_A \circ l_B = l_{AB},$$

denn Gleichheit der Funktionswerte an der Stelle  $\vec{x}$  ist gerade diese Regel, das heißt die Matrizenmultiplikation beschreibt genau die Verkettung linearer Abbildungen. Das ist der wesentliche Grund für die Definition des Matrizenproduktes. Da die Verkettung von Abbildungen offenbar assoziativ ist, sieht man jetzt auch ohne Rechnung, daß das Matrizenprodukt assoziativ ist.

**(6.5) Beispiel.** Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  eine lineare Abbildung. Dann gibt es genau einen Vektor  $\vec{a} \in \mathbb{R}^n$ , so daß gilt  $f(\vec{x}) = \langle \vec{a}, \vec{x} \rangle$ , denn das Skalarprodukt  $\langle \vec{a}, \vec{x} \rangle$  können wir auch als Matrizenprodukt mit der  $1 \times n$ -Matrix  $a$  lesen.  $\diamond$

## 6.2. Beschreibung durch Matrizen

Sei  $(\vec{a}_1, \dots, \vec{a}_n) = A$  eine Basis des  $\mathbb{R}^n$  und  $(\vec{b}_1, \dots, \vec{b}_m) = B$  eine Basis des  $\mathbb{R}^m$ . Ist  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  eine lineare Abbildung, so ist ihre Matrix  $M_A^B(f) = (a_{rs})$  bezüglich dieser Basen definiert durch

$$f(\vec{a}_j) = \sum_{k=1}^m a_{kj} \vec{b}_k.$$

In der  $j$ -ten Spalte stehen also die Komponenten des Bildvektors  $f(\vec{a}_j)$  bezüglich der Basis  $B$ .

Sind  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  und  $g: \mathbb{R}^m \rightarrow \mathbb{R}^p$  lineare Abbildungen und sind  $A, B, C$  Basen von  $\mathbb{R}^n, \mathbb{R}^m, \mathbb{R}^p$ , so gilt

$$(6.6) \quad M_B^C(g) M_A^B(f) = M_A^C(g \circ f),$$

d. h. auch hier beschreibt die Matrizenmultiplikation die Verkettung linearer Abbildungen. Die Basiswechselmatrix erscheint hier als Spezialfall der Matrixbeschreibung einer identischen Abbildung.

Die Matrixbeschreibung ist von der Basiswahl abhängig. Wie ändert sich die Matrix, wenn andere Basen  $A'$  für  $\mathbb{R}^n$  und  $B'$  für  $\mathbb{R}^m$  gewählt werden? Es gilt

$$(6.7) \quad M_{A'}^{B'}(f) = M_{B'}^{B'}(\text{id}) M_A^B(f) M_{A'}^A(\text{id}).$$

Das folgt aus der vorigen Formel, wenn wir die Basiswechselmatrix als Matrix einer identischen Abbildung ansehen.

Ist  $g: \mathbb{R}^m \rightarrow \mathbb{R}^m$  eine bijektive lineare Abbildung und  $S$  ein Unterraum, so haben  $S$  und  $g(S)$  dieselbe Dimension, denn eine Basis wird offenbar wieder auf eine Basis abgebildet.

<sup>9</sup>Man beachte, daß das Symbol  $\vec{e}_i$  etwas anderes bedeutet, je nachdem ob dieser Vektor im  $\mathbb{R}^n$  oder  $\mathbb{R}^m$  liegt.

Wird eine lineare Abbildung  $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$  in der Form  $\vec{x} \mapsto A\vec{x}$  beschrieben, so ist das Bild dieser Abbildung der Spaltenraum von  $A$ . Setzen wir  $f$  mit einer bijektiven linearen Abbildung  $g: \mathbb{R}^m \rightarrow \mathbb{R}^m$  zusammen, so haben also das Bild von  $f$  und  $gf$  dieselbe Dimension. Demnach ist für eine invertierbare Matrix  $B$  der Spaltenrang von  $BA$  und  $A$  derselbe. Wir haben schon früher gesehen, daß sich dabei auch nicht der Zeilenrang ändert. Also ändern sich beide Ränge nicht beim Übergang zur Zeilenstufenform. Für Zeilenstufenmatrizen erkennt man aber, daß beide Ränge gleich der Anzahl der Stufen sind. Folglich haben wir:

**(6.8) Satz.** *Für jede Matrix stimmen Zeilenrang und Spaltenrang überein.*  $\square$

### 6.3. Drehungen und Spiegelungen der Ebene

Wir beschreiben nun anhand der ebenen Geometrie in Beispielen die geometrische Bedeutung invertierbarer linearer Abbildungen.

Wir verwenden die Matrix

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

Multiplikation mit  $D(\varphi)$  dreht einen Vektor um den Winkel  $\varphi$  gegen den Uhrzeigersinn. Diese Aussage kann man als Definition dieser *Drehung* ansehen. Die folgende Rechnung bestätigt aber diese Definition. Sei  $\begin{pmatrix} u \\ v \end{pmatrix}$  ein Einheitsvektor, das heißt gelte  $u^2 + v^2 = 1$ . Dann ist

$$D(\varphi) \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u \cos \varphi - v \sin \varphi \\ u \sin \varphi + v \cos \varphi \end{pmatrix}.$$

Das Skalarprodukt mit dem Ausgangsvektor errechnet sich zu

$$u(u \cos \varphi - v \sin \varphi) + v(u \sin \varphi + v \cos \varphi) = (u^2 + v^2) \cos \varphi = \cos \varphi.$$

Das Skalarprodukt beschreibt aber nach unserer früheren Festlegung den Winkel.

Eine Drehung erst um  $\varphi$  und dann um  $\psi$  soll eine Drehung um  $\varphi + \psi$  sein. Nacheinanderausführung von linearen Abbildungen wird durch das Matrizenprodukt gegeben. Es soll also

$$D(\varphi)D(\psi) = D(\varphi + \psi)$$

gelten. Wenn man das Matrizenprodukt ausrechnet und auf beiden Seiten die sich entsprechenden Eintragungen der Matrizen vergleicht, so ist die letzte Gleichung äquivalent zu den Additionstheoremen für die Funktionen Sinus und Cosinus (siehe dazu II.6).

Eine  $2 \times 2$ -Matrix ist genau dann eine Drehmatrix, wenn sie die Form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad a^2 + b^2 = 1$$

hat.

Wir betrachten nun Matrizen der Form

$$S = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \quad a^2 + b^2 = 1.$$

Es gilt mit dieser Matrix  $S$

$$S \begin{pmatrix} a+1 \\ b \end{pmatrix} = \begin{pmatrix} a+1 \\ b \end{pmatrix}, \quad S \begin{pmatrix} -b \\ a+1 \end{pmatrix} = - \begin{pmatrix} -b \\ a+1 \end{pmatrix}.$$

Also ist  $S$  die Spiegelung an der durch  $\begin{pmatrix} a+1 \\ b \end{pmatrix}$  aufgespannten Geraden, denn diese Gerade bleibt fest und in der dazu senkrechten wird ein Vektor in sein Negatives verwandelt. (Durch diese Eigenschaften ist eine *Spiegelung* definiert.) Insbesondere gilt  $S \circ S =: S^2 = E_2$ . Wir schreiben Spiegelmatrizen in der Form

$$\begin{pmatrix} \cos 2\varphi & \sin 2\varphi \\ \sin 2\varphi & -\cos 2\varphi \end{pmatrix} = S(\varphi).$$

Der Grund ist aus den Relationen

$$S(\varphi) \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}, \quad S(\varphi) \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} = - \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}$$

ersichtlich;  $S(\varphi)$  ist die Spiegelung an der von  $\begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$  aufgespannten Geraden. Es gelten die folgenden Formeln:

$$\begin{aligned} D(\varphi)D(\psi) &= D(\varphi + \psi) \\ S(0)D(\varphi) &= D(-\varphi)S(0) \\ D(\varphi)S(0)D(-\varphi) &= S(\varphi) \\ S(\varphi)S(\psi) &= D(2\varphi - 2\psi) \\ S(\varphi) &= D(2\varphi)S(0) \\ S(\varphi)D(\psi) &= D(2\varphi - \psi)S(0). \end{aligned}$$

Daraus sehen wir: Jede Drehung ist Produkt zweier geeigneter Spiegelungen, und das Produkt zweier Spiegelungen ist immer eine Drehung.

Wir fassen nun die Ebene  $\mathbb{R}^2$  als Gaußsche Zahlenebene  $\mathbb{C}$  der komplexen Zahlen auf. Multiplikation mit  $z = a + bi$  ist eine lineare Abbildung  $\mathbb{C} \rightarrow \mathbb{C}$ , die bezüglich der Basis  $(1, i)$  von  $\mathbb{C}$  die Matrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

hat. Hat  $z$  den Betrag 1, so handelt es sich um eine Drehmatrix. Damit erkennen wir die geometrische Bedeutung der Multiplikation komplexer Zahlen: Multiplikation mit  $\cos \varphi + i \sin \varphi$  ist die Drehung um den Winkel  $\varphi$ .

#### 6.4. Scherungen. Dehnungen und Stauchungen

Scherungen sind durch Multiplikation mit Matrizen der Form

$$\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix}$$

gegeben. Geometrisch läßt sich die Wirkung der ersten Matrix so beschreiben: Die  $x$ -Achse bleibt fest. Eine Parallele zur  $x$ -Achse wird wieder auf dieselbe Parallele geworfen, dabei aber nach rechts ( $c > 0$ ) oder nach links ( $c < 0$ ) verschoben. In der Höhe 1 wird gerade um  $c$  verschoben. Wenn man sich  $x$ - und  $y$ -Achse als die beiden Arme

einer Schere vorstellt mit dem Nullpunkt als Gelenk, so wird also die  $y$ -Achse um dieses Gelenk nach Maßgabe von  $c$  gedreht, und die Parallelen zur  $x$ -Achse werden entsprechend mitverschoben.

Schließlich gibt es noch die Dehnungen und Stauchungen, die durch Diagonalmatrizen vermittelt werden.

## 7 Determinanten

### 7.1. Determinantenfunktion

Die Determinatenfunktion ordnet jeder  $n \times n$ -Matrix  $A$  eine Zahl  $\det(A)$  zu. Die Determinantenfunktion hat gewisse Eigenschaften, durch die sie eindeutig bestimmt ist. Im folgenden ist es zweckmäßig, eine  $n \times n$ -Matrix  $A$  auch durch ihre Spaltenvektoren zu bezeichnen, etwa in der Form  $A = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ , wenn  $\mathbf{a}_k$  die  $k$ -te Spalte ist.

**(7.1) Definition.** Die *Determinantenfunktion* ordnet jeder  $n \times n$ -Matrix  $A = (\mathbf{a}_1, \dots, \mathbf{a}_n)$  eine Zahl  $\det(A) = \det(\mathbf{a}_1, \dots, \mathbf{a}_n)$ , genannt ihre *Determinante*, so zu, daß gilt:

- (1)  $\det(\dots, \lambda' \mathbf{a}'_j + \lambda'' \mathbf{a}''_j, \dots) = \lambda' \det(\dots, \mathbf{a}'_j, \dots) + \lambda'' \det(\dots, \mathbf{a}''_j, \dots)$ .
- (2)  $\det(A) = 0$ , wenn  $A$  zwei gleiche Spalten hat.
- (3)  $\det(E_n) = 1$ .

In (1) stehen in allen drei Determinanten dieselben Spaltenvektoren an den von  $j$  verschiedenen Stellen; wir haben nur die an der Stelle  $j$  stehenden ausgeschrieben. Zur Eigenschaft (1) sagt man: Die Determinantenfunktion ist eine in jeder Spalte lineare Abbildung. Wegen (2) und des nächsten Satzes nennt man die Determinantenfunktion *alternierend*. Hat die Matrix die Einträge  $(a_{ij})$ , so wird die zugehörige Determinante ähnlich wie eine Matrix, nur mit senkrechten Strichen bezeichnet; also

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

für eine  $3 \times 3$ -Matrix  $(a_{ij})$ . ◇

Wir ziehen aus den Axiomen der Definition einige Folgerungen.

**(7.2) Satz.** *Werden zwei Spalten vertauscht, so ändert die Determinante ihr Vorzeichen.*

BEWEIS. Wir betrachten die ersten beiden Spalten und notieren die anderen nicht. Die Behauptung wird durch die folgende Rechnung geliefert:

$$\begin{aligned} 0 &= \det(\mathbf{a}_1 + \mathbf{a}_2, \mathbf{a}_1 + \mathbf{a}_2) \\ &= \det(\mathbf{a}_1, \mathbf{a}_1) + \det(\mathbf{a}_1, \mathbf{a}_2) + \det(\mathbf{a}_2, \mathbf{a}_1) + \det(\mathbf{a}_2, \mathbf{a}_2) \\ &= \det(\mathbf{a}_1, \mathbf{a}_2) + \det(\mathbf{a}_2, \mathbf{a}_1). \end{aligned}$$

Hierbei haben wir in der ersten Gleichung (2), in der zweiten (1) und in der dritten (2) ausgenutzt. □

**(7.3) Satz.** *Sind die Spaltenvektoren linear abhängig, so ist die Determinante Null.*

BEWEIS. Wir nehmen an, daß zum Beispiel  $\mathbf{a}_1 = \sum_{j=2}^n \mu_j \mathbf{a}_j$  ist. Wir setzen diese Summe für  $\mathbf{a}_1$  ein und benutzen dann (1) und (2).  $\square$

Die nächste Eigenschaft ist die Scherungsinvarianz der Determinante.

**(7.4) Satz.** *Wird das  $\lambda$ -fache einer Spalte zu einer anderen addiert, so ändert sich die Determinante nicht.*

BEWEIS. Wir betrachten wieder den typischen Fall von zwei Spalten

$$\det(\mathbf{a}_1 + \lambda \mathbf{a}_2, \mathbf{a}_2) = \det(\mathbf{a}_1, \mathbf{a}_2) + \lambda \det(\mathbf{a}_2, \mathbf{a}_2) = \det(\mathbf{a}_1, \mathbf{a}_2).$$

Darin haben wir (1) und (2) verwendet.  $\square$

**(7.5) Satz.** *Sei  $A$  eine Matrix mit Diagonale  $\lambda_1, \dots, \lambda_n$ . Dann ist ihre Determinante gleich dem Produkt der  $\lambda_j$ .*

BEWEIS. Das ist eine direkte Folge aus den Eigenschaften (1) und (3).  $\square$

Die Determinantenaxiome zeigen also, wie sich die Determinante bei elementaren Spaltenumformungen verhält. Das zeigt, daß man aus den Axiomen die Determinante ausrechnen kann.

**(7.6) Beispiel.** Wir zeigen, wie aus den bislang genannten Eigenschaften der Wert einer  $2 \times 2$ -Determinante entsteht:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Wir nutzen die Eigenschaft (1) der Definition mehrmals aus, sodann die Eigenschaft (2), die Folgerung (7.2) und schließlich (3):

$$\begin{aligned} \det A &= a \det(\vec{e}_1, \mathbf{a}_2) + c \det(\vec{e}_2, \mathbf{a}_2) \\ &= ab \det(\vec{e}_1, \vec{e}_1) + ad \det(\vec{e}_1, \vec{e}_2) + cb \det(\vec{e}_2, \vec{e}_1) + cd \det(\vec{e}_2, \vec{e}_2) \\ &= ad \det(\vec{e}_1, \vec{e}_2) + cb \det(\vec{e}_2, \vec{e}_1) \\ &= (ad - bc) \det(\vec{e}_1, \vec{e}_2) \\ &= ad - bc. \end{aligned}$$

**(7.7) Beispiel.** Für eine  $3 \times 3$ -Matrix ergibt sich in derselben Weise, nur etwa mühsamer hinzuschreiben:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} a_{11}a_{22}a_{33} & + & a_{12}a_{23}a_{31} & + & a_{13}a_{21}a_{32} \\ -a_{12}a_{21}a_{33} & - & a_{11}a_{23}a_{32} & - & a_{13}a_{22}a_{31}. \end{matrix}$$

Diese Formel kann man sich leicht für die praktische Rechnung merken, indem man sich einmal die „Lage“ der Produkte in der links ausgeschriebenen Determinante klarmacht.  $\diamond$

**(7.8) Beispiel.** Seien  $\vec{x} = (x_1, x_2, x_3)$ ,  $\vec{y} = (y_1, y_2, y_3)$ ,  $\vec{z} = (z_1, z_2, z_3)$  drei Vektoren, so daß  $\vec{y} - \vec{x}$  und  $\vec{z} - \vec{x}$  linear unabhängig sind. Dann ist

$$\begin{vmatrix} x & x_1 & x_2 & x_3 \\ y & y_1 & y_2 & y_3 \\ z & z_1 & z_2 & z_3 \\ 1 & 1 & 1 & 1 \end{vmatrix} = 0$$

die Gleichung einer Ebene durch diese drei Punkte.  $\diamond$



## 7.2. Der Entwicklungssatz

Für die Berechnung einer Determinante ist die sogenannte *Entwicklung nach einer Zeile oder Spalte* nützlich. Um sie formulieren zu können, bezeichnen wir mit  $A_{ik}$  die durch Streichen der  $i$ -ten Zeile und  $k$ -ten Spalte aus der  $n \times n$ -Matrix  $A = (a_{ik})$  entstehende  $(n-1) \times (n-1)$ -Matrix.

**(7.9) Satz.** *Die Entwicklung nach der  $k$ -ten Zeile ist die Formel*

$$\det(A) = \sum_{i=1}^n (-1)^{i+k} a_{ki} \det(A_{ki}).$$

*Die Entwicklung nach der  $k$ -ten Spalte ist die Formel*

$$\det(A) = \sum_{i=1}^n (-1)^{i+k} a_{ik} \det(A_{ik}).$$

Für die  $3 \times 3$ -Matrix wie in (6.5) sieht die Entwicklung nach der ersten Zeile also so aus:

$$a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Die Formel für die Entwicklung nach der  $k$ -ten Spalte kann man dazu benutzen, die Determinante induktiv nach  $n$  zu definieren. Durch Induktion weist man dann die drei Eigenschaften der Definition nach. Für die Eigenschaften (1) und (3) ist das unmittelbar klar; und zwei ist auch einfach, man muß es nur sorgfältig hinschreiben.

Für die praktische Berechnung der Determinante verwendet man häufig elementare Zeilen- oder Spaltenumformungen. Das ist dann analog zur Lösung von Gleichungssystemen.

## 7.3. Die Cramersche Regel

Wir setzen  $a_{ji}^{\#} = (-1)^{i+j} \det(A_{ij})$  und nennen die Matrix  $A^{\#} = (a_{ij}^{\#})$  die *Adjunkte* von  $A$ .

**(7.10) Satz.** Für eine  $(n, n)$ -Matrix  $A$  gilt  $A^{\#}A = AA^{\#} = \det(A) \cdot E_n$ .

BEWEIS. Der Eintrag von  $AA^{\#}$  an der Stelle  $(i, i)$  ist

$$\sum_j a_{ij} a_{ji}^{\#} = \sum_i a_{ij} \cdot (-1)^{i+j} \det(A_{ij}) = \det(A).$$

Der Eintrag an der Stelle  $(i, j)$  für  $i \neq j$  ist  $\sum_k (-1)^{j+k} a_{ik} \det(A_{jk})$ , also gleich der Entwicklung nach der  $j$ -ten Zeile einer Matrix, die aus  $A$  entsteht, indem die  $j$ -te Zeile durch die  $i$ -te Zeile ersetzt wird. Also ist dieser Eintrag als Determinante einer Matrix mit zwei gleichen Zeilen gleich Null. Die andere Behauptung folgt ebenso durch Entwicklung nach Spalten.  $\square$

**(7.11) Inverse Matrix.** *Ist  $A$  invertierbar, so liefert der letzte Satz die Formel*

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

für das Inverse von  $A$ .  $\square$

Nach dieser Formel wird man das Inverse fast niemals praktisch berechnen; aber es ist bemerkenswert, daß es überhaupt eine Formel gibt. Wir erhalten daraus auch eine Formel für die Lösung eines linearen Gleichungssystems.

**(7.12) Cramersche Regel.** Sei  $A$  eine invertierbare Matrix. Die Lösung des Gleichungssystems  $Ax = b$  ist  $x = A^{-1}b = |A|^{-1}A^\#b$ , also  $x_i = |A|^{-1} \sum_j (-1)^{i+j} b_j |A_{ji}|$ . Die letzte Summe ist aber nach dem Entwicklungssatz die Determinante einer Matrix, die aus  $A$  entsteht, wenn die  $i$ -te Spalte durch  $b$  ersetzt wird. Die damit gewonnene Formel für die Lösung wird *Cramersche Regel* genannt.  $\diamond$

#### 7.4. Der Produktsatz

**(7.13) Produktsatz.** Für zwei  $n \times n$ -Matrizen  $A$  und  $B$  gilt

$$\det(B) \det(A) = \det(BA).$$

BEWEIS. Sei diese Formel für Matrizen  $A_1$  und  $A_2$  und alle Matrizen  $B$  richtig. Dann gilt

$$\begin{aligned} \det(BA_1A_2) &= \det(BA_1) \det(A_2) \\ &= \det(B) \det(A_1) \det(A_2) \\ &= \det(B) \det(A_1A_2). \end{aligned}$$

Sie ist dann auch für das Produkt  $A_1A_2$  und alle  $B$  richtig. Nach den Grundeigenschaften der Determinante ist die Formel für Elementarmatrizen  $A$  und alle  $B$  richtig. Da jede invertierbare Matrix  $A$  Produkt von Elementarmatrizen ist, so gilt die Formel also für invertierbare  $A$  und alle  $B$ . Ist  $B$  nicht invertierbar, so sind ihre Spalten linear abhängig, und die Determinante ist dann Null. In diesem Fall sind aber auch die Spalten von  $BA$  linear abhängig. Also ist die Formel auch für diese  $B$  und alle  $A$  richtig.  $\square$

**(7.14) Satz.** Eine  $n \times n$ -Matrix  $A$  ist genau dann invertierbar, wenn ihre Determinante von Null verschieden ist. Dieses ist genau dann der Fall, wenn die Spaltenvektoren (bzw. die Zeilenvektoren) linear unabhängig sind.

BEWEIS. Ist  $A$  invertierbar, so ist nach dem Produktsatz die Determinante von Null verschieden, da die Determinante der Einheitsmatrix 1 ist. Ist die Determinante von Null verschieden, so folgt die Existenz des Inversen aus der Cramerschen Regel. Sind die Spaltenvektoren linear abhängig, so haben wir schon gesehen, daß die Determinante Null ist. Sind die Spaltenvektoren linear unabhängig, so bilden sie eine Basis. Basiswechsellmatrizen sind aber invertierbar.

**(7.15) Satz.** Eine Matrix und ihre Transponierte haben dieselbe Determinante  $\det(A) = \det(A^t)$ .

BEWEIS. Das ist richtig für Elementarmatrizen, also auch für deren Produkte, also für invertierbare  $A$ . Im anderen Fall benutzt man Zeilenrang = Spaltenrang.  $\square$

Wegen des letzten Satzes kann man die Axiome der Determinante auch analog für Zeilenvektoren formulieren.

### 7.5. Determinante von Blockmatrizen

Haben wir eine Blockmatrix

$$U = \begin{pmatrix} E_k & 0 \\ B & C \end{pmatrix}$$

mit quadratischer Matrix  $C$ , so liefert Entwicklung nach der ersten Zeile und Induktion nach  $k$  die Gleichung  $\det(U) = \det(C)$ . Ähnlich wird eine Matrix behandelt, in der  $C$  und  $E_k$  vertauscht vorkommen. Aus dem Produktsatz erhalten wir damit

$$\det \begin{pmatrix} A & 0 \\ B & C \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ B & E \end{pmatrix} \cdot \det \begin{pmatrix} E & 0 \\ 0 & C \end{pmatrix} = \det(A) \det(C),$$

wenn  $A$  und  $C$  quadratisch sind. Durch Transposition erhalten wir das analoge Ergebnis für eine Blockmatrix mit der Null links unten. Daraus entnehmen wir durch Induktion nach  $n$  für  $(n, n)$ -Dreiecksmatrizen  $D$  mit Diagonale  $d_1, \dots, d_n$  die Gleichung  $\det(D) = d_1 \cdot d_2 \cdot \dots \cdot d_n$ . Wir erkennen, daß eine solche Matrix genau dann invertierbar ist, wenn alle  $d_k$  von Null verschieden sind.  $\diamond$

### 7.6. Volumen

Eine geometrische Anwendung der Determinanten dient der Volumendefinition; dazu betrachten wir natürlich Vektorräume über  $\mathbb{R}$ . Seien  $b_1, \dots, b_n$  Vektoren des  $\mathbb{R}^n$ . Die Teilmenge

$$P(b_1, \dots, b_n) = \left\{ \sum_{i=1}^n \lambda_i b_i \mid 0 \leq \lambda_i \leq 1 \right\}$$

wird als der von  $b_1, \dots, b_n$  aufgespannte  $n$ -dimensionale Quader bezeichnet. Im Fall  $n = 2$  handelt es sich um ein Parallelogramm. Der Absolutbetrag

$$|\det(b_1, \dots, b_n)| = \text{Vol}P(b_1, \dots, b_n)$$

wird als *Volumen* von  $P(b_1, \dots, b_n)$  definiert. Im Fall  $n = 2$  bedeuten die Grundeigenschaften einer Determinante elementargeometrische Eigenschaften des Flächeninhaltes von Parallelogrammen. Man veranschauliche so etwa die Scherungsinvarianz.

**(7.16) Beispiel.** Seien  $(x_1, y_1), (x_2, y_2), (x_3, y_3)$  drei Punkte in der Ebene. Dann ist

$$\frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

der Flächeninhalt eines Dreiecks mit diesen drei Eckpunkten.  $\diamond$

### 7.7. Orientierung

Was bedeutet das Vorzeichen der Determinante? Es dient zur Definition einer subtilen geometrischen Eigenschaft: der Orientierung. Sei  $\mathcal{B}$  die Menge aller Basen des  $\mathbb{R}^n$ . Wir sagen, ein Paar  $B_1, B_2$  von Basen sei *gleich orientiert*, in Zeichen  $B_1 \simeq B_2$ , wenn die Basiswechsellmatrix von  $B_1$  nach  $B_2$  positive Determinante hat. Die Menge  $\mathcal{B}$  zerfällt in zwei Klassen  $\mathcal{B}_1$  und  $\mathcal{B}_2$ , und je zwei Matrizen in  $\mathcal{B}_j$  sind gleich orientiert. Wir nennen die  $\mathcal{B}_j$  die beiden *Orientierungen* von  $\mathbb{R}^n$ . Ist eine Orientierung  $\mathcal{B}_j$  festgelegt, so heißt

eine Basis in der Klasse  $\mathcal{B}_j$  *positiv* bezüglich dieser Orientierung; und *negativ*, wenn sie zur anderen Klasse gehört.

Im Eindimensionalen ist eine Orientierung anschaulich eine Durchlaufrichtung der Geraden. Im Zweidimensionalen ist eine Orientierung anschaulich ein Drehsinn (Uhrzeigersinn); die Determinante mit den Spaltenvektoren  $\vec{x}_1, \vec{x}_2$  ist genau dann positiv, wenn  $\vec{x}_1$  gegen den Uhrzeigersinn um einen Winkel kleiner als  $\pi$  in ein positives Vielfaches von  $\vec{x}_2$  gedreht werden kann. Im Dreidimensionalen ist eine Orientierung anschaulich ein Schraubensinn.

Ferner dient die Orientierung dazu, die Begriffe Rechts und Links festzulegen.

Eine Orientierung ist eine zusätzliche Struktur, eine „Vereinbarung“. Rechts-Links-Vereinbarungen kann man nicht durch das Telefon mitteilen sondern nur durch Vorzeigen.

## 8 Das Vektorprodukt

Wir betrachten den  $\mathbb{R}^3$  mit dem Standardskalarprodukt  $\langle -, - \rangle$ . Für  $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^3$  sei  $\det(\vec{x}, \vec{y}, \vec{z})$  die Determinante mit den Zeilenvektoren  $\vec{x}, \vec{y}$  und  $\vec{z}$ . Für feste  $\vec{x}, \vec{y} \in \mathbb{R}^3$  ist  $\mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $\vec{z} \mapsto \det(\vec{x}, \vec{y}, \vec{z})$  eine lineare Abbildung. Es gibt deshalb genau einen (von  $\vec{x}$  und  $\vec{y}$  abhängigen) Vektor  $w(\vec{x}, \vec{y})$ , so daß diese lineare Abbildung durch  $\vec{z} \mapsto \langle w(\vec{x}, \vec{y}), \vec{z} \rangle$  gegeben wird. Wir schreiben  $w(\vec{x}, \vec{y}) =: \vec{x} \times \vec{y}$  und nennen  $\vec{x} \times \vec{y}$  das *Vektorprodukt* des geordneten Paares  $(\vec{x}, \vec{y})$ . (Im Gegensatz zum Skalarprodukt ist also das Resultat wiederum ein Vektor). Definitionsgemäß gilt für alle  $\vec{x}, \vec{y}, \vec{z}$  die Gleichung

$$(8.1) \quad \langle \vec{x} \times \vec{y}, \vec{z} \rangle = \det(\vec{x}, \vec{y}, \vec{z}).$$

Aus der Definition folgen unmittelbar weitere Eigenschaften:

$$(8.2) \quad (\vec{x}_1 + \vec{x}_2) \times \vec{y} = \vec{x}_1 \times \vec{y} + \vec{x}_2 \times \vec{y}, \quad \lambda \vec{x} \times \vec{y} = \lambda(\vec{x} \times \vec{y}).$$

Entsprechendes gilt für die zweite Variable.

Da die Determinante alternierend ist, folgt

$$(8.3) \quad \vec{x} \times \vec{y} = -\vec{y} \times \vec{x}, \quad \vec{x} \times \vec{x} = \vec{0}$$

für alle  $\vec{x}$  und  $\vec{y}$ ; die Verknüpfung ist also nicht kommutativ. Nach der Definition (7.1) folgt  $\langle \vec{x} \times \vec{y}, \lambda \vec{x} + \mu \vec{y} \rangle = 0$ . Also gilt:

**(8.4) Satz.** *Der Vektor  $\vec{x} \times \vec{y}$  ist orthogonal zu dem von  $\vec{x}$  und  $\vec{y}$  aufgespannten Unterraum.*  $\square$

Sei  $(x_1, x_2, x_3) = \vec{x}$  und entsprechend für  $\vec{y}$ . Dann gilt:

**(8.5) Satz.** *Der Vektor  $\vec{x} \times \vec{y}$  hat die Komponentendarstellung*

$$\begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} \vec{e}_1 - \begin{vmatrix} x_1 & x_3 \\ y_1 & y_3 \end{vmatrix} \vec{e}_2 + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \vec{e}_3.$$

BEWEIS. Man bildet das Skalarprodukt mit  $\vec{z}$  und sieht, daß die Entwicklung von

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}$$

nach der dritten Zeile herauskommt.  $\square$

**(8.6) Folgerung.**  $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$ ,  $\vec{e}_1 \times \vec{e}_3 = -\vec{e}_2$ ,  $\vec{e}_2 \times \vec{e}_3 = \vec{e}_1$ .  $\square$

Jedes  $\vec{y}$  liefert eine lineare Abbildung  $\mathbb{R}^3 \rightarrow \mathbb{R}^3$ ,  $\vec{z} \mapsto \vec{y} \times \vec{z}$ . Setzen wir  $\vec{y} = (a, b, c)$ , so errechnet man die Matrix dieser linearen Abbildung zu

$$(8.7) \quad \varphi(\vec{y}) := \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}.$$

Für diese Matrix  $A$  gilt  $A^t = -A$ . Matrizen mit dieser Eigenschaft heißen *schiefsymmetrisch*. Das Vektorprodukt ist nicht assoziativ. Stattdessen gilt die

**(8.8) Jacobi-Identität.**  $\vec{x} \times (\vec{y} \times \vec{z}) + \vec{y} \times (\vec{z} \times \vec{x}) + \vec{z} \times (\vec{x} \times \vec{y}) = \vec{0}$  (zyklische Vertauschung von  $\vec{x}, \vec{y}, \vec{z}$  in den Summanden).

**(8.9) Satz.** *Das Vektorprodukt hat die folgenden Eigenschaften:*

- (1)  $\vec{x} \times \vec{y} = \vec{0}$  genau dann, wenn  $\vec{x}, \vec{y}$  linear abhängig sind.
- (2) Sind  $\vec{x}, \vec{y}$  linear unabhängig, so ist  $\vec{x}, \vec{y}, \vec{x} \times \vec{y}$  eine positiv orientierte Basis, d. h.  $\det(\vec{x}, \vec{y}, \vec{x} \times \vec{y}) > 0$ .
- (3) Aus  $\vec{x} \times \vec{y} = \vec{0}$  für alle  $\vec{y}$  folgt  $\vec{x} = \vec{0}$ .
- (4)  $\langle \vec{x} \times \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{y} \times \vec{z} \rangle$ .
- (5)  $\langle \vec{x}, \vec{y} \rangle^2 + \|\vec{x} \times \vec{y}\|^2 = \|\vec{x}\|^2 \|\vec{y}\|^2$  (Verschärfte Cauchy-Schwarzsche Ungleichung.)
- (6)  $\vec{x} \times (\vec{y} \times \vec{z}) = \langle \vec{x}, \vec{z} \rangle \vec{y} - \langle \vec{x}, \vec{y} \rangle \vec{z}$  (Graßmann-Identität).
- (7)  $(\vec{a} \times \vec{b}) \times (\vec{c} \times \vec{d}) = \det(\vec{a}, \vec{b}, \vec{c}) \vec{d} - \det(\vec{a}, \vec{b}, \vec{d}) \vec{c}$ .

BEWEIS. (1) Ist  $\vec{x} = \lambda \vec{y}$ , so folgt  $\vec{x} \times \vec{y} = \lambda \vec{y} \times \vec{y} = \vec{0}$ . Sind  $\vec{x}, \vec{y}$  linear unabhängig, so gibt es eine Basis der Form  $\vec{x}, \vec{y}, \vec{z}$ . Für diese ist  $\det(\vec{x}, \vec{y}, \vec{z}) \neq 0$  und deshalb ist  $\vec{x} \times \vec{y} \neq \vec{0}$ .

(2) Ist  $\vec{x} \times \vec{y} \neq \vec{0}$ , also  $\vec{x}, \vec{y}, \vec{x} \times \vec{y}$  eine positiv orientierte Basis.

(3) Sei  $\vec{x} \neq \vec{0}$ . Wir wählen  $\vec{x}, \vec{y}$  linear unabhängig. Dann ist  $\vec{x} \times \vec{y} \neq \vec{0}$ .

(4)  $\langle \vec{x} \times \vec{y}, \vec{z} \rangle = \det(\vec{x}, \vec{y}, \vec{z}) = \det(\vec{y}, \vec{z}, \vec{x}) = \langle \vec{y} \times \vec{z}, \vec{x} \rangle = \langle \vec{x}, \vec{y} \times \vec{z} \rangle$ .

(6) Die Abbildungen  $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , die  $(\vec{x}, \vec{y}, \vec{z})$  auf  $\vec{x} \times (\vec{y} \times \vec{z})$  bzw. auf  $\langle \vec{x}, \vec{z} \rangle \vec{y} - \langle \vec{x}, \vec{y} \rangle \vec{z}$  abbilden, sind linear in jeder Variablen. Um ihre Gleichheit zu zeigen, genügt es deshalb, Vektoren der Standardbasis einzusetzen. Das ist leicht, z. B.  $\vec{e}_2 \times (\vec{e}_2 \times \vec{e}_3) = \vec{e}_2 \times \vec{e}_1 = -\vec{e}_3$  und  $\langle \vec{e}_2, \vec{e}_3 \rangle \vec{e}_2 - \langle \vec{e}_2, \vec{e}_2 \rangle \vec{e}_3 = -\vec{e}_3$  etc.

(5) Nach (4) gilt  $\|\vec{x} \times \vec{y}\|^2 = \langle \vec{x} \times \vec{y}, \vec{x} \times \vec{y} \rangle = \langle \vec{x}, \vec{y} \times (\vec{x} \times \vec{y}) \rangle$ . Man setzt (6) ein und erhält das Gewünschte.

(7) Man wendet (6) an und danach (7.1).  $\square$

**(8.10) Satz.** *Sei  $A$  eine reelle  $(3, 3)$ -Matrix. Dann gilt*

$$A^t(A\vec{x} \times A\vec{y}) = \det(A)\vec{x} \times \vec{y}.$$

BEWEIS. Für alle  $\vec{z} \in \mathbb{R}^3$  haben wir die folgende Gleichungskette

$$\begin{aligned} \langle A^t(A\vec{x} \times A\vec{y}), \vec{z} \rangle &= \langle A\vec{x} \times A\vec{y}, A\vec{z} \rangle = \det(A\vec{x}, A\vec{y}, A\vec{z}) \\ &= \det(A) \det(\vec{x}, \vec{y}, \vec{z}) = \det(A) \langle \vec{x} \times \vec{y}, \vec{z} \rangle. \end{aligned}$$

Es folgt die Behauptung.  $\square$

Sei  $B = \{b_1, b_2, b_3\}$  eine Basis des  $\mathbb{R}^3$ . Für jeden Index  $k$  wählen wir  $i, j$  so, daß  $(i, j, k)$  eine gerade Permutation ist und setzen

$$b_i \times b_j = \sum_{\ell=1}^3 \gamma_{k\ell} b_\ell.$$

Wir erhalten nach (7.1).

$$\det(b_i, b_j, b_t) = \sum_{\ell} \gamma_{k\ell} \langle b_\ell, b_t \rangle = \det(B) \delta_{kt}.$$

Sei  $g = (g_{ij})$  mit  $g_{ij} = \langle b_i, b_j \rangle$ . Dann ist also die Matrix  $\gamma = (\gamma_{k\ell})$  durch

$$(8.11) \quad \gamma = \det(B) g^{-1}$$

gegeben. Die Matrix  $\gamma$  beschreibt, wie sich das Kreuzprodukt bezüglich irgendeiner Basis ausrechnen läßt. Ist  $B$  eine positive Orthonormalbasis, so ist  $g = \gamma$  die Einheitsmatrix.

**(8.12) Satz.** Für  $\vec{x}, \vec{y} \in \mathbb{R}^3$  ist  $\|\vec{x} \times \vec{y}\|$  der Flächeninhalt des von  $\vec{x}$  und  $\vec{y}$  aufgespannten Parallelogramms.

BEWEIS. Sind  $\vec{x}$  und  $\vec{y}$  linear abhängig, so sind  $|\vec{x} \times \vec{y}|$  und der Flächeninhalt gleich Null. Andernfalls sei  $E$  die von  $\vec{x}, \vec{y}$  aufgespannte Ebene und  $\vec{b}_1, \vec{b}_2, \vec{b}_3$  eine positive Orthonormalbasis mit  $\vec{b}_1, \vec{b}_2 \in E$ . Sind  $\vec{x} = \sum x_i \vec{b}_i, \vec{y} = \sum y_i \vec{b}_i$  die Komponentendarstellungen, so ist  $x_3 = y_3 = 0$  und  $\vec{x} \times \vec{y} = (x_1 y_2 - x_2 y_1) \vec{b}_3$ . Es ist aber definitionsgemäß

$$|x_1 y_2 - x_2 y_1| = \left| \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \right|$$

der Flächeninhalt des durch  $\vec{x}$  und  $\vec{y}$  aufgespannten Parallelogramms. Mittels  $\langle \vec{x}, \vec{y} \rangle = \|\vec{x}\| \|\vec{y}\| \cos \alpha$  mit dem Winkel  $\alpha$  zwischen  $\vec{x}$  und  $\vec{y}$  folgt  $\|\vec{x} \times \vec{y}\| = \|\vec{x}\| \|\vec{y}\| \sin \alpha, 0 \leq \alpha < \pi$ .  $\square$

## 9 Eigenwerte

### 9.1. Eigenvektoren. Eigenwerte

Sei  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  eine lineare Abbildung.

**(9.1) Definition.** Ein *Eigenvektor* von  $f$  zum *Eigenwert*  $\lambda$  ist ein von Null verschiedener Vektor  $\vec{x}$ , der die Gleichung  $f(\vec{x}) = \lambda \vec{x}$  erfüllt. Ein Eigenwert von  $f$  ist eine Zahl  $\lambda$ , für die der Unterraum  $V(\lambda) = \{\vec{x} \mid f(\vec{x}) = \lambda \vec{x}\}$  nicht der Nullraum ist; er heißt dann ein *Eigenraum*.  $\diamond$

Gibt es eine Basis aus Eigenvektoren, so wird  $f$  bezüglich dieser Basis durch eine Diagonalmatrix beschrieben. Das ist die denkbar einfachste Form einer Matrix.

Eine Drehmatrix in der Ebene hat keine Eigenwerte, es sei denn, sie vermittelt die identische Abbildung oder die antipodische  $x \mapsto -x$ .

## 9.2. Das charakteristische Polynom

Die Determinante ist dazu geeignet, die Eigenwerte einer  $n \times n$ -Matrix  $A$  zu bestimmen. Zu diesem Zweck betrachten wir die Determinante

$$\det(xE_n - A).$$

Der Wert dieser Determinante werde als Funktion von  $x$  betrachtet. Es gilt

**(9.2) Notiz.**  $\det(xE_n - A)$  ist ein Polynom vom Grad  $n$  in  $x$ . Es wird charakteristisches Polynom von  $A$  genannt.

BEWEIS. Das folgt durch Induktion nach  $n$  mittels des Entwicklungssatzes.  $\square$

**(9.3) Beispiel.** Das charakteristische Polynom einer  $2 \times 2$ -Matrix  $A$  wird durch

$$\det(xE_2 - A) = x^2 - \operatorname{Sp}(A)x + \det(A)$$

gegeben, wie eine Ausrechnung sofort zeigt. Hierin ist  $\operatorname{Sp}(A)$  die Summe der Diagonalelemente, die als *Spur* der Matrix bezeichnet wird.  $\diamond$

**(9.4) Satz.** Die Eigenwerte einer quadratischen Matrix sind genau die Nullstellen ihres charakteristischen Polynoms.

BEWEIS. Ist  $\lambda$  ein Eigenwert, so hat das Gleichungssystem  $(\lambda E_n - A)\vec{x} = \vec{0}$  eine von Null verschiedene Lösung; die Koeffizientenmatrix ist daher nicht invertierbar und hat deshalb die Determinante Null. Und umgekehrt.  $\square$