

Brückenkurs Mathematik

T. von der Twer

Inhaltsverzeichnis

Kapitel 1. Elementare Vorbereitungen	1
1. Grundlegende Zahlenmengen, Grundrechenarten	1
2. Umformen von Rechenausdrücken, Gleichungen und Ungleichungen	4
3. Rollen von Buchstaben	6
4. Arten von Formeln und Gleichungen	7
5. Übersicht zu den Rollen	9
6. Gleichungen für Geraden und Parabeln in der Ebene	9
Kapitel 2. Vektorrechnung und Analytische Geometrie	11
1. Die geometrischen Grundräume	11
2. Koordinatensysteme und Koordinatendarstellungen	12
3. Punktmengen und deren Quantifizierung als Koordinatentripelmengen	13
4. Die linearen Vektorraumoperationen	17
5. Schnitte linearer Gebilde und lineare Gleichungssysteme	27
6. Betrag, Skalarprodukt, Vektorprodukt, Spatprodukt	38
Kapitel 3. Die komplexen Zahlen	47
1. Der Körper \mathbb{C} der komplexen Zahlen	47
2. Motivierung der komplexen Zahlen	47
3. Konstruktion von \mathbb{C}	48
4. Das anschauliche Verständnis von \mathbb{C}	50
5. Multiplikation usw. in Polarkoordinaten	51
6. Lösen von einfachen Gleichungen in \mathbb{C}	52
Kapitel 4. Reelle Funktionen	55
1. Reelle Funktionen und ihre Graphen	55
2. Ableitung (Differentiation) von Funktionen	66
3. Integrale	84
Index	93

Elementare Vorbereitungen

1. Grundlegende Zahlenmengen, Grundrechenarten

Unter einer Menge verstehen wir naiv die Zusammenfassung von diversen Objekten zu einem Ganzen. Gehört ein Objekt x zur Menge M , so sagen wir: „ x ist Element von M “ und schreiben kurz dafür: $x \in M$. Gehört x nicht zu M , so schreiben wir kurz: $x \notin M$. Eine Menge kann man festlegen durch Aufzählen der Elemente, z.B. $M = \{1, 2, 3\}$. Wichtiger ist die Möglichkeit, eine Menge durch eine Eigenschaft zu definieren, z.B. $\{x \mid x \text{ ganze Zahl, } -10 \leq x \leq 10\}$. Lies: „Menge aller x , für die gilt: x ist ganze Zahl und...“ Gewöhnlich hat man dabei eine Grundmenge anzugeben, aus der die Objekte mit einer gewissen Eigenschaft auszusondern sind. Im Beispiel ist das die Menge der ganzen Zahlen, die man gern mit \mathbb{Z} bezeichnet. Also $\mathbb{Z} = \{\dots - 10, -9, -8, \dots - 1, 0, 1, \dots, 8, 9, 10, \dots\}$ (beidseitig ins Unendliche). Dann schreibt man etwa $M = \{z \in \mathbb{Z} \mid z = 3k \text{ für ein } k \in \mathbb{Z}\}$ (oder dasselbe mit einem Doppelpunkt oder Semikolon anstelle des senkrechten Strichs) für die Menge der durch 3 teilbaren ganzen Zahlen. Also $2 \notin M$, $-36 \in M$. Man mache sich bei dieser symbolischen Bezeichnung klar: *Hinter dem Strich spielt die Musik*, da kommt die wesentliche Eigenschaft, davor steht gar keine Information oder lediglich die Information, in welcher Grundmenge diese Eigenschaft betrachtet wird. (In \mathbb{R} , der Menge aller reellen Zahlen wäre Teilbarkeit durch 3 uninteressant, wieso?) Generell sollte man nie einfach „von links nach rechts“ lesen - man läuft dann immer Gefahr, die wesentlichen Informationen zu verpassen, nicht richtig zu ordnen usw. Das gilt insbesondere auch für Rechenausdrücke.

Einige Zahlen-Grundmengen werden sehr oft benutzt und haben daher Standard-Bezeichnungen:

- \mathbb{N} = $\{1, 2, 3, 4, 5, \dots\}$ Menge der natürlichen Zahlen, zuweilen bei 0 beginnend, manchmal auch \mathbb{N}_0 geschrieben für $\{0, 1, 2, 3, \dots\}$.
- \mathbb{Z} = Menge der ganzen Zahlen, also die negativen Zahlen zu \mathbb{N}_0 hinzu.
- \mathbb{Q} = Menge der rationalen Zahlen = $\left\{ \frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0 \right\}$
- \mathbb{R} = Menge der reellen Zahlen (sehr schwierig exakt zu definieren - grob gesagt sind das alle Zahlen, die sich beliebig gut durch rationale Zahlen nähern lassen; wichtige Beispiele sind $\sqrt{2}$, π , e . Sie sind reell, aber nicht rational. Man denke an Dezimaldarstellungen - diejenigen mit unendlicher nichtperiodischer Dezimalentwicklung sind reell, aber nicht rational.
- \mathbb{C} = Menge der komplexen Zahlen = $\{a + jb \mid a, b \in \mathbb{R}\}$, wobei $j^2 = -1$

Wesentliche Eigenschaften: In \mathbb{N} kann man nicht *allgemein* subtrahieren, aber in \mathbb{Z} . In \mathbb{Z} kann man nicht *allgemein* durch Zahlen $\neq 0$ dividieren, aber in \mathbb{Q} kann man es. \mathbb{Q} , \mathbb{R} und \mathbb{C} sind *Körper*. In allen Körpern beherrscht man das Rechnen über folgende Gesetze (Axiome), die speziell in \mathbb{Q} , \mathbb{R} und \mathbb{C} und definitionsgemäß

in allen Körpern erfüllt sind:

$$\begin{aligned}
 a + (b + c) &= (a + b) + c \\
 0 + a &= a \\
 -a + a &= 0 \\
 a + b &= b + a \\
 a \cdot (b \cdot c) &= (a \cdot b) \cdot c \\
 a \cdot (b + c) &= a \cdot b + a \cdot c \\
 1 \cdot a &= a \\
 a^{-1} \cdot a &= 1 \quad (a \neq 0) \\
 a \cdot b &= b \cdot a
 \end{aligned}$$

Die Allgemeingültigkeit dieser Formeln ist so zu verstehen und anzuwenden, daß man für jeden Buchstaben jede beliebige konkrete Zahlbezeichnung (für eine Zahl des jeweiligen Bereiches), aber auch jeden Rechenausdruck einsetzen kann und damit stets eine gültige Aussage bekommt. (Die Buchstaben mit diesem Zweck in allgemeingültigen Formeln nennt man **freie Variablen**.) Insbesondere folgt daraus das ganze Bruchrechnen, binomische Formeln usw. Noch ein wichtiger Unterschied: \mathbb{Q} und \mathbb{R} sind *angeordnete* Körper, nicht so aber \mathbb{C} . Was Anordnung bei einem Körper bedeutet, folgt im nächsten Abschnitt. Zunächst stellen wir noch die grundlegenden Gesetze der Bruchrechnung zusammen:

DEFINITION 1. $\frac{a}{b}$ ist für $b \neq 0$ definiert als $a \cdot b^{-1}$.

Der Sinn der Definition ist dieser: $\frac{a}{b}$ ist für $b \neq 0$ die *einzige, also eindeutige* Lösung der Bestimmungsgleichung $x \cdot b = a$. Denn: $\frac{a}{b}b = (ab^{-1})b = a(b^{-1}b) = a \cdot 1 = 1 \cdot a = a$. Und aus $x \cdot b = a$ folgt: $(xb)b^{-1} = ab^{-1}$, also $x = ab^{-1}$. Dagegen hat die Gleichung $x \cdot 0 = a$ für $a = 0$ jede reelle Zahl als Lösung, für $a \neq 0$ überhaupt keine.

Aus der Definition folgen mit den Körperaxiomen alle diese Grundgesetze der Bruchrechnung:

$$\begin{aligned}
 \frac{a}{b} + \frac{c}{d} &= \frac{ad + bc}{bd} \\
 \frac{a}{b} \cdot \frac{c}{d} &= \frac{ac}{bd} \\
 \frac{\frac{a}{b}}{\frac{c}{d}} &= \frac{ad}{bc} \\
 \frac{ac}{bc} &= \frac{a}{b} \\
 -\frac{a}{b} &= \frac{-a}{b} = \frac{a}{-b}
 \end{aligned}$$

Bei Doppelbrüchen ist stets darauf zu achten, dass die Klammerung klar ist - was also der Hauptbruchstrich ist. Das wird in obenstehender Formel einmal durch die größere Länge angezeigt, vor allem aber durch die Höhe des Gleichheitszeichens. Beim praktischen Addieren von Brüchen achte man auf die Verwendung des kleinsten gemeinsamen Vielfachen der Nenner als Hauptnenner. Die vierte Regel oben gibt sowohl für das Kürzen wie das Erweitern die Basis. Leiten wir als Beispiel die Additionsregel her: Zunächst beweise der Leser, dass stets $(ab)^{-1} = a^{-1}b^{-1}$. Dann haben wir:

$$\begin{aligned}
 \frac{ad + bc}{bd} &= (ad + bc)(bd)^{-1} = (ad + bc)(b^{-1}d^{-1}) \\
 &= adb^{-1}d^{-1} + bcb^{-1}d^{-1} = ab^{-1} + cd^{-1} = \frac{a}{b} + \frac{c}{d}.
 \end{aligned}$$

Dabei haben wir noch freien Gebrauch von Kommutativ- und Assoziativgesetz gemacht, ohne diese Schritte ins Einzelne „aufzudröseln“. Weiter nennen wir noch die wichtigen binomischen Formeln und ihre Verallgemeinerungen:

$$\begin{aligned}(a \pm b)^2 &= a^2 \pm 2ab + b^2, \quad (a+b)(a-b) = a^2 - b^2 \\(a+b)^n &= \sum_{i=0}^n \binom{n}{i} a^i b^{n-i}, \quad n \in \mathbb{N}. \text{ Dabei sind definiert:} \\ \binom{n}{i} &= \frac{n!}{i!(n-i)!}, \text{ lies: „}n \text{ über } i\text{“, wobei} \\ 0! &= 1, \quad (n+1)! = (n+1)n!, \text{ lies } n!: \text{ „}n \text{ Fakultät“,} \\ \sum_{i=0}^n a_i &= a_0 + \dots + a_n.\end{aligned}$$

Bemerkung: Sie sog. rekursive Definition von $n!$ in der dritten Zeile funktioniert so: Die erste Gleichung definiert, was $0!$ ist, die zweite führt den Wert von $(n+1)!$ auf den für $n!$ zurück. Damit ist die Sache für alle natürlichen Zahlen definiert, da sich diese aus der Null durch Nachfolgerbildung *alle* ergeben. Es ist natürlich sofort festzustellen, dass $n!$ für $n \geq 1$ einfach $1 \cdot \dots \cdot n$ ist. Also $2! = 2$, $3! = 6$, $4! = 24$, $5! = 120$. Die Binomialkoeffizienten $\binom{n}{i}$ sind dabei genau die, welche Sie aus dem sog. Pascalschen Dreieck kennen:

Exponent	Koeffizienten	Binom ausgeschrieben
$n = 0$	1	$(a+b)^0 = 1 \cdot a^0 b^0 = 1$
$n = 1$	1,1	$(a+b)^1 = 1 \cdot a^0 b^1 + 1 \cdot a^1 b^0 = a + b$
$n = 2$	1,2,1	$(a+b)^2 = 1 \cdot a^0 b^2 + 2 \cdot a^1 b^1 + 1 \cdot a^2 b^0$
$n = 3$	1,3,3,1	$(a+b)^3 = b^3 + 3ab^2 + 3a^2b + a^3$
$n = 4$	1,4,6,4,1	$(a+b)^4 = b^4 + 4ab^3 + 6a^2b^2 + 4a^3b + a^4$

Die einfache binomische Gleichung hängt eng mit quadratischer Ergänzung, diese wiederum z.B. mit der Lösungsformel für quadratische Gleichungen zusammen: Für $a \neq 0$ haben wir:

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

und aus

$$x^2 + px + q = 0, \quad p, q \in \mathbb{R}$$

folgt mit quadratischer Ergänzung

$$\begin{aligned}\left(x + \frac{p}{2}\right)^2 - \left(\frac{p}{2}\right)^2 + q &= 0, \text{ also} \\ x &= -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}, \text{ 3 Fälle:} \\ \left(\frac{p}{2}\right)^2 - q &> 0: \text{ Zwei reelle Lösungen,} \\ \left(\frac{p}{2}\right)^2 - q &= 0: \text{ Einzige Lösung } -\frac{p}{2} \\ \left(\frac{p}{2}\right)^2 - q &< 0: \text{ Keine reelle Lösung.}\end{aligned}$$

Abschließend verallgemeinern wir noch die Rechengesetze Distributiv- Assoziativ- und Kommutativgesetz auf mit großem Summenzeichen geschriebene Summen, auch Doppelsummen:

$$\begin{aligned} \sum_i ca_i &= c \sum_i a_i, \quad \sum_{i=1}^n c = n \cdot c \quad (n \in \mathbb{N}, n \geq 1) \\ \sum_i a_i + \sum_i b_i &= \sum_i (a_i + b_i) \\ \left(\sum_i a_i \right) \cdot \left(\sum_j b_j \right) &= \sum_i \left(\sum_j a_i b_j \right) = \sum_i a_i \left(\sum_j b_j \right) \\ &= \sum_j \left(\sum_i a_i b_j \right) = \sum_j b_j \left(\sum_i a_i \right) = \sum_{i,j} a_i b_j \end{aligned}$$

Dabei läuft i stets über dieselbe Indexmenge, ebenso j stets über dieselbe Indexmenge, die jedoch nicht mit der für i übereinstimmen muss. Eine Summe über leere Indexmenge ist definitionsgemäß Null, damit stimmt die zweite Formel auch für $n = 0$, wenn man die Indexmenge der natürlichen Zahlen ≥ 1 , die ≤ 0 sind, richtig als leere Menge versteht. Indizes (so heißt der Plural von *Index*) nimmt man stets ganzzahlig, aber man kann im Prinzip jede endliche Menge nehmen. Verallgemeinerungen werden jeweils aus dem Zusammenhang klar. Konkretisierendes Beispiel zur letzten Formel:

$$\begin{aligned} (a_1 + a_2)(b_1 + b_2 + b_3) &= a_1 b_1 + a_1 b_2 + a_1 b_3 + a_2 b_1 + a_2 b_2 + a_2 b_3 \\ &= a_1(b_1 + b_2 + b_3) + a_2(b_1 + b_2 + b_3) \\ &= b_1 a_1 + b_1 a_2 + b_2 a_1 + b_2 a_2 + b_3 a_1 + b_3 a_2 \\ &= b_1(a_1 + a_2) + b_2(a_1 + a_2) + b_3(a_1 + a_2) \end{aligned}$$

2. Umformen von Rechenausdrücken, Gleichungen und Ungleichungen

Basis des Umgangs mit Rechenausdrücken ist das *gleichwertige Umformen*. Man vollzieht das in Gleichungsketten, macht dabei gewöhnlich mehrere Schritte auf einmal. Konzentrierter Überblick ist wichtig! Z.B. $\frac{(x+y)^3}{x^2-y^2} + \frac{1}{x-y} = \frac{(x+y)^2+1}{x-y}$, $(a+b+c)^2 = a^2 + b^2 + c^2 + 2ab + 2ac + 2bc$. Jeder Einzelschritt besteht im Einsetzen in eine als allgemeingültig bekannte Gleichung. Wichtig ist ferner die allgemeingültige Regel, daß man stets Teilausdrücke durch gleichwertige ersetzen darf. Zu den Gesetzen des Rechnens mit Brüchen und Potenzen gesellt sich weiterhin die Nutzung von Formeln zu speziellen Funktionen, z.B. $\ln(3x^5) = \ln(3) + 5 \ln(x)$ (zum Logarithmus später mehr) sowie Formeln zum Ableiten etc.

Ähnlich funktioniert das gleichwertige Umformen von Gleichungen (gewöhnlich Bestimmungsgleichungen, die man lösen möchte). Dabei nutzt man jedoch nicht nur gleichwertiges Umformen beteiligter Rechenausdrücke, sondern zusätzlich spezielle Rechentechniken für spezielle Gleichungstypen (quadratische Gleichungen, Gleichungen wie $e^{2x+3} = 5$). Bewußt sollte man ein wichtiges allgemeines Mittel zur Umformung von Gleichungen verwenden und vorsichtig verstehen - das gilt nur für Gleichungen, nicht für Rechenausdrücke, und es ist vorsichtig und verständnisvoll zu verwenden: „Auf beiden Seiten dasselbe tun“. Dies ist zunächst einmal keine korrekt formulierte Regel, z.B. darf man sicher nicht auf beiden Seiten einer Gleichung immer dort, wo eine 2 steht, eine Eins hinsetzen. Hier ist ein besonders gravierender Fehler zu nennen, der von Anfängern *in der Mehrzahl (!)* tatsächlich gemacht wird: Man hat auf einer Seite einer Gleichung einen Bruch stehen, auf der andern Seite eine Summe (etwa auch von Brüchen) und möchte den Kehrwert bilden, da die Unbekannte etwa im Nenner steht: Der rücksichtslose Anfänger bildet gliedweise die Kehrwerte auf beiden Seiten, hat im oberflächlichen Sinne „dasselbe auf beiden Seiten“ gemacht und damit völligen Unsinn produziert. Gemeint sein muß genauer: Man darf auf beiden Seiten auf den dort stehenden Rechenausdruck dieselbe Funktion anwenden und erhält eine Gleichung, die aus der ursprünglichen folgt. (Achtung: Nicht immer folgt die ursprüngliche

auch aus der so umgeformten, diese Umformungen sind nur dann *gleichwertig*, wenn die angewandte Funktion eine umkehrbare ist.) Beispiel: Aus der Gleichung

$$\begin{aligned}\sqrt{x+2} &= x+1 \text{ folgt} \\ x+2 &= x^2+2x+1.\end{aligned}$$

Die Lösungen der letzteren sind $-\frac{1}{2} \pm \frac{1}{2}\sqrt{5}$. Aber unter $\sqrt{x+2}$ versteht man definitionsgemäß eine positive Zahl. Also ist nur $-\frac{1}{2} + \frac{1}{2}\sqrt{5}$ eine Lösung der ersten Gleichung. Die zweite Gleichung folgt aus der ersten, aber nicht umgekehrt. Aber immerhin konnten wir die einzige Lösung der ersten Gleichung finden, indem wir die zweite lösten und die für die erste Gleichung unbrauchbare Lösung wegwarfen. Das ist ganz typisch so.

Addiert man dagegen auf beiden Seiten einer Gleichung dieselbe Zahl, multipliziert man beide Seiten mit derselben Zahl $\neq 0$, so erhält man eine gleichwertige Umformung. Beispiel: $x+2 = \frac{3}{7}$, also $x = -\frac{11}{7}$. Es handelt sich um gleichwertige Umformung *der Gleichung*; man beachte: Nicht etwa sind die *Rechenausdrücke* dabei gleichwertig umgeformt worden, $x+2 = x$ ist offenbar Unsinn. Man vermeide den Anfängerfehler, dies Vorgehen auf Rechenausdrücke fälschlich zu übertragen: $\frac{f(x)}{x^2+1} = 0$ ist gleichwertig zu $f(x) = 0$, aber natürlich $\frac{f(x)}{x^2+1} \neq f(x)$, wenn nur $f(x) \neq 0$. Noch ein abschließendes Beispiel eines besonders wichtigen Typs (wieder gleichwertige Umformungen der Gleichungen):

$$\begin{aligned}2^{2x+3} &= 5 \\ (2x+3)\ln 2 &= \ln 5 \\ x &= \frac{\ln 5 - 3\ln 2}{2}\end{aligned}$$

Analog funktioniert es, wenn man eine Unbekannte aus einem Nenner, einer Wurzel herausholt usw.

Basis für den Umgang mit Ungleichungen sind die Gesetze (Axiome) für das Rechnen in angeordneten Körpern (wie \mathbb{Q} und \mathbb{R}). Diese lauten ($a \leq b$ bedeutet definitionsgemäß $a < b$ oder $a = b$):

$$\begin{aligned}a &< b \text{ oder } b < a \text{ oder } a = b \text{ (genau eins von diesen dreien trifft zu)} \\ \text{Wenn } a &< b \text{ und } b < c, \text{ dann } a < c \text{ (ebenso für } \leq \text{)} \\ \text{Wenn } a &< b, \text{ dann } a + c < b + c \text{ (ebenso für } \leq \text{)} \\ \text{Wenn } a &< b \text{ und } c > 0, \text{ dann } ac < bc\end{aligned}$$

Die letzteren beiden Regeln besagen, daß man beim Addieren und Multiplizieren auf beiden Seiten einer Ungleichung analog wie bei Gleichungen verfahren kann. Lediglich hat man darauf zu achten, daß beim Multiplizieren mit einer negativen Zahl auf beiden Seiten Umkehrung des Kleiner-Zeichens eintritt:

$$\text{Wenn } a < b \text{ und } c < 0, \text{ dann } ac > bc.$$

Übrigens kann man diese Regel aus den vorigen bereits herleiten, wir stellen sie nur wegen ihrer Bedeutung zur Vermeidung von Fehlern gesondert heraus. Denn aus $a > 0$ folgt $-a < 0$, da $-a = 0$ ausgeschlossen ist und aus $-a > 0$ und $a > 0$ folgen würde: $0 = -a + a > 0$, im Widerspruch zum ersten Axiom. Es bleibt also nur $-a < 0$ möglich, wenn $a > 0$. Aus $a < b$ und $c < 0$ folgt damit $a(-c) < b(-c)$ und daraus $ac > bc$.

Weitere Folgerungen sind etwa: Wenn $a < b$ und $c < d$, dann $a + c < b + d$; wenn $0 \leq a < b$ und $0 \leq c < d$, dann $ac < bd$.

Zu bemerken ist ferner, daß man wichtige weitere Ungleichungen aus der Kenntnis der Monotonie von Funktionen erhält, z.B. weiß: $0 < x < y$, dann $0 < x^2 < y^2$ oder: Wenn $x < y$, dann $e^x < e^y$. Übrigens erreicht man mit den feineren Mitteln der Analysis gerade die Herleitung von wichtigen Ungleichungen zur Abschätzung von Fehlern etc., die nicht so einfach einzusehen sind. Zunächst einmal werden wir Ungleichungen vor allem zur beschreibenden Einschränkung von Bereichen antreffen. In dieser Form sollten sie sich von selbst verstehen. Möchte man auch in bereits komplizierteren Fällen wissen, wie ein Bereich aussieht, der durch eine Ungleichung beschrieben wird, z.B. die Menge aller reellen Zahlen x , für die gilt: $x^2 - 2x - 3 < 0$, so wird man vielfach so

vorgehen, daß man stattdessen die zugehörige Gleichung löst, im Beispiel also $x^2 - 2x - 3 = 0$, $x_{1,2} = 1, 3$. Dann kennt man die Grenzen und weiß etwa im Beispiel: $x^2 - 2x - 3 < 0$ gilt genau für $1 < x < 3$.

Die schließlich noch wichtigen Rechengesetze für Exponenten besprechen wir im Zusammenhang mit Exponential- und Logarithmusfunktionen.

3. Rollen von Buchstaben

Die nützlichen elementaren Vorkenntnisse gliedern sich in drei Teile, die man zusammenwirken müssen, um zunächst einfachere, dann auch komplexere Probleme behandeln zu lassen: Rechnen, Umgang mit einfachsten geometrischen Gebilden wie Geraden und Parabeln in der Ebene, schließlich der verständige Einsatz von verschiedenen logischen Grundoperationen durch den sinnvollen Gebrauch von Buchstaben (sowohl beim Rechnen als auch bei der Beschreibung geometrischer Sachverhalte). Dazu ein

Beispiel: Warum ist das Rechteck mit dem größten Flächeninhalt bei gegebenem Umfang u ein Quadrat (der Seitenlänge $u/4$)? Wir rechnen dazu aus: Mit den Seitenlängen a, b kommt man auf den Flächeninhalt $F(a, b) = ab$ und hat die Bedingung $2a + 2b = u$, also $F(a, b) = a(u - 2a)/2 = -a^2 + au/2 = -(a - u/4)^2 + u^2/16$, und nun ist (ohne Differentialrechnung!) klar, dass $F(a, b)$ maximal wird für $a = b = u/4$. Derartige Überlegungen sollte man bei *neuen* Problemen ohne weiteres selbständig anstellen und die zugehörigen Rechnungen praktisch und korrekt ausführen können. Typisch treten jedoch Probleme bei Anfängern gerade in dieser Hinsicht auf, und das liegt daran, dass eben die besagten Vorkenntnisse fehlen. Analysieren wir, was man für das Beispiel tatsächlich benötigt: Zunächst ist u ein **äußerer Parameter**, mit einem beliebig festgesetzten positiven Wert, der durch die ganze Aufgabe gezogen wird. Mit a und b werden die Seitenlängen eines beliebigen Rechtecks bezeichnet. Diese Buchstaben treten jedoch ganz anders auf: Es sind **unabhängige Variablen**, genauer ist nur eine unabhängig, der Wert der anderen mit vorgegebenem Umfang u fixiert. (Das ist der Sinn des Rechenschrittes, der die Variable b eliminiert.) $F(a, b)$ tritt als **abhängige Variable** auf - daher die Funktionsschreibweise ' $F(a, b)$ '. So weit die Situationsbeschreibung: Flächeninhalt eines Rechtecks in Abhängigkeit von einer Seitenlänge bei vorgeschriebenem Umfang. Nun verlangt die Aufgabe, die variable Seitenlänge so einzurichten, dass der Flächeninhalt maximal wird. Damit tritt ein sehr typischer **Rollenwechsel** ein: a wird zur **Unbestimmten** (oder auch Unbekannten). Die Bedingung lautet: $-(a - u/4)^2 + u^2/16$ soll maximalen Wert erhalten. Diese Bedingung hat eine eindeutige Lösung, da jedes Quadrat reeller Zahlen positiv ist: $a = u/4$. Aus diesem einzigen Wert besteht die **Lösungsmenge** (oder Erfüllungsmenge) unserer Bedingung. Betrachten wir nun die Rechnung näher: Zunächst werden Flächeninhalt und Umfang eines Rechtecks mit den Seitenlängen a, b *allgemein* ausgedrückt. $F = ab$ und $u = 2a + 2b$ sind hier **unter einer speziellen Interpretation der beteiligten Größen allgemeingültige Formeln**. Die darin auftretenden Buchstaben sind **freie Variablen**, d.h. solche, für die man beliebige Werte *im Rahmen der vorgegebenen Deutung* einsetzen kann und dann stets gültige Aussagen erhält: Etwa für $a = 4$ und $b = 5$ erhält man $F = 20$, aber auch für $F = 3$ und $a = 1$ folgt $b = 3$. In eine solche allgemeingültige Formel kann man jedoch nicht nur konkrete Werte einsetzen, sondern viel wichtiger ist es, allgemeine Ausdrücke einzusetzen. Genau dies geschah oben: Die Gleichung $2a + 2b = u$ wurde nach b aufgelöst zu $b = (u - 2a)/2$, und dann wurde dieser Ausdruck für b in die Gleichung $F(a, b) = ab$ eingesetzt, mit dem Resultat $F(a, b) = a(u - 2a)/2$, was nur noch von der einen unabhängigen Variablen a abhängt und auch als $F(a)$ geschrieben werden könnte. Man beachte, dass hier bereits a einen ersten Rollenwechsel von der freien Variablen in der Formel zur unabhängigen Variablen hatte! Der nächste Schritt war das Ausmultiplizieren $a(u - 2a)/2 = -a^2 + au/2$. Dabei wurde die (ohne eine vorgeschriebene Deutung für alle Zahlen) allgemeingültige Rechenformel (Distributivgesetz!) $x(y + z) = xy + xz$ benutzt, wiederum mit Einsetzen von Ausdrücken: $a/2$ für x , u für y und $-2a$ für z . Zur Umordnung wurde noch die Formel $x + y = y + x$ verwandt. Dann wurde quadratische Ergänzung vorgenommen, gemäß der binomischen Formel $(x - y)^2 = x^2 - 2xy + y^2$. Für x setzen wir a ein, für y : $u/4$, dann folgt $(a - u/4)^2 = a^2 - au/2 + u^2/16$. Subtraktion von $u^2/16$ auf beiden Seiten und Multiplikation der Gleichung mit -1 ergibt das Resultat, an dem sich die Lösung der Aufgabe ohne weiteres ab lesen ließ. (Damit sind noch immer nicht alle einzelnen Formelanwendungen genannt.)

Einige weitere Rollen von Buchstaben kamen im Beispiel noch nicht vor:

Konstanten wie π oder die Eulerzahl e , auch Naturkonstanten wie die Lichtgeschwindigkeit bezeichnet man mit Buchstaben. Stets sollte man auch damit Rechnungen ausführen und allenfalls im Endresultat Näherungswerte einsetzen; denn diese Zahlen hören in ihrer Dezimalentwicklung hinter dem Komma nicht auf, und sie sind nicht als Brüche darstellbar. Rechnungen mit Näherungswerten produzieren alsbald untragbare Fehler. (Ähnlich verhält es sich mit $\sqrt{2}$ usw.)

Hilfsvariablen dienen dazu, Rechenausdrücke abzukürzen, zu vereinfachen, gezielt ein Verhalten zu untersuchen oder strategisch Rechnungen auszuführen. Zum Beispiel wird man in der Bestimmungsgleichung $x^4 - 4x^2 + 3 = 0$ die neue Unbestimmte $u = x^2$ einführen, um eine quadratische Gleichung zu erhalten. Oder man setzt im Ausdruck

$$f(x) = \frac{1 + \sqrt{1 + \left(\frac{x}{x+1}\right)^2}}{\sqrt{1 + \left(\frac{x}{x+1}\right)^2}}$$

für die Wurzel die neue unabhängige Variable a und erhält den einfachen Ausdruck $(a+1)/a$. Daran kann man sofort sehen, dass für $a \geq 1$ nur Werte resultieren im Bereich $]1, 2]$ (das ist das Intervall von 1 bis zwei, wobei 1 ausgeschlossen und 2 eingeschlossen ist). Also hat f genau alle Werte in diesem Intervall, da die Wurzel alle Werte im Bereich $[1, \infty[$ annimmt für $x \in \mathbb{R}$.

Freie Parameter benutzt man gerade in der Naturwissenschaft gern, um Mengen zu beschreiben. Die Lösungsmenge der Gleichung $2x + 3y = 1$ im Reellen ist die Menge aller reellen Zahlenpaare, welche diese Gleichung erfüllen, also $\{(x, y) \mid 2x + 3y = 1\}$. Nun rechnet man diese Menge aus, indem man feststellt, dass man mit beliebig festgelegtem $x \in \mathbb{R}$ genau ein Lösungspaar $(x, (1 - 2x)/3)$ erhält. Nun schreibt man gern die Lösungsmenge auf in der Form

$$L(x) = \left(x, \frac{1}{3} - \frac{2}{3}x\right), \quad x \in \mathbb{R}.$$

Die Werte des freien Parameters x durchlaufen alle reellen Zahlen, und damit durchläuft $L(x)$ die Lösungsmenge der Gleichung.

Stumme oder gebundene Variablen treten z.B. dann auf, wenn man davon spricht, mit allen Objekten eines Bereiches eine Operation vorzunehmen. Etwa bedeutet der Rechenausdruck

$$\sum_{i=1}^{10} i \quad (= 55),$$

dass *alle* ganzen Zahlen von 1 bis 10 zu addieren sind, mit dem Resultat 55. Es wäre also Unfug, zu fragen, welche Zahl i denn nun bedeute. Ebenso wenig wäre eine Gleichung mit dieser Summe nach i aufzulösen. Ebenso bedeutet $\int_0^1 x^2 dx$, dass die Funktion $f(x) = x^2$ im Bereich von 0 bis 1 zu integrieren ist. Wieder spielen *alle* zugehörigen Funktionswerte hinein.

Man wird bemerkt haben, dass die Rollen von Buchstaben eng mit Arten von Gleichungen oder Formeln verknüpft sind, in denen die Buchstaben vorkommen. Der logische Gebrauch dieser Gleichungen muss unbedingt verstanden werden.

4. Arten von Formeln und Gleichungen

Allgemeingültige Formeln: Man setzt beliebige Werte oder Rechenausdrücke für die **freien Variablen** ein und erhält stets eine wahre bzw. allgemeingültige Aussage. Beispiel: Aus der Formel $(x+y)^2 = x^2 + 2xy + y^2$ erhält man durch Einsetzen von $-y$ für y die wiederum allgemeingültige Formel $(x-y)^2 = x^2 - 2xy + y^2$. Man beachte, dass man *jeden* Ausdruck einsetzen darf, nicht etwa wird hier die Behauptung $y = -y$ gemacht, die außer für $y = 0$ falsch ist. Beim Einsetzen ist darauf zu achten, dass man zunächst das Eingesetzte einzuklammern hat und dann zuweilen überlegt diese Klammern ersparen kann, Beispiel: Setzt man oben $a + b$

für y , so kommt $(x + (a + b))^2 = x^2 + 2x(a + b) + (a + b)^2$ heraus - die Klammer um $a + b$ auf der linken Seite ist überflüssig, auf der rechten Seite benötigt man beide Klammern.

Allgemeingültige Formeln unter einer gegebenen Deutung: Die Formel $c^2 = a^2 + b^2$ (Pythagoras) gilt nicht allgemein - setzt man beliebige Zahlen für a, b, c , so wird man eine falsche Gleichung erhalten. Aber unter der Voraussetzung, dass a, b die Kathetenlängen und c die Hypotenusenlänge eines rechtwinkligen Dreiecks sind, gilt die Formel *allgemein*.

Definitivische (allgemeine) Formeln: Man definiert z.B. $f(x) = x^2$ für einen Zusammenhang. Solche definitivischen Formeln sind wie allgemeingültige zu behandeln. Im Beispiel folgt etwa $f(a + b) = (a + b)^2$. (Wieder die Klammer!)

Bestimmungsgleichungen: Eine Gleichung wie $x^2 + 2x - 1 = 0$ ist nicht etwa allgemeingültig. Vielmehr stellt sie eine Bedingung an x , die in diesem Falle von genau den beiden Zahlen $x_{1,2} = -1 \pm \sqrt{2}$ erfüllt wird. Man hat eine Bedingung an eine Größe und sucht die Lösungsmenge. Entscheidend für die logische Behandlung ist die Frage: Welche Buchstaben sind Unbestimmte? Eine Bestimmungsgleichung kann durchaus noch äußere Parameter enthalten, und dann sollte man nicht etwa nach diesen auflösen. Dazu ein

Beispiel: Welche Gerade der Form $y = mx + b$ geht durch die Punkte (x_0, y_0) und (x_1, y_1) , und unter welchen Umständen gibt es überhaupt eindeutig eine solche Gerade? In dieser Aufgabe treten m, b als Unbestimmte und $x_{0,1}$ sowie $y_{0,1}$ als äußere Parameter auf. Wir haben ein System von zwei Bestimmungsgleichungen:

$$\begin{aligned} y_0 &= mx_0 + b \\ y_1 &= mx_1 + b \end{aligned}$$

und rechnen nach Subtraktion beider Gleichungen aus:

$$\begin{aligned} m &= \frac{y_1 - y_0}{x_1 - x_0}, \text{ für } x_1 \neq x_0. \\ b &= y_0 - \frac{y_1 - y_0}{x_1 - x_0} x_0. \end{aligned}$$

Für $x_1 \neq x_0$ haben wir also eine eindeutige Lösung, und die hängt in der angegebenen Weise von den äußeren Parametern ab. Der Fall $x_1 = x_0$ bleibt noch zu untersuchen, in dem der Bruch keinen Sinn macht. Für $y_1 \neq y_0$ haben wir eine Gerade senkrecht zur x -Achse, und es gibt *keinen* Ausdruck $y = mx + b$, der die Gerade beschreibt. Für $y_1 = y_0$ dagegen haben wir unendlich viele Geraden der Form $y = mx + b$, welche durch den einen vorgeschriebenen Punkt gehen, unser Bedingungssystem reduziert sich auf die einzige Gleichung $y_0 = mx_0 + b$, und für *jede* Zahl m erhalten wir eine solche Gerade, indem wir $b = y_0 - mx_0$ setzen. Solche Fallunterscheidungen treten typisch bei Problemen mit äußeren Parametern auf.

Man beachte, dass das Gleichungssystem

$$\begin{aligned} y_0 &= mx_0 + b \\ y_1 &= mx_1 + b \end{aligned}$$

mit veränderten Rollen eine völlig andere Aufgabe definiert: Seien nun m, b sowie x_0, x_1 äußere Parameter, y_0, y_1 Unbestimmte. Dann handelt es sich inhaltlich um die Frage, welche y -Werte zu den vorgegebenen x -Werten gehören auf der Geraden mit Steigung m und dem y -Achsenabschnitt b . Es ist nichts weiter zu rechnen, die Gleichungen sind in Endform für diese Aufgabe.

Noch einmal zurück zur ersten Aufgabe: Man hat zwei Unbestimmte, m, b . Die **Lösungsmenge des Gleichungssystems** ist dann die Menge alle *Paare* (m, b) , welche das System erfüllen. Zusammengefasst lautet das Resultat (typisches Auftreten von Fallunterscheidungen bei äußeren Parametern!):

$$\begin{aligned} \text{Für } x_0 &\neq x_1 \text{ ist } L(x_0, x_1, y_0, y_1) = \left\{ \left(\frac{y_1 - y_0}{x_1 - x_0}, y_0 - \frac{y_1 - y_0}{x_1 - x_0} x_0 \right) \right\}, \\ \text{für } x_0 &= x_1, y_0 = y_1 \text{ ist } L(x_0, x_0, y_0, y_0) = \{(m, y_0 - mx_0) \mid m \in \mathbb{R}\}, \\ \text{für } x_0 &= x_1, y_0 \neq y_1 \text{ ist } L(x_0, x_0, y_0, y_1) = \emptyset \text{ (leere Menge)}. \end{aligned}$$

5. Übersicht zu den Rollen

Nun zur Übersicht über die Rollen von Buchstaben in der Mathematik und ihren Anwendungen in entsprechenden Formeln und Gleichungen:

Art der Gleichung	typische Rollen darin	typische Aktionen
definitorische Formel allgemeingültige Formel (Axiome oder herleitbar)	freie Variablen → unabhängige Var. → abhängige Variablen	Einsetzung von Zahlen und vor allem Rechen- ausdrücken (Termen)
Bestimmungsgleichung (auch System davon)	Unbestimmte → 1 äußere Parameter → 2	1 Auflösen danach 2 Mitschleppen
Lösungsformel bzw. Formel für Lösungen	freie Parameter → 1 äußere Parameter → 2	1 Einsetzen 2 bleiben in ihrer Funktion

Weitere Rollen von Buchstaben in Formeln und Gleichungen

Rolle	Vorkommen	Aktion
Konstante	überall möglich	Mitschleppen, ev. (Näherungs-) Wert einsetzen
Bezeichnung	überall	ein beliebiges gedachtes Objekt wird bezeichnet
Hilfsvariable	überall	ein komplizierter Ausdruck wird abgekürzt, eine Gleichung vereinfacht, usw.
stumme bzw. gebundene Variable	vielfach, insbesondere in Summen und Integralen	keine speziellen Einsetzungen, sondern Durch- führen der zugeh. Aktion mit <i>allen</i> Objekten des zugehörigen Bereichs

6. Gleichungen für Geraden und Parabeln in der Ebene

Jede der Formen von Geradengleichungen hat ihre eigene Nützlichkeit:

$$\begin{aligned}
 ax + by &= c, a \neq 0 \text{ oder } b \neq 0 : \text{Bestimmungsgleichung, implizite Form} \\
 \frac{x}{a} + \frac{y}{b} &= 1, a, b \neq 0 : \text{Achsenabschnittsform} \\
 y &= mx + b : \text{explizite Form, analog } f(x) = mx + b \text{ (Funktionsform)} \\
 y &= y_0 + \frac{y_1 - y_0}{x_1 - x_0} (x - x_0), x_0 \neq x_1 : \text{Zweipunkte-Form} \\
 y &= y_0 + m(x - x_0), x_0 \neq x_1 : \text{Punkt-Richtungs-Form}
 \end{aligned}$$

Zugehörige Lösungsmengen (Unbestimmte x, y) - Beispiele für das Aufschreiben unendlicher Lösungsmengen mittels freier Parameter:

$$L(x) = (x, mx + b), x \in \mathbb{R}, \text{ für } y = mx + b,$$

$$L(y) = \left(\frac{c}{a}, y\right), x \in \mathbb{R}, \text{ für } ax = c \text{ (Gerade } \parallel y\text{-Achse)}$$

Für Parabeln in der Ebene besprechen wir hier nur die einfachen Formen, die geeignet sind, Parabeln zu beschreiben, welche ihre Achse parallel zur y -Achse haben (später mehr in der Vektorrechnung):

$$y = ax^2 + bx + c, a \neq 0 : \text{ allgemeine algebraische Form}$$

$$y = \alpha(x - x_0)^2 + y_0, \alpha \neq 0 : \text{ Scheitelpunktsform,}$$

mit Scheitelpunkt (x_0, y_0) in Koord.-Darstellung

Die Scheitelpunktsform entwickelt man leicht durch quadratische Ergänzung aus der ersten Form. Lösungsmenge der ersten Gleichung (mit x, y als Unbestimmten) ist $L(x) = (x, ax^2 + bx + c), x \in \mathbb{R}$.

Vektorrechnung und Analytische Geometrie

Es sind physikalisch-technische Probleme und Sachverhalte mathematisch zu beschreiben. Zu diesem Zweck bringt man zunächst einmal geometrische Situationen in geeigneten Räumen unter. Von Anfang hat man dazu zwei Möglichkeiten, die man stets kombinieren sollte, im Sinne der wichtigsten gedanklichen Verbindung, die man geistig zur Verfügung hat: Räumliche Anschauung und Rechnung. Entsprechend hat man die rein geometrischen *Punkträume* sowie *geometrischen Vektorräume* und auf der andern Seite die *Zahlentupelräume* zur quantitativen Erfassung. Im elementaren mathematischen Unterricht fehlt diese Trennung: Bei einfachsten eindimensionalen Problemen kommt man damit aus, eine unmittelbare Anschauung (die nicht näher mathematisch erfasst wird) mit einer Rechnung zu verbinden. Aus der geometrischen Sachlage werden lediglich eindimensionale Eigenschaften für Zahlbeschreibungen herausgezogen, die selber keine weitere geometrische Struktur haben und daher nicht eigens in einem mathematischen geometrischen Raum zu deuten sind. Bei komplexeren mehrdimensionalen Problemen benötigt man dagegen eine mathematische Erfassung der geometrischen Sachverhalte selbst, die *ohne Zahlbeschreibungen* auskommt. Warum ist das so? Tatsächlich gibt es eine Fülle von Möglichkeiten (konkret: Koordinatensysteme), eine zahlenmäßige Beschreibung anzusetzen, und man hat sich oft innerhalb einer Aufgabe mehrerer verschiedener zu bedienen. Das kann man aber nicht tun, wenn man die Zahlenbeschreibung überhaupt nicht vom Sachverhalt selbst getrennt und diese Möglichkeiten verstanden hat. Überdies gibt es viele Probleme, die man eleganter mit absoluten geometrischen Beschreibung selbst lösen kann, ohne überhaupt zu einer Koordinatenbeschreibung überzugehen.

1. Die geometrischen Grundräume

1.1. Die Räume E^2 und E^3 . Darunter verstehen wir zunächst einfach die Anschauungsräume, die Menge aller Punkte, und zwar ist E^2 die Menge aller Punkte in einer anschaulichen Ebene, E^3 die Menge aller Punkte im anschaulichen dreidimensionalen Raum. Tatsächlich gehört noch die geometrische Struktur dazu, die wir später mit den Verbindungsvektoren zwischen zwei Punkten erhalten. Die Elemente der Räume $E^{2,3}$, also die Punkte bezeichnen wir mit P, Q, R, P_1 usw.

1.2. Die Räume V_O^2 und V_O^3 der anschaulichen Ortsvektoren. Wir sprechen nur über V_O^3 , für zwei Dimensionen funktioniert die Sache analog. Wir fixieren einen Punkt O im E^3 als Ursprung. Dann kann man zu jedem Punkt $P \in E^3$ den Ortsvektor von P (bezüglich O) \vec{x}_P bilden: Das ist der geometrische Pfeil, der von O ausgeht und mit seiner Spitze in P endet. V_O^3 ist die Menge all dieser Pfeile. Offenbar gehört zu jedem Punkt genau ein Ortsvektor (bei fixiertem $O!$) und zu jedem Ortsvektor genau ein Punkt. Ein Ortsvektor ist an einen festen Ursprung gebunden, im Gegensatz zu den gleich einzuführenden freien Vektoren.

Warum ist das nützlich? Wir werden sehen, dass man mit Ortsvektoren und freien Vektoren wichtige rechnerische Verbindungen zwischen Punkten herstellen kann, die mit Punkten keinen Sinn machen. Überdies werden ganz andere Dinge als Punkte mittels freier Vektoren beschrieben, und es wird sich zeigen, dass man mit beiderlei Vektoren in gleicher Weise umgehen kann.

1.3. Die Räume V_O^2, V_O^3 der freien Vektoren. Es gibt eine Fülle von Objekten, die man angemessen mit freien Vektoren zu beschreiben hat: Z. B. gehört zu einer Geschwindigkeit eines Teilchens ein Betrag und eine Richtung im Raum. Diese Information lässt sich zu einem geometrischen Pfeil zusammenfassen, dessen Länge den Betrag angibt und dessen Richtung eben die Richtung der Geschwindigkeit angibt. Oder man möchte

eine Parallelverschiebung im Raum angeben: Jeder Punkt des Raums wird in derselben Richtung um dieselbe Streckenlänge verschoben. Oder ganz ähnlich: Die Beziehung der Punkte P, Q kann man durch den *Vektor von P nach Q* (Bezeichnung: \overrightarrow{PQ}) beschreiben, den Pfeil, der in P beginnt und in Q endet. Diesen Pfeil fasst man als Verschiebungsvektor auf. Ein besonders wichtiges Beispiel ist die Beschreibung physikalischer Kräfte: Auch zu einer Kraft gehört ein Betrag und eine Richtung im Raum.

Warum sprechen wir von *freien* Vektoren in Beispielen wie den angegebenen? Der Grund liegt darin, dass eine Ortsfixierung nicht zur gewünschten Information gehört: Verschiebt man einen Kraftpfeil parallel, so zeigt er immer noch dieselbe Länge und Richtung. Daher ist ein freier Vektor auch genau genommen kein Pfeil, sondern er wird nur *repräsentiert*, d.h. etwa in einer Skizze dargestellt mit einem Pfeil. Verschiebt man den darstellenden Pfeil *parallel*, so erhält man eine *andere Darstellung desselben freien Vektors!*

Freie Vektoren bezeichnen wir mit \vec{x} , \vec{y} , \vec{x}_1 usw.

1.4. Ein gefährlicher Anfängerfehler im Umgang mit Vektoren: Zeichnet man einen Pfeil, so zeichnet man insbesondere eine Strecke. Diese Tatsache verführt immer wieder Anfänger zu der Ansicht, mit einem Vektor sei insbesondere eine Strecke beschrieben. Aber das ist nicht so: Der Pfeil eines freien Vektors ist beliebig verschiebbar, der freie Vektor enthält also keinerlei Ortsinformation! (Es kommen sogar Argumentationen bei Anfängern vor der Art: „Setze ich den (freien!) Vektor \vec{a} so..., dann...“. Das ist reiner Unsinn und resultiert in fehlerhaften Lösungen bzw. im Nichtlösen einer gestellten Aufgabe.) Wie steht es mit gebundenen Ortsvektoren? Die enthalten doch tatsächlich Ortsinformation, aber man sollte verstehen, welche das genau ist. Zunächst ist ein Ursprung O fixiert, und dann gibt der Ortsvektor die Lage seines Endpunktes P relativ zu O . Das ist alles. Wenn man aber eine Strecke beschreiben will, so dass man sie etwa mit einer gegebenen Geraden schneiden kann, dann benötigt man eine Beschreibung, die später unter dem Namen *Parameterdarstellung* einzuführen ist. Dabei stellt man *jeden Punkt* einer Strecke durch seinen eigenen Ortsvektor dar. Zwar kann man diese Beschreibung jeweils aus der Information der Endpunkte der fraglichen Strecke gewinnen, aber man muss sie doch eigens daraus herstellen. Dagegen ist es völlig sinnlos, wie viele Anfänger etwa zu formulieren: „Vektor \vec{a} schneidet Vektor \vec{b} “, aufgrund des Irrtums, beide Vektoren seien etwa geeignet, zwei Strecken zu beschreiben. Noch einmal ausdrücklich dazu: Dass man zwei freie Vektoren (oder einen Ortsvektor und einen freien) etwa in einer Skizze graphisch so darstellt, dass sie einander schneiden, dies hat *überhaupt keine Bedeutung* - man hätte sie nämlich ebensogut durch einander nicht schneidende Pfeile repräsentieren können, weil ein freier Vektor sich eben nicht damit ändert, dass man seinen darstellenden Pfeil parallel verschiebt.

2. Koordinatensysteme und Koordinatendarstellungen

2.1. Koordinatensysteme für E^2 und E^3 . Im Raum E^3 legt man ein Koordinatensystem K fest durch einen Ursprungspunkt O (Koordinatenursprung) und drei Achsen (Geraden), welche durch O gehen. Man nennt sie gern x -, y - und z - Achse. Die Achsen dürfen nicht in einer Ebene liegen. Ferner legt man auf den Achsen eine positive Richtung fest sowie je eine Einheit. Man kann auch sagen, dass mit einem Dreibein von Einheitsvektoren (in positiver Achsenrichtung zeigend) ein solches System vorgegeben ist. Speziell nennt man ein solches System **kartesisch**, wenn diese Einheitsvektoren gleich lang sind und rechte Winkel miteinander bilden. Solche Systeme kennen Sie für E^2 und haben sie in der Schule benutzt. Ein Koordinatensystem für E^2 benötigt natürlich nur ein Zweibein von Einheitsvektoren, die in verschiedene Richtungen zeigen.

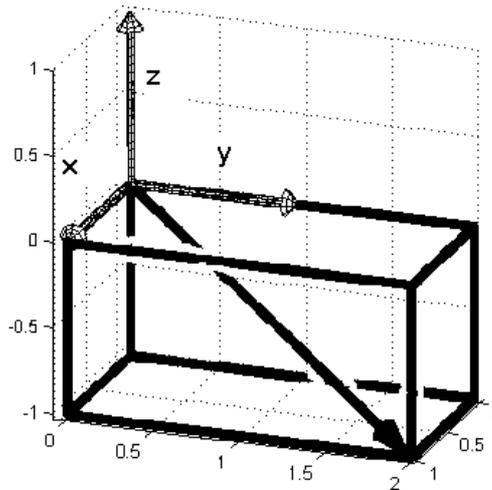
2.2. Koordinatendarstellungen bezüglich K für Punkte, Ortsvektoren und freie Vektoren. Geben wir ein Koordinatensystem K für E^3 vor. Dann gehört zu jedem Punkt P im E^3 genau ein *Zahlentripel* (geordnete Folge von drei Zahlen), das wir $\vec{x}_P^K = (x_P^K, y_P^K, z_P^K)$ nennen:

Man projiziert den Punkt P parallel zur z - Achse auf die xy - Ebene. Nun projiziert man den resultierenden Punkt auf der xy -Ebene parallel zur y - Achse auf die x - Achse und erhält einen Punkt auf der x - Achse, der die x -Koordinate x_P^K ergibt. Vertauschen der Achsen ergibt analog y_P^K und z_P^K . Umgekehrt gelangt man von jedem Zahlentripel bezüglich eines festen Koordinatensystems eindeutig zu einem Punkt.

Da nach der Wahl eines Ursprungs jedem Punkt umkehrbar eindeutig ein Ortsvektor entspricht, liegt es nahe, dem Ortsvektor \vec{x}_P dieselbe Koordinatendarstellung wie dem Punkt P zu geben, also das Tripel (x_P^K, y_P^K, z_P^K) auch als Koordinatendarstellung des Ortsvektors \vec{x}_P in K zu lesen.

Schließlich erhalten freie Vektoren folgende Koordinatendarstellung: Ein freier Vektor $\vec{x} \in V^3$ hat einen eindeutig bestimmten Pfeil-Repräsentanten, der im Koordinatenursprung beginnt. Dieser Repräsentant ist ein Ortsvektor, und \vec{x} erhält dessen bereits definierte Koordinatendarstellung, die wir konsequent mit \vec{x}^K bezeichnen.

Folgende Grafik illustriert dies:



Man sieht das (hier kartesische) Koordinatensystem mit dem Dreibein, das die Achsen x, y, z mit Richtungen und Einheiten vorgibt. Der eingezeichnete Vektor kann als Ortsvektor \vec{x}_P des Punktes P an der Pfeilspitze betrachtet werden oder auch als Repräsentant eines freien Vektors. Man liest problemlos die Koordinatendarstellung ab: $\vec{x}_P^K = (1, 2, -1)$. Ferner ist der Quader zu P bezüglich K eingezeichnet: Dessen Seitenflächen liegen parallel zu den Koordinatenebenen, und die Strecke \overline{OP} bildet eine Diagonale. (Bei schiefwinkligem Koordinatensystem hätte man ein Spat oder Parallelepiped anstelle des Quaders.)

Hinweis: Man verwechsle niemals einen Vektor mit seiner Koordinatendarstellung in einem System - in einem anderen System ergäbe sich eine andere Darstellung. Wohl aber werden wir sehen, dass man mit Zahlenpaaren und Zahlentripeln wie mit Vektoren rechnen kann und dass man sie geometrisch deuten kann.

Wir bezeichnen die Menge der Paare reeller Zahlen mit \mathbb{R}^2 , die Menge der Tripel reeller Zahlen mit \mathbb{R}^3 . Lies: „R zwei, R drei“. Man verwechsle nicht $\{1, 2, 2\} = \{1, 2\}$ mit $(1, 2, 2)$: Im letzteren Zahlentripel kommt es auf die Reihenfolge an, und man kann kein Glied fortlassen. In der ersteren Menge werden dagegen einfach die Elemente aufgezählt, die Reihenfolge ist beliebig, und Überflüssiges kann man weglassen. *Hinweis: Wir schreiben Zahlentripel systemantisch als Spalten, nicht als Zeilen. Letztere werden hier nur aus Platzgründen verwendet.*

3. Punktmengen und deren Quantifizierung als Koordinatentripelmengen

Geometrische Gebilde können wir nunmehr als Mengen von Zahlenpaaren bzw. Zahlentripeln quantifizieren und umgekehrt diese geometrisch verstehen. Dabei lernen wir die beiden grundlegenden Darstellungsformen geometrischer Gebilde kennen.

Beginnen wir mit ein paar sehr einfachen Beispielen:

Legen wir ein kartesisches System K in der Ebene E^2 fest und betrachten die Gerade im Winkel von 45 Grad zu den Achsen, welche durch den ersten Quadranten geht. Diese Gerade ist die Menge aller Punkte, deren Koordinatendarstellungen (x, y) bezüglich K folgende Bestimmungsgleichung lösen: $y = x$. Zur Geraden gehört also die Zahlenpaarmenge

$$\{(x, y) \in \mathbb{R}^2 \mid y = x\} = \{(x, x) \mid x \in \mathbb{R}\}.$$

Wir können auch t (für x) als freien Parameter alle reellen Zahlen durchlaufen lassen und mit $\vec{x}(t) = (t, t)$, $t \in \mathbb{R}$, die Lösungsmenge der Gleichung durchlaufen lassen. Das nennt man Parameterdarstellung. Die zugehörigen Ortsvektoren durchlaufen mit ihren Pfeilspitzen dann die gesamte Gerade. Die weitere Gerade, die Sie traditionell mit $y = 2x + 3$ beschrieben haben, hätte analog die Parameterdarstellung

$$\vec{x}(t) = (t, 2t + 3), \quad t \in \mathbb{R}.$$

Wir beobachten: Die Gleichung $y = 2x + 3$ ist linear in den beiden Unbestimmten, und die Komponentenfunktionen in der Parameterdarstellung sind linear im Parameter t : Die Ausdrücke lauten t , $2t + 3$. Wir erwarten, dass auch

$$\vec{y}(t) = (3t - 1, 2t + 3), \quad t \in \mathbb{R}$$

eine Parameterdarstellung einer Geraden ist, weil die Komponentenfunktionen linear in t sind. Tatsächlich ist $2t + 3 = ((3t - 1) + 1)/3 * 2 + 3$, also lautet die zugehörige traditionelle Geradengleichung

$$y = 2 \frac{x + 1}{3} + 3 = \frac{2}{3}x + \frac{11}{3}.$$

Begrenzen wir den Wertebereich des freien Parameters t , etwa $0 \leq t \leq 1$, so erhalten wir eine Strecke, in diesem Falle einschließlich der Endpunkte.

Weiter beobachten wir an diesen Beispielen: Eine Bestimmungsgleichung führt im zweidimensionalen Raum auf ein eindimensionales Gebilde, und diese Eindimensionalität drückt sich in einem freien Parameter bei einer Parameterdarstellung aus. Wir erwarten daher, dass

$$\vec{z}(t) = (2t + 3, t - 1, 3t + 4), \quad t \in \mathbb{R}$$

ebenfalls eine Gerade parametrisiert, diesmal im dreidimensionalen Raum: Ein freier Parameter ohne Beschränkung, und alle Komponentenfunktionen linear im freien Parameter. Solche Ausdrücke werden wir nach Einführung der linearen Vektorraumoperationen geometrisch besser verstehen lernen. Wie sähe es mit einer Darstellung derselben Geraden durch Bestimmungsgleichungen aus? Eine Gleichung wird es nicht tun, da wir im dreidimensionalen Raum erwarten, dass eine einzige Gleichung erst ein zweidimensionales Gebilde aussondert. Beispiel: $z = 0$ bestimmt als Lösungsmenge die gesamte xy -Ebene. Aber wir können wie oben im Ausdruck $\vec{z}(t) = (2t + 3, t - 1, 3t + 4)$ die zweite Komponente durch die erste, dann die dritte durch die erste ausdrücken und erhalten so ein System von zwei linearen Bestimmungsgleichungen. Nach unserer bereits eingeführten Systematik sollten wir erwarten, dass jede einzelne Gleichung eine Ebene bestimmt (da sie linear ist und die Lösungsmenge im \mathbb{R}^3 gesucht wird). Und der Schnitt zweier Ebenen bildet eine Gerade, sofern diese Ebenen nicht parallel liegen. (Solche Ebenengleichungen werden wir später genauer geometrisch verstehen nach Einführung des Skalarproduktes.) Geben wir aber noch ein einfaches Beispiel im Dreidimensionalen, bei dem die Verhältnisse sofort zu übersehen sind: Das Gleichungssystem

$$\begin{aligned} y &= x \\ z &= x \end{aligned}$$

bedeutet geometrisch: Die Ebene (nicht etwa Gerade - wir suchen Lösungen im \mathbb{R}^3 !) $y = x$ besteht aus allen Punkten mit beliebiger z -Koordinate und gleichen x - und y -Koordinaten. Analog versteht sich die Ebene $z = x$. Der Schnitt dieser beiden Ebenen ist die Gerade mit der Parameterdarstellung $\vec{u}(t) = (t, t, t)$, $t \in \mathbb{R}$.

Geben wir aber noch eine Parameterdarstellung für die Ebene $y = x$ - wieder Lösungsmenge im \mathbb{R}^3 : Sie lautet offenbar am einfachsten:

$$\vec{v}(x, z) = (x, x, z), \quad x, z \in \mathbb{R}.$$

Es sind also zwei freie Parameter nötig entsprechend der Dimension 2 der Ebene.

Bereits bei der verbalen Beschreibung einer gemeinten geometrischen Figur muss man recht sorgfältig vorgehen: Betrachten wir einen Kreis im E^2 um den Mittelpunkt P vom Radius r , so könnten wir genauer die gesamte Kreisfläche meinen (mit oder ohne Rand) oder aber gerade den Rand. Sehen wir letzteren an: Das geometrische Gebilde ist die Menge aller Punkte dieser Kurvenbahn. Wir führen ein möglichst einfaches Koordinatensystem K ein, um sie zu beschreiben: Als Ursprung wählen wir $O = P$, dazu ein kartesisches System. Dann haben wir als zugehörige Menge von Zahlenpaaren:

$$\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\}.$$

Hätten wir die Einheit r gewählt, so wäre die einfachere Gleichung $x^2 + y^2 = 1$ herausgekommen. Umgekehrt können wir diese Menge von reinen Zahlenpaaren als Kreisrand geometrisch interpretieren, wenn wir ein entsprechendes Koordinatensystem wählen, aber auch als Ellipse, wenn wir die Einheitslängen auf beiden Achsen verschieden setzen. Wählen wir nicht P als Ursprung unseres wiederum kartesischen Systems K , sondern einen beliebigen Punkt O , so nimmt die zum Kreisrand gehörige Menge von Zahlenpaaren folgende Form an:

$$\{(x, y) \in \mathbb{R}^2 \mid (x - x_P)^2 + (y - y_P)^2 = r^2\}.$$

Dabei ist $\vec{x}_P^K = (x_P, y_P)$ die Koordinatendarstellung von P in K . Diese Beschreibung stimmt auch speziell für $O = P$.

Wir haben damit einen Kreisrand (genauer die Menge der zugehörigen Koordinatendarstellungen) als Lösungsmenge einer Bestimmungsgleichung wie $x^2 + y^2 = r^2$ dargestellt. Wenn wir jedoch beispielsweise eine solche Kurve auf dem Computerbildschirm zeichnen wollen, so ist das unpraktisch. Viel handlicher dafür ist die zweite grundlegende Darstellungsform: Parameterdarstellung. Dazu müssen wir nur wissen, dass die Lösungen der Gleichung $x^2 + y^2 = r^2$ sämtlich die Form $(r \cos(\varphi), r \sin(\varphi))$ haben mit $0 \leq \varphi < 2\pi$. Dann können wir die Lösungsmenge so umschreiben:

$$\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\} = \{(r \cos(\varphi), r \sin(\varphi)) \mid 0 \leq \varphi < 2\pi\}.$$

Was bringt das? Wir können etwa die Werte $0, 0.01, 0.02, \dots$ für φ in den Ausdruck $(r \cos(\varphi), r \sin(\varphi))$ einsetzen und erhalten damit lauter Koordinatenpaare von Kreispunkten, die wir dann zeichnen können. Gerade so funktioniert das Computer-Zeichnen. Letztere Darstellung schreibt man gewöhnlich nicht als Menge sondern so:

$$\vec{x}(\varphi) = (r \cos(\varphi), r \sin(\varphi)), \quad 0 \leq \varphi < 2\pi.$$

Hier ist φ ein freier Parameter. Durchläuft er die angegebene Menge von Werten, so durchläuft $\vec{x}(\varphi)$ die gewünschte Menge von Zahlenpaaren. Selbstverständlich ist es auch möglich, φ als unabhängige und $\vec{x}(\varphi)$ als abhängige Variable zu betrachten, was für weitergehende Fragen nützlich ist. Wollen wir analog einen Kreis mit Radius r und Mittelpunkt P beschreiben, so ergibt sich:

$$\vec{y}(\varphi) = (x_P + r \cos(\varphi), y_P + r \sin(\varphi)), \quad 0 \leq \varphi < 2\pi,$$

wobei (x_P, y_P) die Koordinatendarstellung von P ist.

Betrachten wir als nächstes Beispiel die gesamte Kreisfläche mit Rand, wieder mit Mittelpunkt im Koordinatenursprung. Analog zur ersten Darstellung hätten wir eine Ungleichung zu bilden anstelle der Gleichung:

$$\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r^2\}.$$

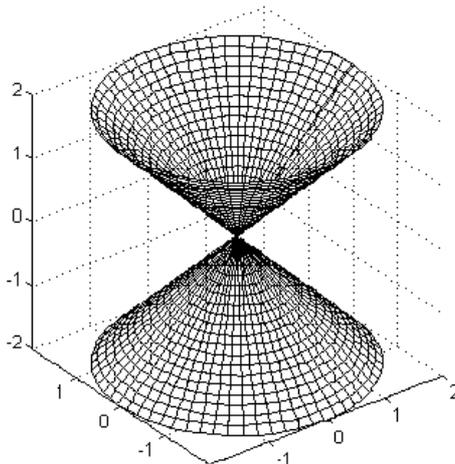
Wir denken uns die Kreisfläche aus allen Kreiskurven mit Radius $\leq r$ zusammengesetzt. Wie sähe eine Parameterdarstellung dafür aus? Zusätzlich zum Winkel hätten wir den Radius im Bereich von 0 bis r variieren zu

lassen, also etwa

$$\vec{x}(\varphi, s) = (s \cos(\varphi), s \sin(\varphi)), \quad 0 \leq \varphi < 2\pi, \quad 0 \leq s \leq r.$$

Zusammengefasst beobachten wir: Nichtlinearen Bestimmungsgleichungen entsprechen nichtlineare („krumme“) geometrische Gebilde. In zugehörigen Parameterdarstellungen treten nichtlineare Komponentenfunktionen auf. Wieder gilt die oben angesprochene Zählweise für die Dimension.

Ein weiteres Beispiel: Unbegrenzter Doppelkegelmantel im E^3 . Betrachten wir ein kartesisches System und dafür die Koordinaten-Bestimmungsgleichung $z^2 = x^2 + y^2$. Nach unseren allgemeinen Beobachtungen erwarten wir, dass damit eine krumme Fläche beschrieben wird. (Die Gleichung ist nicht linear, und die Dimension 3 des Grundraums sollte sich um 1 erniedrigen durch eine Gleichung.) Ein solches Gebilde kann man gut verstehen, indem man einfache Schnitte bildet: Schnitt mit der Ebene $z = 4$ ergibt $x^2 + y^2 = 4$ und $z = 4$, also den Kreis aller Punkte $(2 \cos(\varphi), 2 \sin(\varphi), 4)$, $0 \leq \varphi < 2\pi$. Und so für alle Ebenen parallel zur xy -Ebene: Stets ergeben sich Kreise mit Mittelpunkten auf der z -Achse. Speziell für $z = 0$ nur der Ursprung, und die Kreise werden größer mit dem Abstand von der xy -Ebene. Wir brauchen nur noch durch Schnittbildung mit der yz -Ebene einzusehen, in welchem Maße sie größer werden: Schneiden mit der yz -Ebene bedeutet Nullsetzen von x , also $z^2 = y^2$. Lösungen sind die Geraden $z = y$, $z = -y$ auf der yz -Ebene. Diese steigen in 45 Grad-Winkeln zur xy -Ebene an. Also haben wir den versprochenen unbegrenzten Doppelkegelmantel, genauer Kreiskegel mit Spitze im Ursprung und der z -Achse als Achse des Kegels. Folgendes Bild zeigt einen begrenzten Abschnitt davon:



3.1. Bestimmungsgleichungen und Parameterdarstellungen. Wir geben einen zusammenfassenden Überblick über die Beobachtungen an den Beispielen: Man beachte, dass man ein zweidimensionales Gebilde eine Fläche nennt (und etwa vom Flächeninhalt unterscheidet). Ein eindimensionales heißt Kurve oder genauer Bahn einer Kurve. Ebenen sind spezielle Flächen, und Geraden sind spezielle Kurven. Punkte sind nulldimensional, ebenso endliche Mengen von Punkten und sogar solche unendlichen Mengen, welche kein Kontinuum enthalten. Für höhere Dimensionen ist das Wort „Körper“ üblich. Bei den folgenden Angaben muss man ein wenig aufpassen: Als Anzahl der Gleichungen ist die Zahl der unabhängigen Gleichungen gemeint, was wir im linearen Falle näher verstehen werden. Entsprechend ist mit der Anzahl der freien Parameter die notwendige Zahl unabhängiger Parameter gemeint. Schließlich wiederholen wir noch einmal die Möglichkeit der Einschränkungen für Parameterbereiche. Damit kann man etwa aus einer Geraden eine Halbgerade oder Strecke aussondern, aus

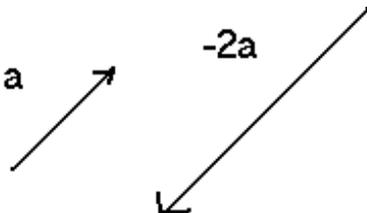
einer Ebene eine Halbebene oder ein Parallelogramm usw.

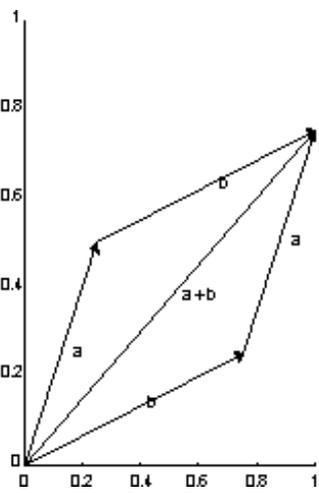
Darstellungsform:	Art des geometrischen Gebildes:		
	linear (gerade)	nichtlinear (krumm)	Dimension
Bestimmungsgleichung(en)	linear in Unbest.	nichtlinear in Unbest.	Raumdimension minus Anzahl der Gleichungen
Parameterdarstellung	lineare Komponentenfunktionen in den freien Parametern	nichtlineare Komponentenfunktionen in den freien Parametern	Anzahl der freien Parameter

4. Die linearen Vektorraumoperationen

Wichtig ist es immer wieder, aus einfacheren Dingen kompliziertere zusammenzusetzen. Gerade dazu leisten die linearen Operationen Beträchtliches. Man stellt etwa zwei Kräfte durch Vektoren dar: Die resultierende Kraft ergibt sich durch Addition der Vektoren. Analog für Geschwindigkeiten, einander überlagernde elektrische Feldstärken, usw. Besonders nützlich ist es, dass man eine geometrische Version der Vektoroperationen sowie eine rechnerische auf der Seite der Zahlentripel hat. (Wir erwähnen nicht jedes Mal eigens die Zahlenpaare.) Vor allem kann man die Vektorraumoperationen dazu benutzen, aus einfachen Elementen eine kompliziertere Situation aufzubauen, etwa aus dem Ortsvektor eines Kreispunktes den Ortsvektor eines Punktes eines weiteren Kreises gewinnen, der auf ersterem abrollt. Oder man möchte von einem Punkt aus zum neuen Punkt, der aus ersterem durch eine Projektion oder eine Spiegelung entsteht. Später werden Sie noch lernen, dass auch Funktionen Vektorräume bilden und dass der Addition von Funktionen überragende Bedeutung zukommt.

4.1. Vektoraddition und Multiplikation eines Vektors mit einer Zahl. Das sind die sogenannten linearen Vektorraumoperationen. Wir stellen sie in beiderlei Form gegenüber:

Multiplikation eines Vektors mit einer Zahl (einem Skalar)	
geometrisch ($\lambda \in \mathbb{R}$, $\vec{a} \in V^3$)	rechnerisch ($\lambda \in \mathbb{R}$, $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$)
$\lambda \vec{a}$ entsteht aus \vec{a} durch Streckung mit λ , wobei $\lambda < 0$ zusätzlich ein Umdrehen der Pfeilspitze bewirkt. Zum Beispiel:	$\lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \\ \lambda z \end{pmatrix}$, z.B.:
	$-2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \\ 0 \end{pmatrix}$

Addition zweier Vektoren:	
geometrisch ($\vec{a}, \vec{b} \in V^3$)	rechnerisch $\left(\begin{pmatrix} u \\ v \\ w \end{pmatrix}, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \right)$
$\vec{a} + \vec{b}$ entsteht durch Hintereinandersetzen bzw. Parallelogramm-ergänzung. Zum Beispiel:	$\begin{pmatrix} u \\ v \\ w \end{pmatrix} + \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} u+x \\ v+y \\ w+z \end{pmatrix}$, z.B.:
	$\begin{pmatrix} 1/4 \\ 1/2 \\ 0 \end{pmatrix} + \begin{pmatrix} 3/4 \\ 1/4 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3/4 \\ 0 \end{pmatrix}$

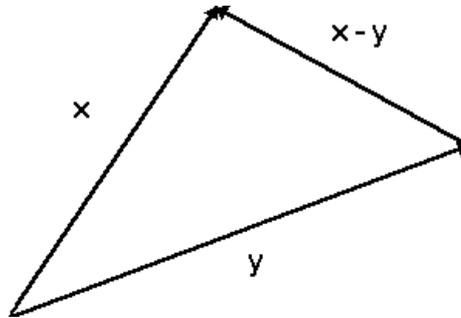
Entscheidend ist nun der Zusammenhang zwischen den rechnerischen Operationen einerseits und den geometrischen andererseits: Man beobachtet sogleich im Beispiel, wie sich bei geometrischer Addition auch die Koordinaten addieren. Das ist generell so. Für die Multiplikation mit einem Skalar wird der Zusammenhang vom Strahlensatz gesichert. Erst diese Sachverhalte garantieren, dass man etwa auch Kräfte richtig rechnerisch addiert und damit die Koordinatendarstellung der tatsächlichen resultierenden Kraft erhält, usw. Wir fassen sie daher noch einmal ausdrücklich zu Formeln zusammen und unterscheiden dazu an dieser Stelle notwendigerweise auch einmal ausdrücklich die rechnerische von der geometrischen Operation durch Anfügen eines Index (r bzw. g). Zusätzlich verwenden wir einen Multiplikationspunkt, den man sonst gewöhnlich weglässt.

$$\begin{aligned} (\lambda \cdot_g \vec{x})^K &= \lambda \cdot_r (\vec{x}^K) \\ (\vec{x} +_g \vec{y})^K &= \vec{x}^K +_r \vec{y}^K \end{aligned}$$

Diese allgemeingültigen Formeln besagen gerade, dass die Rechenoperationen den geometrisch-physikalischen wirklich entsprechen.

Warnung: Es gibt noch eine andere „Operation“, die dem Anfänger vielfach naheliegt: Das komponentenweise Multiplizieren zweier Vektoren zu einem neuen Vektor. Dieser Vektor ist zwar dann aufgrund einer wohldefinierten Operation klar definiert, nur ist er zu nichts Gutem zu gebrauchen. Damit wird stets Unsinn produziert. Gefährlich ist vor allem, das später einzuführende Skalarprodukt damit zu verwechseln.

Recht wichtig dagegen ist es, die Differenz zweier Vektoren zu bilden und eine unmittelbare Anschauung dafür zu gewinnen, obwohl man einfach lapidar $\vec{x} - \vec{y} = \vec{x} + (-\vec{y}) = \vec{x} + (-1)\vec{y}$ definieren kann und diese Operation daher nichts eigentlich Neues ist. Hier die Illustration:



Zunächst mache man sich am Bild einfach klar, dass $\vec{x} - \vec{y}$ allein schon deswegen so aussehen sollte, weil sicher gelten sollte: $\vec{y} + (\vec{x} - \vec{y}) = \vec{x}$. Die Differenz zweier Vektoren bildet man also, indem man ihre Anfänge aneinanderlegt und den zwischen die Pfeilspitzen eingespannten Vektor bildet. Man hat nur zu beachten, dass die Spitze des Differenzvektors an die des Vektors stößt, von dem abgezogen wird. Das bedeutet insbesondere, dass man mit der Differenz der Ortsvektoren $\vec{x}_Q - \vec{x}_P$ den Vektor von P nach Q bekommt, den man auch mit \vec{PQ} bezeichnet. Man achte auf die Reihenfolgen!

4.2. Die abstrakten Rechenregeln für das Rechnen mit Vektoren (Vektorraumaxiome). In Worten lässt sich knapp zusammenfassen: Mit Vektoren rechnet man wie mit Zahlen, einzig und allein das Dividieren durch einen Vektor entfällt (und ist auch nicht sinnvoll im allgemeinen Fall zu definieren!). Man kann jedoch auch einen vollständig ausreichenden Satz von Regeln (alle weiteren kann man daraus herleiten) knapp aufschreiben, hier ist er. Zu verstehen sind diese Formeln als allgemeingültige, d.h. man darf für λ, μ jeden Skalar(ausdruck) und für $\vec{x}, \vec{y}, \vec{z}$ jeden Vektor(ausdruck) einsetzen, erhält stets wieder allgemeingültige Formeln. Der Bereich der Zahlen muss nicht \mathbb{R} sein, es kann jeder Körper sein, der Fall der komplexen Zahlen ist besonders wichtig. Man spricht dann allgemein von einem Vektorraum über einem Körper K . Nun nennt man definitionsgemäß jede Struktur mit einer inneren Verknüpfung und einer äußeren mit den Elementen eines Körpers K genau dann einen Vektorraum über K , wenn diese Regeln erfüllt sind. Was hat man davon? Man verifiziert diese Regeln und weiß dann, dass auch alle Folgerungen gelten und man also „wie üblich“ bequem und korrekt rechnen kann.

DEFINITION 2. Eine Menge V mit einer zweistelligen Verknüpfung $+$ zusammen mit einem Körper (vgl. die Definition dafür durch die Körperaxiome in Kap. 1) K und einer äußeren Verknüpfung \cdot , welche jedem Paar (λ, \vec{x}) , $\lambda \in K$, $\vec{x} \in V$, genau ein Element $\lambda \cdot \vec{x} \in V$ zuordnet, heißt Vektorraum über K , wenn folgende

Rechengesetze (die Vektorraumaxiome) allgemein erfüllt sind:

$$\begin{aligned}
 \vec{x} + (\vec{y} + \vec{z}) &= (\vec{x} + \vec{y}) + \vec{z} \text{ (Assoziativität der Addition)} \\
 \vec{0} + \vec{x} &= \vec{x} \text{ (neutrales Element, Nullvektor)} \\
 -\vec{x} + \vec{x} &= \vec{0} \text{ (inverse Elemente bzgl. Addition)} \\
 \vec{x} + \vec{y} &= \vec{y} + \vec{x} \text{ (Kommutativität)} \\
 (\lambda\mu)\vec{x} &= \lambda(\mu\vec{x}) \text{ (Klammern überflüssig)} \\
 (\lambda + \mu)\vec{x} &= \lambda\vec{x} + \mu\vec{x} \text{ (erstes Distributivgesetz)} \\
 \lambda(\vec{x} + \vec{y}) &= \lambda\vec{x} + \lambda\vec{y} \text{ (zweites Distributivgesetz)} \\
 1 \cdot \vec{x} &= \vec{x} \text{ (Normierung der Mult. mit Zahl)}
 \end{aligned}$$

Die Elemente von V nennt man dann Vektoren, die von K Skalare (für: Zahlen).

Man sieht die zwei Gruppen: Erstere handelt nur von $+$, letztere auch von der Multiplikation mit einem Skalar, insbesondere von der Verbindung beider in den Distributivgesetzen. Ersetzen Sie einmal die Vektoren durch Zahlen, so sehen Sie Ihnen längst bekannte Rechengesetze.

Eigentlich müssten wir jetzt für unsere eingeführten Strukturen V^3 (analog V_O^3) sowie \mathbb{R}^3 zeigen, dass diese Axiome alle erfüllt sind. Damit erst ist klar, dass wir mit Recht von Vektoren gesprochen haben. Aber man beachte, dass diese Nachprüfung etwa für \mathbb{R}^3 sehr selbstverständlich ist, da sich jedes Axiom auf Entsprechendes innerhalb der Zahlen sofort zurückführt. Ein Beispiel:

$$\begin{aligned}
 \lambda \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} u \\ v \\ w \end{pmatrix} \right) &= \lambda \begin{pmatrix} x+u \\ y+v \\ z+w \end{pmatrix} = \begin{pmatrix} \lambda(x+u) \\ \lambda(y+v) \\ \lambda(z+w) \end{pmatrix} \\
 &= \begin{pmatrix} \lambda x + \lambda u \\ \lambda y + \lambda v \\ \lambda z + \lambda w \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \\ \lambda z \end{pmatrix} + \begin{pmatrix} \lambda u \\ \lambda v \\ \lambda w \end{pmatrix} \\
 &= \lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \lambda \begin{pmatrix} u \\ v \\ w \end{pmatrix}.
 \end{aligned}$$

Dabei wurden nur die Definitionen der linearen Operationen im \mathbb{R}^3 sowie das Distributivgesetz der reellen Zahlen verwandt. Analog funktioniert die Sache mit den andern Gesetzen. Wir erwähnen nur, dass wir im \mathbb{R}^3 haben:

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad - \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -x \\ -y \\ -z \end{pmatrix}.$$

Wenn wir nunmehr die Rechengesetze für den Vektorraum \mathbb{R}^3 verifiziert haben, so können wir sofort schließen, dass sie auch für V^3 gelten: Denn die obenstehenden Formeln der Entsprechung zwischen rechnerischer und geometrischer Operation machen klar, dass man über K in gewisser Weise V^3 mit \mathbb{R}^3 identifizieren kann. (Entsprechen sich Elemente umkehrbar eindeutig und ebenfalls die Operationen, dann spricht man von isomorphen Strukturen, und man kann ganz allgemein zeigen, dass dann stets dieselben Rechengesetze gelten.) Dasselbe Argument gilt dann auch für V_O^3 . Andererseits ist es interessant, zu beobachten, dass ein Gesetz wie $\lambda(\vec{x} + \vec{y}) = \lambda\vec{x} + \lambda\vec{y}$ geometrisch immerhin so viel wie den Strahlensatz bedeutet, das heisst: Wendet man dies Rechengesetz einfach formal an, so hat man eine sonst im allgemeinen viel kompliziertere geometrische Überlegung gespart. Insgesamt werden wir sehen: Kompliziertere geometrische Probleme lassen sich mit Vektorrechnung viel leichter lösen als ohne sie. Halten wir noch einmal unser Hauptresultat fest:

SATZ 1. V^3, V_O^3, \mathbb{R}^3 sind paarweise isomorphe Vektorräume über dem Körper der reellen Zahlen. Das entsprechende Resultat gilt für Dimension 2. Analog hat man auch die Räume \mathbb{R}^n der sogenannten „ n -Tupel“ (also Folgen der Länge n , $n \in \mathbb{N}$, $n \geq 1$) von reellen Zahlen, in denen ebenfalls komponentenweise gerechnet wird.

Allerdings fehlen für $n \geq 4$ die zugehörigen anschaulichen Räume, wenn auch nicht der unmittelbare Bezug zur Physik (Raumzeit, aber auch mechanische Systeme mit riesigen Freiheitsgraden usw.)

4.3. Einige praktisch wichtige Folgerungen. Folgende Regeln sind herleitbar aus den Axiomen, man hat sie aber zweckmäßig im Kopf und benutzt sie ohne weiteres:

$$\begin{aligned} 0 \cdot \vec{x} &= \vec{0} \\ -(\lambda \vec{x}) &= (-\lambda) \vec{x} \\ \lambda(\vec{x} - \vec{y}) &= \lambda \vec{x} - \lambda \vec{y} \end{aligned}$$

Man beachte: Die zweite Gleichung ist nicht von vornherein selbstverständlich, sie besagt, dass der additiv inverse (negative) Vektor zu $\lambda \vec{x}$ dasselbe ist wie der Vektor \vec{x} mit $(-\lambda)$ multipliziert (Inversenbildung im Körper der reellen Zahlen).

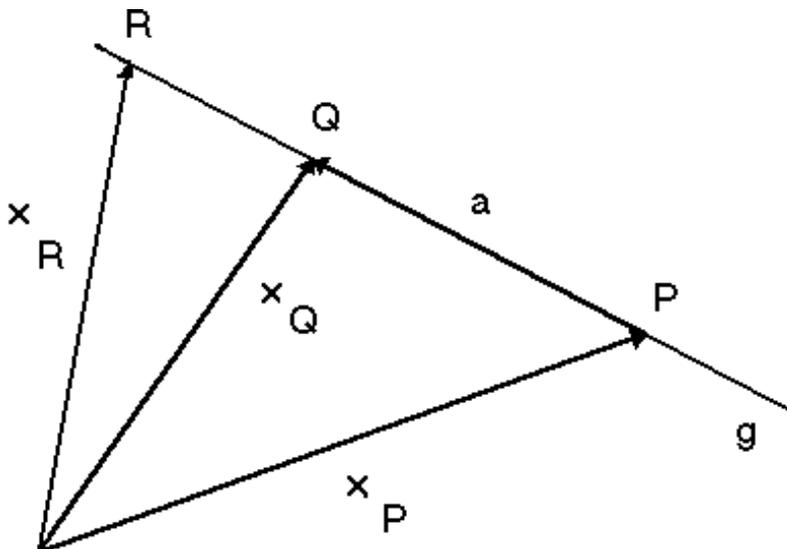
Eine immer wieder nützliche Anwendung der Rechenregeln ist diese Folgerung: Jeder Rechenausdruck der Vektorraumstruktur in den Vektoren $\vec{a}_1, \dots, \vec{a}_n$ lässt sich auf die Form $\sum_{i=1}^n \lambda_i \vec{a}_i = \lambda_1 \vec{a}_1 + \dots + \lambda_n \vec{a}_n$ bringen. Das ist dann auch die ordentliche Endform eines solchen Ausdrucks, die man für viele Zwecke benötigt und stets produzieren sollte. Man nennt diesen Ausdruck eine **Linearkombination der Vektoren** $\vec{a}_1, \dots, \vec{a}_n$. Beispiel:

$$2(\vec{x} + \lambda \vec{y}) - 3\vec{x} + \mu \frac{\vec{y}}{3} = -\vec{x} + \left(2\lambda + \frac{\mu}{3}\right) \vec{y}.$$

Hier wären also $n = 2$, $\vec{a}_1 = \vec{x}$, $\vec{a}_2 = \vec{y}$, $\lambda_1 = -1$, $\lambda_2 = 2\lambda + \frac{\mu}{3}$. Beachten Sie, wie man hierzu insbesondere die beiden Distributivgesetze benötigt (und zwar jeweils in beiden Richtungen gelesen, das eine Mal Ausmultiplizieren, das andere Mal Ausklammern!).

4.4. Parameterdarstellungen für Geraden und Ebenen. Mittels der Vektorrechnung gelingt es leicht, die oben bereits angesprochene Parameterdarstellung für Geraden und Ebenen geometrisch zu verstehen.

Sei g eine Gerade im E^3 . Dann kann man g festlegen durch einen Punkt und eine Richtung. Die Richtung von g kann wiederum durch einen freien Vektor $\vec{a} \neq \vec{0}$ parallel zu g bestimmt werden. Sei also P ein fester Punkt auf g und \vec{a} ein solcher Richtungsvektor für g . Betrachten wir folgende Skizze:



Daraus wird klar: Wenn man $\vec{x}_P + \lambda \vec{a}$ bildet mit dem Ortsvektor \vec{x}_P von P und einer beliebigen Zahl $\lambda \in \mathbb{R}$, so erhält man stets einen Ortsvektor eines Punktes der Geraden. (Beachte Konvention: Ortsvektor plus freier (Richtungs-) Vektor gleich Ortsvektor!) Zum Beispiel ist in der Skizze $\vec{x}_P + 1 \cdot \vec{a} = \vec{x}_Q$ und etwa $\vec{x}_P + 1 \cdot \vec{a} = \vec{x}_R$. Umgekehrt erhält man zu jedem Punkt S der Geraden eine eindeutig bestimmte Zahl $\lambda(S)$, so dass $\vec{x}_S = \vec{x}_P + \lambda(S) \vec{a}$ ist. Die Menge aller Ortsvektoren von Geradenpunkten wird also durchlaufen mit

$$\begin{aligned} \vec{x}_g(\lambda) &= \vec{x}_P + \lambda \vec{a}, \quad \lambda \in \mathbb{R}. \quad (\lambda: \text{Freier Parameter}), \\ &\quad \text{(Parameterdarstellung einer Geraden } g, \text{ geometrische Form.)} \\ \vec{x}_g^K(\lambda) &= \vec{x}_P^K + \lambda \vec{a}^K, \quad \lambda \in \mathbb{R}. \quad \text{(Koordinatenform.)} \\ \text{Dabei sind} &: \quad \vec{x}_P \text{ Aufpunktvektor,} \\ &\quad \vec{a} \neq \vec{0} \text{ Richtungsvektor.} \end{aligned}$$

Man beachte hier die beiden Formen. $\vec{x}_g(\lambda)$ ist der geometrische Pfeil „Ortsvektor des Geradenpunktes zum Parameterwert von λ “. Er ergibt sich durch die geometrische Vektoroperationen als $\vec{x}_P + \lambda \vec{a}$. Diese Form hatten wir für unsere allgemeine anschaulich-geometrische Einsicht benutzt. Dagegen werden wir für Fragen wie nach dem Schnittpunkt einer Geraden mit der xy -Ebene usw. die Koordinatenform benutzen. Dabei rechnet man rein formal mit Zahlentripeln.

Ein Beispiel dazu: Geben wir die Gerade g mit folgender Parameterdarstellung in Koordinatenform vor:

$$\vec{x}_g^K(\lambda) = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} -3 \\ 2 \\ 2 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Suchen wir den Schnitt mit der xy -Ebene, so ergibt sich die Bedingung

$$\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} -3 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}.$$

Man beachte: Zwei Vektoren aus \mathbb{R}^3 sind genau dann gleich, wenn sie in allen drei Komponenten übereinstimmen. Unsere Vektorgleichung läuft also auf folgendes System von drei Zahlengleichungen hinaus:

$$\begin{aligned} 2 - 3\lambda &= x \\ -1 + 2\lambda &= y \\ 2 + 2\lambda &= 0 \end{aligned}$$

Das ist ein lineares Gleichungssystem in den Unbestimmten λ, x, y . Aber es ist besonders einfach (zur Systematik mehr im nächsten Abschnitt). Aus der letzten Gleichung liest man sofort $\lambda = -1$ ab, und mit beiden andern Gleichungen sind dann x, y auch eindeutig bestimmt. Wir erhalten also einen einzigen Schnittpunkt, es ist der Geradenpunkt zum Parameterwert $\lambda = -1$. Der gesuchte Schnitt besteht also aus dem einzigen Punkt S mit

$$\vec{x}_S^K = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} + (-1) \begin{pmatrix} -3 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \\ 0 \end{pmatrix}.$$

Man wird in solchen Zusammenhängen später den Index K weglassen - man sieht ja, dass es sich um eine Koordinatendarstellung handelt.

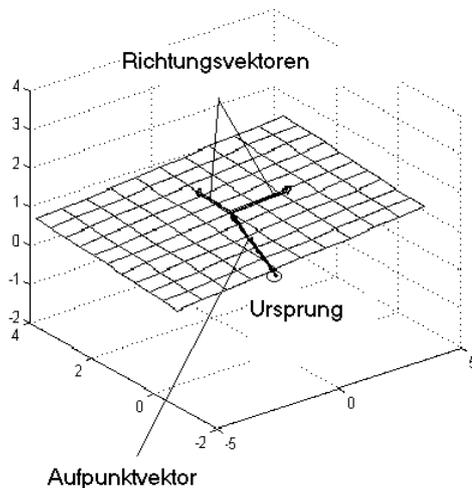
Es sollte deutlich geworden sein, dass die Spaltenschreibweise für die Zahlentripel für solche Rechnungen besonders übersichtlich ist. (Noch einmal: Wenn wir hier Zeilenvektoren schreiben, dann nur aus Gründen der Platzerparnis.)

Mit diesen Mitteln können wir auch sofort einsehen, dass so etwas wie

$$\vec{x}(\lambda) = \begin{pmatrix} 2 - 3\lambda \\ -1 + 2\lambda \\ -3\lambda \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -3 \\ 2 \\ -3 \end{pmatrix}, \lambda \in \mathbb{R},$$

eine Parameterdarstellung einer Geraden ist. Die ausgeführte Umformung macht die Aufspaltung der Standardform deutlich, und wir können nunmehr Aufpunkt- und Richtungsvektor identifizieren und die Gerade uns vorstellen sowie in ein Koordinatensystem einzeichnen.

Nun das entsprechende Programm für Ebenen. Wieder startet die Argumentation mit der geometrischen Vektorrechnung: Geben wir einen Punkt P einer Ebene E im E^3 vor sowie zwei unabhängige (d.h. hier: beide ungleich Null und nicht parallel) Richtungsvektoren \vec{a}, \vec{b} parallel zur Ebene. Folgende Skizze zeigt dann, dass man mit $\vec{x}_P + \lambda \vec{a} + \mu \vec{b}$, $\lambda, \mu \in \mathbb{R}$, stets einen Ortsvektor eines Punktes von E erhält und dass man umgekehrt zu jedem Punkt Q aus E auf diese Weise (sogar eindeutig) seinen Ortsvektor bekommt. Denn man kann stets auf der Ebene ein passendes Parallelogramm mit Seiten parallel zu \vec{a}, \vec{b} ansetzen, so dass eine Ecke in P , die andere im gewünschten Punkt Q landet.



Also hat man ganz analog zu den Geraden folgende Parameterform für eine Ebene mit zwei freien Parametern:

$$\vec{x}_E(\lambda, \mu) = \vec{x}_P + \lambda \vec{a} + \mu \vec{b}, \quad \lambda, \mu \in \mathbb{R}, \quad \vec{a}, \vec{b} \text{ linear unabhängige Richtungsvektoren, } \vec{x}_P \text{ Aufpunktvektor.}$$

Entsprechend hat man wiederum die Koordinatenform mit Zahlentripeln. Zwei Vektoren \vec{a}, \vec{b} sind definitionsgemäß genau dann linear unabhängig, wenn keiner ein Vielfaches des andern ist, also wenn sie beide nicht Null sind und nicht parallel sind, also verschiedene Richtungen angeben. (Dies ist ein Spezialfall des später einzuführenden allgemeinen Begriffes der linearen Unabhängigkeit.) Nun sollte klar sein, dass man eine Parameterdarstellung für eine gewünschte Ebene (analog: Gerade) einfach aufstellt, indem man dies Schema ausfüllt mit einem Aufpunktvektor und einem Richtungsvektor.

Beispiele: Eine Parameterdarstellung ist gesucht für die Ebene E , welche durch P geht, $\vec{x}_P^K = (1, 2, -2)$, und parallel zur yz -Ebene liegt. Offenbar sind $(0, 1, 0)$ und $(0, 0, 1)$ passende Richtungsvektoren, und damit können wir eine Parameterdarstellung für E in Koordinatenform sofort hinschreiben:

$$\vec{x}_E(\lambda, \mu) = (1, 2, -2) + \lambda(0, 1, 0) + \mu(0, 0, 1), \quad \lambda, \mu \in \mathbb{R}.$$

Natürlich gibt es unendlich viele Parameterdarstellungen für *dieselbe* Gerade/Ebene, in unserm Beispiel wäre auch

$$\vec{y}_E(\lambda, \mu) = (1, 3, -2) + \lambda(0, 1, 1) + \mu(0, -1, 1), \quad \lambda, \mu \in \mathbb{R},$$

eine andere Parameterdarstellung derselben Ebene. (Wie könnte man das nachprüfen?)

Weiteres Beispiel: Wir suchen eine Parameterdarstellung der Ebene H , welche durch die drei Punkte P, Q, R geht mit $\vec{x}_P^K = (1, 2, -2)$, $\vec{x}_Q^K = (1, 4, 2)$, $\vec{x}_R^K = (-2, 1, 3)$. Wiederum finden wir leicht einen Aufpunktvektor. Wir wählen den ersten dazu. Richtungsvektoren erkennen wir leicht als Differenzen zwischen den Ortsvektoren. Im Rahmen der Dreiecksituation liegt es nahe, dazu die Vektoren $\vec{PQ} = \vec{x}_Q - \vec{x}_P$ und $\vec{PR} = \vec{x}_R - \vec{x}_P$ zu wählen. Damit erhalten wir folgende Parameterdarstellung für H :

$$\vec{x}_H(\lambda, \mu) = (1, 2, -2) + \lambda(0, 2, 4) + \mu(-3, -1, 5), \quad \lambda, \mu \in \mathbb{R}.$$

Rechentchnik: Im Zusammenhang mit den Parameterdarstellungen von Geraden und Ebenen ist folgende typische Rechnung *in beiden Richtungen* wichtig:

$$\begin{pmatrix} 1 + \lambda - 2\mu \\ 2\lambda + 3\mu \\ 2 + 3\lambda \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix}.$$

Es handelt sich lediglich (von rechts nach links) um die Ausführung der linearen Operationen mit den Zahlentripeln. In der andern Richtung ist es das „Auseinanderziehen“, und das macht man mit dem allenthalben wichtigen Vergleichen von Koeffizienten. Der Plan ist eine Zerlegung wie auf der rechten Seite, also fasst man alle konstanten Glieder zu einem Vektor (Fehlen: Koeffizient Null!), dann alle Koeffizienten bei λ usw. Nunmehr können wir auch die linke Seite, die aus einer Rechnung resultieren könnte, sofort als Ausdruck einer Parameterdarstellung einer Ebene verstehen und mit der rechten Seite diese Ebene geometrisch in ihrer Lage verstehen. Die andere Richtung braucht man etwa beim Aufstellen eines linearen Gleichungssystems, um den Schnitt zweier linearer Gebilde zu berechnen, welche durch Parameterdarstellungen gegeben sind. Schneiden wir g mit E , wobei $\vec{x}_g(\alpha) = (1, 1, 1) + \alpha(2, 1, 1)$, $\alpha \in \mathbb{R}$, $\vec{y}_E(\lambda, \mu) = \lambda(2, -1, 1) + \mu(1, 2, 2)$, $\lambda, \mu \in \mathbb{R}$, so haben wir folgende Gleichung zu lösen:

$$\vec{x}_g(\alpha) = \vec{y}_E(\lambda, \mu)$$

(„Gleichsetzen“ - dabei ist auf Trennung der Parameter aus den verschiedenen Parameterdarstellungen zu achten, ein Parameter λ in \vec{x}_g hätte umbenannt werden müssen). Im Beispiel sollte man sofort das daraus entstehende lineare Gleichungssystem überblicken:

$$\begin{aligned} 2\lambda + \mu - 2\alpha &= 1 \quad (I) \\ -\lambda + 2\mu - \alpha &= 1 \quad (II) \\ \lambda + 2\mu - \alpha &= 1 \quad (III). \end{aligned}$$

Nunmehr ist dies zu lösen. Man achte auf die *im allgemeinen* sinnvolle Technik: Es wird das lineare Gleichungssystem durch Kombination der Zeilen in jedem Schritt um eine Gleichung und eine Unbekannt „kürzer“ gemacht. Wir werden hier λ, μ hinauswerfen (warum?). Zunächst λ , man erhält:

$$\begin{aligned} 5\mu - 4\alpha &= 3, \quad (I') = (I) + 2(II) \\ 2\mu - \alpha &= 1, \quad (II') = ((II) + (III))/2. \end{aligned}$$

Nunmehr kann man μ hinauswerfen:

$$3\alpha = -1, \quad 5(II') - 2(I').$$

Das ergibt $\alpha = -1/3$, und das reicht bereits aus, die gestellte Aufgabe zu lösen: Einsetzen in die Parameterdarstellung der Geraden ergibt den einzigen Schnittpunkt $\vec{x}_S = \vec{x}_g(-1/3) = (1/3, 2/3, 2/3)$. Man kann und sollte allerdings zur Kontrolle noch durch Einsetzen die Werte der andern Parameter ausrechnen: $\lambda = 0$, $\mu = 1/3$. Die einzige Lösung des Gleichungssystems lautet also $(0, 1/3, -1/3)$ (zur Reihenfolge μ, λ, α). Das ist nicht etwa die Koordinatendarstellung des Schnittpunktes, sondern man muss diese gewinnen, indem man $\alpha = -1/3$ in die Parameterdarstellung der Geraden oder aber $\lambda = 0, \mu = 1/3$ in die Parameterdarstellung der Ebene einsetzt. Beide Male kommt dasselbe heraus, wie man überprüfe. Damit ist eine wirksame Kontrolle gemacht. Wir werden im nächsten Abschnitt 5. über Schnittpunktaufgaben und lineare Gleichungssysteme allerdings noch sehen, dass die Gleichungsform für solche Aufgaben wesentlich günstiger ist, auch eine Kombination Gleichung/Parameterdarstellung.

4.5. Zur Nutzung von koordinatenfreier Vektorrechnung. Das rein geometrische Rechnen mit den geometrischen Vektorraumoperationen insbesondere in V^2 , V^3 hat für gewisse Aufgaben eine Eleganz und Einfachheit, die sich mit dem Rechnen mittels Koordinatendarstellungen bei weitem nicht erreichen lässt. Dazu zählt insbesondere das Herausarbeiten bzw. Beweisen allgemeiner Sachverhalte, die jedoch für Anwendungen durchaus konkrete Bedeutung haben. (Es ist ein Irrglaube, unter „theoretisch“ dasselbe wie „uninteressant, da

für praktische Anwendungen irrelevant“ zu verstehen.) Ein erstes Beispiel haben wir schon mit der allgemeinen Überlegung gegeben, dass man eine Parameterdarstellung für die Ebene E durch drei nicht auf einer Geraden liegende Punkte P, Q, R sofort angeben kann mit $\vec{x}_E(\lambda, \mu) = \vec{x}_P + \lambda(\vec{x}_Q - \vec{x}_P) + \mu(\vec{x}_R - \vec{x}_P)$, $\lambda, \mu \in \mathbb{R}$. Dagegen ist die konkrete Ausführung für Beispiele reichlich langweilig (!). Wir geben ein prägnantes weiteres Beispiel, dazu das ungemein wichtige Prinzip des Koeffizientenvergleichs, das sich in unendlich vielen Aufgaben nutzen lässt:

Die Aufgabe. Zu drei Punkten P, Q, R , welche nicht auf einer Geraden liegen, ist der Schnittpunkt der Seitenhalbierenden des zugehörigen Dreiecks allgemein zu berechnen. Dieser Punkt ist von Bedeutung, weil er der Schwerpunkt des Dreiecks ist bei homogener Massenverteilung. Dazu bilden wir zunächst die *freien* Vektoren $\vec{a} = \overrightarrow{PQ}$ und $\vec{b} = \overrightarrow{PR}$. Diese Vektoren sind linear unabhängig nach unserer Voraussetzung eines nicht ausgearteten Dreiecks. Das hat eine Konsequenz von allgemeiner Bedeutung: Man kann Koeffizientenvergleich machen, d.h.

$$\begin{aligned} \text{Aus } \alpha \vec{a} + \beta \vec{b} &= \lambda \vec{a} + \mu \vec{b} \text{ folgt stets („Koeffizientenvergleich“)} \\ \alpha &= \lambda, \beta = \mu, \text{ wenn } \vec{a}, \vec{b} \text{ linear unabhängig sind.} \end{aligned}$$

Das können wir sehr leicht begründen: Sei etwa $\alpha \neq \lambda$, aber $\alpha \vec{a} + \beta \vec{b} = \lambda \vec{a} + \mu \vec{b}$. Dann ist $(\alpha - \lambda) \vec{a} = (\mu - \beta) \vec{b}$, also mit $\alpha \neq \lambda$ $\vec{a} = \vec{b} (\mu - \beta) / (\alpha - \lambda)$. Somit wäre \vec{a} ein Vielfaches von \vec{b} , also \vec{a}, \vec{b} nicht linear unabhängig.

Nummehr nennen wir den Schwerpunkt S , arbeiten jedoch mit dem *freien* Vektor $\vec{s} = \overrightarrow{PS}$. Diesen Vektor stellen wir auf zwei Arten dar, wir haben, wenn wir die seitenhalbierenden Geraden durch \overline{PQ} und \overline{PR} betrachten:

$$\vec{s} = \frac{1}{2} \vec{a} + \lambda \left(\vec{b} - \frac{1}{2} \vec{a} \right) = \frac{1}{2} \vec{b} + \mu \left(\vec{a} - \frac{1}{2} \vec{b} \right).$$

Schreiben wir beide Seiten als Linearkombinationen ordentlich hin:

$$\left(\frac{1}{2} - \frac{\lambda}{2} \right) \vec{a} + \lambda \vec{b} = \mu \vec{a} + \left(\frac{1}{2} - \frac{\mu}{2} \right) \vec{b},$$

so erhalten wir mit Koeffizientenvergleich sofort das einfache lineare Gleichungssystem

$$\begin{aligned} \left(\frac{1}{2} - \frac{\lambda}{2} \right) &= \mu \\ \lambda &= \left(\frac{1}{2} - \frac{\mu}{2} \right). \end{aligned}$$

Die eindeutige Lösung ist $\lambda = \mu = 1/3$. Also haben wir die Lösung des Problems, letztlich auch mittels der ursprünglichen Ortsvektoren ausgedrückt:

$$\vec{s} = \frac{1}{3} \vec{a} + \frac{1}{3} \vec{b}, \text{ und } \vec{x}_S = \vec{x}_P + \vec{s} = \frac{1}{3} (\vec{x}_P + \vec{x}_Q + \vec{x}_R).$$

(Überprüfen Sie, dass dies Resultat mittels Einführung von Koordinatendarstellungen bei weitem nicht so leicht zu gewinnen wäre.) Unser Resultat für den Schwerpunkt sollte auch intuitiv einleuchten: Der Schwerpunkt ergibt sich einfach rechnerisch über arithmetische Mittelbildung. Wir werden die Verallgemeinerung erwarten, dass der Schwerpunkt eines Systems von n Massenpunkten P_i , $1 \leq i \leq n$, in denen jeweils dieselbe Masse sitzt, ganz entsprechend zu bilden ist:

$$\vec{x}_S = \frac{1}{n} \sum_{i=1}^n \vec{x}_{P_i}.$$

5. Schnitte linearer Gebilde und lineare Gleichungssysteme

Im Titel sind die Zusätze „linear“ zu beachten - nur dafür kann man ein so einfaches geschlossenes Bild sowie völlig problemlose Lösungsmethoden erhalten, nicht so im nichtlinearen Fall. Wir beginnen mit einer Reihe konkreter Beispiele. Darin finden sich auch nichtlineare, die man in besonders einfachen Fällen durchaus auch von Hand rechnend bewältigt. Insbesondere soll hervorgehoben werden, dass die *Logik* der Schnittbildung im linearen wie nichtlinearen Fall ganz dieselbe ist. Nur werden im allgemeinen Fall sowohl Rechnung als auch theoretischer Überblick über Lösungsmengen extrem viel schwieriger. Vorab sei einmal klargestellt, was genau unter dem Schnitt zweier geometrischer Figuren, also Punktmengen F und G zu verstehen ist:

DEFINITION 3. *Seien A, B beliebige Mengen. Dann ist der Durchschnitt (oder die Schnittmenge) von A und B die folgende Menge (gelesen: „ A geschnitten mit B “): $A \cap B := \{x \mid x \in A \text{ und } x \in B\}$. Der Schnitt zweier geometrischer Figuren ist einfach der Durchschnitt der Figuren als Punktmengen.*

Beispiele: $\{1, 2, 3, 4\} \cap \{-1, 2, 3, 5, 6\} = \{2, 3\}$. Der Durchschnitt eines Kegelmantels und einer Ebene ist einer der Kegelschnitte: Ellipse, Parabel, Hyperbel.

5.1. Beispiele (ausschließlich linear). Hier sollte man zunächst die Logik wahrnehmen, die differenziert ist je nach dem, ob beide Gebilde in Parameterdarstellung, beide in Form von Bestimmungsgleichungen (bzw. bereits Systemen davon) oder das eine in der einen, das andere in der andern Form gegeben ist. Zu dieser Logik gehört auch, wie man von der technischen Lösung eines Gleichungssystems zur Darstellung der Schnittmenge gelangt. Schließlich ist das Verständnis wichtig, dass der Schnitt zweier linearer Gebilde stets ein lineares Gleichungssystem hervorbringt, dessen Dimensionierung (Zahl der Bedingungen und der Unbekannten) man aus der Aufgabenstellung vorhersagen können sollte. Gerade an der Dimensionierung sollte man feststellen, dass Gleichungsform oder Mischform für Schnittaufgaben wesentlich günstiger sind als Parameterformen. Der zweite Hauptpunkt des Interesses ist dann natürlich auch die Technik des Lösens der resultierenden Gleichungssysteme.

5.1.1. *Beide (linearen) Gebilde in Parameterform.* (Vorbemerkung: Es handelt sich hier um eine sehr unpraktische Art, zwei Ebenen miteinander zu schneiden, es geht jedoch um die Systematik des Lösens linearer Gleichungssysteme und um die Logik der Schnittaufgaben.) Wir schneiden die Ebenen E und F , beide in Parameterform gegeben durch

$$\begin{aligned}\vec{x}_E(\alpha, \beta) &= (1, 1, 1) + \alpha(1, 2, -1) + \beta(2, 1, 3), \quad \alpha, \beta \in \mathbb{R}, \\ \vec{x}_F(\lambda, \mu) &= (2, 1, 2) + \lambda(-1, 2, 2) + \mu(2, -2, 3), \quad \lambda, \mu \in \mathbb{R}.\end{aligned}$$

Dabei ist zunächst zu beachten, dass die Parameter beider Darstellungen getrennt sind. Andernfalls muss man umbenennen. Nun zur Logik: Nehmen wir einen beliebigen Schnittpunkt $S \in E \cap F$, so haben wir folgende Aussagen:

1. Es gibt $\alpha, \beta \in \mathbb{R}$, so dass $\vec{x}_S = \vec{x}_E(\alpha, \beta)$.
2. Es gibt $\lambda, \mu \in \mathbb{R}$, so dass $\vec{x}_S = \vec{x}_F(\lambda, \mu)$.

Das bedeutet zusammen, dass wir eine Lösung der Gleichung

$$(3) \quad \vec{x}_E(\alpha, \beta) = \vec{x}_F(\lambda, \mu)$$

haben. Das besagt aber nichts etwa darüber, dass z.B. $\alpha = \lambda$ sein sollte. (Dies war der logische Grund für die Parametertrennung.) Umgekehrt gehört zu jeder solchen Lösung ein Schnittpunkt, den wir dann durch einen der Rechenausdrücke dieser Gleichung darstellen können. Nun lösen wir Gleichung (3) im Beispiel. Sie lautet konkret

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \beta \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix}$$

Dies sollte man *unmittelbar* zu folgendem linearen Gleichungssystem ummünzen können (gleiche Unbekannte geordnet, untereinander, die Konstanten isoliert auf der anderen Seite):

$$\begin{array}{l} (I) \quad \alpha + 2\beta + \lambda - 2\mu = 1 \\ (II) \quad 2\alpha + \beta - 2\lambda + 2\mu = 0 \\ (III) \quad -\alpha + 3\beta - 2\lambda - 3\mu = 1 \end{array} .$$

Dies System hat drei Bedingungen für vier Unbestimmte, es ist also im Normalfall eine unendliche Lösungsmenge mit einem freien Parameter zu erwarten. Strategisch verläuft die Lösung so, dass man jeweils eine Unbekannte hinauswirft und eine Gleichung weniger bekommt. Hinsichtlich unserer Aufgabenstellung genügt es, auf eine Gleichung zu kommen, in der nur noch λ, μ vorkommen. (α, β wäre ebenso gut, nicht etwa α, λ .) Wir werfen α hinaus und erhalten:

$$\begin{array}{l} (I') = 2(I) - (II) \quad 3\beta + 4\lambda - 6\mu = 2 \\ (II') = (I) + (III) \quad 5\beta - \lambda - 5\mu = 2 \end{array} .$$

Nunmehr β hinaus:

$$(I'') = 5(I') - 3(II') \quad 23\lambda - 15\mu = 4$$

Der zweite Schritt im Lösen eines solchen Gleichungssystems besteht nach dieser Arbeit darin, die Lösungsmenge aufzuschreiben. Wir beobachten: Zu jedem beliebig vorgegebenen Wert von λ können wir eindeutig μ, β, α ausrechnen. Unsere Lösungsmenge hat also hier einen freien Parameter. (Warum ist es günstiger, λ dafür zu wählen als μ ?) Wir haben also alle andern Unbestimmten durch λ auszudrücken und erhalten aus $(I''), (I'), (I)$ der Reihe nach:

$$\mu = \frac{23}{15}\lambda - \frac{4}{15}, \quad \alpha = -\frac{7}{5}\lambda + \frac{1}{5}, \quad \beta = \frac{26}{15}\lambda + \frac{2}{15}.$$

Die Lösungsmenge in Parameterform aufgeschrieben lautet damit (unter Beibehaltung der Reihenfolge $\alpha, \beta, \lambda, \mu$):

$$L(\lambda) = \frac{1}{5}(1, 2, 0, -4) + \lambda \left(-\frac{7}{5}, \frac{26}{15}, 1, \frac{23}{15} \right), \quad \lambda \in \mathbb{R}.$$

Dies kann man geometrisch interpretieren als eine Gerade im \mathbb{R}^4 , und wir werden sehen, dass jede Lösungsmenge eines linearen Gleichungssystems entweder leer ist oder aber geometrisch als lineares Gebilde (Verallgemeinerung von Punkt, Gerade, Ebene,...) zu deuten ist und eine entsprechende Parameterdarstellung besitzt. Aber man verwechsle das in unserem Fall nicht mit der gesuchten Schnittmenge! Um diese zu erhalten, hätten wir mit (I'') und der Beziehung $\mu = (23\lambda - 4)/15$ aufhören können. Das haben wir einzusetzen (so stets der letzte Schritt bei Schnittaufgaben mit Parameterformen) in den Ausdruck $\vec{x}_F(\lambda, \mu)$. Also ist die Schnittmenge, schon in Parameterdarstellung:

$$\begin{aligned} \vec{x}_{g_S}(\lambda) &= (2, 1, 2) + \lambda(-1, 2, 2) + \left(\frac{23}{15}\lambda - \frac{4}{15} \right) (2, -2, 3) \\ &= \left(\frac{22}{15}, \frac{23}{15}, \frac{6}{5} \right) + \lambda \left(\frac{31}{15}, -\frac{16}{15}, \frac{33}{5} \right). \end{aligned}$$

Es resultiert die geometrisch zu erwartende Schnittgerade g_S der beiden Ebenen im dreidimensionalen Raum. Man sollte noch daran denken, eine umparametrisierung vorzunehmen mit dem einfacheren Richtungsvektor $(31, -16, 99)$, der ohne Brüche auskommt. Fassen wir das ganze Verfahren zusammen:

1. Parameter trennen
2. Aufstellen eines linearen Gleichungssystems nach Gleichsetzen der Terme (d.h. Rechenausdrücke) der Parameterdarstellungen
3. Lösen des Gleichungssystems
4. Einsetzen der Lösungsmenge in der Parameterfolge nur *einer* der Parameterdarstellungen in die zugehörige Parameterdarstellung ergibt eine Parameterdarstellung des Schnittgebildes. (Auf Endform achten.)

5.1.2. *Beide (linearen) Gebilde in Gleichungsform.* Betrachten wir dieselbe Aufgabe, zwei Ebenen zu schneiden, die jedoch in Gleichungsform gegeben sind:

$$\begin{aligned} E: & 2x - y + 3z = 1 \\ F: & x + 3y - 4z = 1 \end{aligned}$$

Viel günstiger, nur zwei Gleichungen mit drei Parametern. Zur Logik: Sei $\vec{x}_S = (x_S, y_S, z_S)$ der Ortsvektor eines beliebigen Schnittpunktes. Dann bedeutet $S \in E$ gerade $2x_S - y_S + 3z_S = 1$, analog bedeutet $S \in F$ die Erfüllung der zweiten Gleichung, $S \in E \cap F$ bedeutet also, dass *die Koordinaten des Schnittpunktes das Gleichungssystem erfüllen*. Somit ist die Lösungsmenge des Gleichungssystems *unmittelbar* die Schnittmenge! Das Gleichungssystem wird nun nach demselben Muster wie oben gelöst: Hinauswerfen von x bringt

$$7y - 11z = 1.$$

Wir erhalten den erwarteten einen freien Parameter, wählen dazu z . Das ergibt $y = 1/7 + 11/7z$, $x = 1 + 4z - 3y = 1 + 4z - 3/7 - 33/7z = 4/7 - 5/7z$ (aus der zweiten Gleichung). Die Lösungsmenge in ordentlicher Parameterform:

$$\vec{x}_{g_S}(z) = \frac{1}{7}(4, 1, 0) + z \left(-\frac{5}{7}, \frac{11}{7}, 1 \right), \quad z \in \mathbb{R}.$$

Damit ist die Schnittgerade ausgerechnet. Später werden wir lernen, Parameterdarstellungen von Ebenen systematisch in Gleichungsform zu bringen, im Zusammenhang mit Skalarprodukt und Vektorprodukt.

Das Verfahren zusammengefasst:

1. Die Gleichungen (oder auch schon Systeme) beider Gebilde zu einem Gleichungssystem zusammenfassen. (Die Koordinaten-Unbestimmten müssen überall *gleich* benannt sein.)
2. Lösen des entstehenden linearen Gleichungssystems (Unbekannte sind die Koordinaten!).
3. Die Lösungsmenge in parametrisierter Form ist unmittelbar eine Parameterdarstellung des Schnittgebildes.

5.1.3. *Mischform: Eine Parameterdarstellung und eine Gleichung, linearer Fall.* Wir schneiden g mit E , gegeben durch $\vec{x}_g(\lambda) = (1, 2, 2) + \lambda(2, -1, 3)$, $\lambda \in \mathbb{R}$, E mit der Gleichung $2x - y - z = 0$. Logik: Jeder Punkt $S \in g \cap E$ muss auf der Geraden liegen, also mit einem Wert λ_S muss $\vec{x}_S = \vec{x}_g(\lambda_S)$ gelten. Dessen Koordinaten müssen nun auch die Ebenengleichung erfüllen. Dies ergibt folgende Gleichung für λ_S :

$$2 + 4\lambda - 2 + \lambda - 2 - 3\lambda = 0, \text{ also } \lambda = 1.$$

Dies ist nun wieder in die Parameterdarstellung einzusetzen. Es gibt also genau einen Schnittpunkt S , und es ist $\vec{x}_S = (3, 1, 5)$. Man sollte sich nun vorstellen können, dass der Schnitt zweier Ebenen bei Mischform darauf hinausläuft, eine lineare Parametergleichung mit zwei Parametern als Unbestimmten zu lösen. Es bleibt ein freier Parameter, Einsetzen funktioniert dann wie im ersten Beispiel (zwei Ebenen, beide in Parameterform). Das Verfahren ist also:

1. Einsetzen der Koordinaten des Terms der Parameterdarstellung des einen Gebildes in die Gleichung(en) des andern Gebildes.
2. Lösen des entstehenden Gleichungssystems. (Die Unbekannten sind die freien Parameter.)
3. Einsetzen der Lösungsmenge in den Term der Parameterdarstellung ergibt eine Parameterdarstellung des Schnittgebildes.

5.2. Zwei nichtlineare Beispiele. Wir schneiden die Parabel, die durch $\vec{x}(t) = t(1, 1, 2) + t^2(1, 2, 1)$, $t \in \mathbb{R}$, gegeben ist, mit der Ebene, deren Gleichung $2x - y + z = 1$ lautet. Die Logik funktioniert offenbar genau wie im vorigen Beispiel: Die Koordinaten von $\vec{x}(t)$ sind in die Ebenengleichung einzusetzen, das ergibt die quadratische Gleichung $2t + 2t^2 - t - 2t^2 + 2t + t^2 = 1$, und man erhält die Lösungen $t_{1,2} = -\frac{3}{2} \pm \frac{1}{2}\sqrt{13}$, die wiederum in $\vec{x}(t)$ einzusetzen sind. Entsprechend dem Lösungsverhalten quadratischer Gleichungen verhalten sich die geometrischen Möglichkeiten beim Schnitt einer Parabel mit einer Ebene: Kein Schnittpunkt, ein Schnittpunkt (Berührungspunkt) oder zwei. Dazu kommt noch der Fall, dass die Parabel ganz in der Ebene

liegt, was sich beim beschriebenen Rechenvorgang darin äußern würde, dass die Gleichung $0 = 0$ für t entstünde. Dann ist *jede* Zahl $t \in \mathbb{R}$ Lösung, somit der Schnitt die gesamte Parabel.

Auch in den anderen Fällen bleibt die Logik dieselbe, das Unangenehme sind nur die entstehenden nichtlinearen Gleichungen. Aber in einfachen Fällen geht das. Schneiden wir etwa das Paraboloid $z = x^2 + y^2$ mit der Ebene $x + y - z = 1$, so erhalten wir für die Koordinaten der Schnittpunkte die Gleichung $x^2 + y^2 - x - y + 1 = 0$. Fassen wir y als freien Parameter auf, so haben wir eine quadratische Gleichung mit den Lösungen

$$x_{1,2}(y) = \frac{1}{2} \pm \sqrt{y - y^2 - \frac{3}{4}},$$

wenn denn der Term unter der Wurzel positiv ist. Man sieht leicht, dass er stets < 0 ist, also gibt es keine Lösung, die Schnittmenge ist leer, Ebene und Paraboloid laufen knapp aneinander vorbei. An diesem Beispiel sollte man auch erkennen, dass Unlösbarkeit einer Gleichung durchaus einen Sinn haben kann. Die Verfahren wären wie oben in den linearen Fällen zu nennen, allerdings mit dem gewaltigen Unterschied, dass die entstehenden Gleichungen bzw. Gleichungssysteme nichtlinear sind, was das *exakte* Lösen im Normalfall sogar *prinzipiell* unmöglich macht, nicht jedoch das Finden von Näherungslösungen.

5.2.1. *Lineare Gleichungssysteme mit äußeren Parametern.* Stellen wir uns die Aufgabe, etwa eine Schar (oder Familie) von Geraden mit einer Ebene zu schneiden, sieht das etwa so aus: In der Parametrisierung der Geraden kommt neben dem freien Parameter ein äußerer Parameter vor, sagen wir a . Zu jedem Wert von a bekommen wir durch Einsetzen eine konkrete Parameterdarstellung einer speziellen Geraden, allgemein gehört zu a die Gerade g_a , das könnte so aussehen - entsprechend eine Schar von Parameterdarstellungen:

$$\vec{x}_{g_a}(\lambda) = (1, 2, 1) + \lambda(a, 2, 3), \quad \lambda \in \mathbb{R}, \quad \text{für alle } a \in \mathbb{R}.$$

Nun stellen wir uns die Aufgabe, allgemein *jede* dieser Geraden mit der Ebene $x + y - z = 1$ zu schneiden. Am bisherigen Verfahren ändert sich überhaupt nichts, und wie angekündigt wird ein äußerer Parameter einfach „mitgeschleppt“, lediglich sind Fallunterscheidungen zu erwarten und gegebenenfalls zu behandeln. Einsetzen des Ausdrucks der Parameterdarstellungsschar in die Ebenengleichung ergibt: $\lambda(a + 2 - 3) + 1 + 2 - 1 = 1$, also $\lambda(a) = 1/(1 - a)$ für $a \neq 1$. Also ist in diesen Fällen der einzige Schnittpunkt $\vec{x}_{s_a} = \vec{x}_{g_a}(1/(1 - a))$. Was ist bei $a = 1$ der Fall? Dies untersuchen wir, indem wir speziell diesen kritischen Wert in die Gleichung einsetzen. Es entsteht die Gleichung $2 = 1$, also ein Widerspruch. Es gibt also gar keinen Schnittpunkt. Das entspricht wiederum völlig der Geometrie: Die Gerade g_1 hat den Richtungsvektor $(1, 2, 3)$, und der ist parallel zur Ebene, wie wir später eleganter mittels des Skalarproduktes ausrechnen können, im Augenblick aber auch schon nach Bilden etwa der Parameterdarstellung $\vec{y}_E(x, y) = (x, y, x + y - 1) = (0, 0, -1) + x(1, 0, 1) + y(0, 1, 1)$, $x, y \in \mathbb{R}$ und mit Darstellung von $(1, 2, 3)$ als $(1, 0, 1) + 2(0, 1, 1)$ (Linearkombination der Richtungsvektoren der Ebene) belegen können. Dass es keinen Schnittpunkt im Falle $a = 1$ geben kann, folgt nun daraus, dass $(1, 2, 1)$ nicht auf der Ebene liegt. Das wissen wir aber alles schon nach unserer Behandlung des linearen Gleichungssystems mit viel weniger Mühe!

Noch ein Beispiel zur Eliminationstechnik (zum Herauswerfen von Unbestimmten also) bei Vorliegen äußerer Parameter:

$$\begin{array}{rcl} 2ax & -y & = 1 \\ 3x & +5ay & = 0 \end{array}$$

Herauswerfen von y : Man bildet $5a(I) + (II)$ und erhält $(10a^2 + 3)x = 5a$. Offenbar ist für jedes a die Lösung eindeutig auszurechnen, also keine verschiedenen Fälle zu behandeln. Stünde jedoch in der ersten Zeile $2ax + y = 1$, so käme man auf $(10a^2 - 3)x = 5a$, so zwei kritische Fälle $a_{1,2} = \pm\sqrt{30}/10$ resultieren (in allen andern Fällen eindeutige Lösbarkeit). Man geht nun in der Rechnung zurück bis unmittelbar vor den ersten in den kritischen Fällen unerlaubten Schritt (Division) und findet die Gleichung $(10a^2 - 3)x = 5a$. Einsetzen von $a_{1,2}$ ergibt damit eine Gleichung $0 = 5a_{1,2}$, die in jedem Falle ein Widerspruch ist. Hätte dagegen in der ersten Zeile $2ax + y = 0$ gestanden, so wäre man in diesen Fällen (es sind dieselben!) auf unendliche Lösungsmengen gekommen. „Kritischer Fall“ bedeutet also nicht immer Unlösbarkeit.

5.3. Zur Systematik linearer Gleichungssysteme.

5.3.1. *Vorbereitung: Grundbegriffe der linearen Algebra.* Lineare Gleichungssysteme zeigen in ihrem Lösungsverhalten eine einfache Systematik. Wir werden sehen, dass stets ein lineares Gebilde wie Punkt, Gerade, Ebene, dreidimensionales Analogon,... als Lösungsmenge herauskommt. Eine solche Menge nennt man einen affinen Unterraum (eines Grundraums, in dem die Lösungsmenge gesucht wird). Ein affiner Unterraum ist einfach ein Unter-Vektorraum, der durch einen Aufpunktvektor vom Ursprung (der Null) abgestützt wird. Hier die allgemeine

DEFINITION 4. *Sei V ein Vektorraum beliebiger Definition. Dann ist ein affiner Unterraum eine solche Teilmenge von V , die entweder nur einen einzigen Vektor \vec{x}_0 enthält oder aber einer Parametrisierung folgender Form fähig ist:*

$$\vec{x}(\lambda_1, \dots, \lambda_n) = \vec{x}_0 + \sum_{i=1}^n \lambda_i \vec{a}_i, \text{ mit festen Vektoren } \vec{a}_i, 1 \leq i \leq n.$$

Diejenigen affinen Unterräume, welche nur $\vec{0}$ enthalten oder eine Parametrisierung der angegebenen Form mit $\vec{x}_0 = \vec{0}$ zulassen, sind genau die (Vektor-) Unterräume, die man wie folgt definiert:

DEFINITION 5. *Eine Teilmenge M eines Vektorraums V über K heißt Unterraum (ausführlicher: Untervektorraum), wenn folgende Bedingungen gelten:*

- (1) $M \neq \emptyset$.
- (2) Für alle $\vec{x}, \vec{y} \in M$ ist $\vec{x} + \vec{y} \in M$.
- (3) Für alle $\lambda \in K, \vec{x} \in M$ ist $\lambda \vec{x} \in M$.

Man beachte: Diese Bedingungen sichern genau, dass M mit den von V „geerbten“ linearen Operationen wieder ein Vektorraum ist. Dazu müssen nämlich nur die Operationen in M ausführbar sein, d.h. stets wieder Resultate in M liefern. Die Gültigkeit der Vektorraum-Axiome ist dann klar, weil sie ja bereits für die größeren Menge V gelten. Zur Illustration beider Begriffe: Die affinen Unterräume von V^3 bzw. \mathbb{R}^3 haben eine der folgenden Gestalten: $\{\vec{x}_0\}, \{\vec{x}_0 + \lambda \vec{a} \mid \lambda \in \mathbb{R}\}$, mit $\vec{a} \neq \vec{0}$, $\{\vec{x}_0 + \lambda \vec{a} + \mu \vec{b} \mid \lambda, \mu \in \mathbb{R}\}$, mit \vec{a}, \vec{b} linear unabhängig, schließlich der ganze Raum selbst. Es sind also die Mengen aus einem einzigen Punkt, die Geraden und die Ebenen, dann der ganze Raum. Die Untervektorräume sind dementsprechend $\{\vec{0}\}$, alle Ursprungsgeraden, alle Ursprungsebenen, der ganze Raum selbst.

Wir haben bereits gesehen, dass der Grund-Rechenausdruck der Vektorrechnung die Linearkombination ist. Dementsprechend ranken sich auch die Grundbegriffe der Theorie der Vektorräume (man nennt sie „lineare Algebra“) um die Analyse der Eigenschaften von Linearkombinationen. So insbesondere der folgende wichtigste algebraische Grundbegriff, den wir speziell für einen Vektor oder zwei Vektoren für die Geraden und Ebenen schon benutzt haben, den der linearen Unabhängigkeit eines Systems von endlich vielen Vektoren.

DEFINITION 6. *Ein System $\vec{a}_1, \dots, \vec{a}_n$ ($n \in \mathbb{N}, n \geq 1$) heißt linear unabhängig, wenn für alle Zahlen $\lambda_1, \dots, \lambda_n \in K$ aus $\sum_{i=1}^n \lambda_i \vec{a}_i = \vec{0}$ stets folgt, dass $\lambda_i = 0$ für alle $i, 1 \leq i \leq n$. Ein solches System heißt linear abhängig, wenn es nicht linear unabhängig ist. Man nennt die Linearkombination $0 \vec{a}_1 + 0 \vec{a}_2 + \dots + 0 \vec{a}_n$, also mit lauter Koeffizienten Null, die triviale Linearkombination. Sie ergibt stets den Nullvektor, trivialerweise.*

Beispiel: $(1, 0, 0)$ und $(0, 1, 0)$ sind linear unabhängig. Denn $\lambda_1(1, 0, 0) + \lambda_2(0, 1, 0) = (\lambda_1, \lambda_2, 0)$, und das wird $\vec{0}$ nur dann, wenn $\lambda_1 = \lambda_2 = 0$. Eine wichtige Verallgemeinerung: Im $\mathbb{R}^n, n \geq 1$, ist stets das System $\vec{e}_1 = (1, 0, \dots, 0), \vec{e}_2 = (0, 1, 0, \dots, 0), \dots, \vec{e}_n = (0, 0, \dots, 0, 1)$ linear unabhängig. (\vec{e}_i hat also in der i -ten Komponente eine Eins, sonst lauter Nullen.) Der Grund: $\sum_{i=1}^n \lambda_i \vec{e}_i = (\lambda_1, \lambda_2, \dots, \lambda_n)$, und das ist $\vec{0}$ nur dann, wenn jede Komponente, also jeder Koeffizient λ_i , Null ist. Folgendes Beispiel sollte den typischen Anfängerfehler vermeiden lassen: Das System $(2, 1), (1, 3), (-4, 3)$ im \mathbb{R}^2 ist nicht linear unabhängig. Denn es ist z.B. $-3(2, 1) + 2(1, 3) + (-1)(-4, 3) = (0, 0)$. Aber die Vorfaktoren sind nicht alle Null! Der unvorsichtige Anfänger beobachtet, dass kein Vektor der drei ein Vielfaches des andern ist (wie im Beispiel) und schließt

fälschlich auf lineare Unabhängigkeit. Man beachte, dass lineare Unabhängigkeit eine Eigenschaft des ganzen Systems ist, nicht eine, die sich in Eigenschaften der Paare bereits ausdrücken ließe! Übrigens besteht im \mathbb{R}^2 jedes linear unabhängige System aus maximal zwei Vektoren. Dazu merken wir (ohne jede Begründung) ein wichtiges allgemeines Resultat und einen wichtigen weiteren Begriff an:

SATZ 2. *Zu jedem Vektorraum V über K gibt es stets eine eindeutig bestimmte maximal mögliche Länge von Systemen linear unabhängiger Vektoren. Diese Länge heißt Dimension von V über K . (Begriffsbildung und Resultat gelingen übrigens auch für unendlich-dimensionale Vektorräume.)*

Beispiele: V^2, V^3 haben Dimensionen 2,3, allgemein hat \mathbb{R}^n Dimension n .

Es gibt ein paar verschiedene Versionen der Begriffe der linearen Unabhängigkeit, die praktisch vielfach nützlich sind:

SATZ 3. *Ein System $\vec{a}_1, \dots, \vec{a}_n$ von Vektoren ist genau dann **linear unabhängig**, wenn eine der folgenden gleichwertigen Bedingungen erfüllt ist:*

- (1) *Jede Linearkombination der Vektoren ist eindeutig.*
- (2) *Keiner der Vektoren ist als Linearkombination der anderen darstellbar.*

Aus (1) folgt sicher die lineare Unabhängigkeit, da diese definitionsgemäß nur (1) speziell für den Nullvektor fordert. Nun folgt aber (1) aus (2): Wenn $\sum_{i=1}^n \lambda_i \vec{a}_i = \sum_{i=1}^n \mu_i \vec{a}_i$ und etwa $\lambda_1 \neq \mu_1$, dann ist $(\lambda_1 - \mu_1) \vec{a}_1 = \sum_{i=2}^n (\mu_i - \lambda_i) \vec{a}_i$, also $\vec{a}_1 = \frac{1}{\lambda_1 - \mu_1} \sum_{i=2}^n (\mu_i - \lambda_i) \vec{a}_i$. Also folgt die Verneinung von (2) aus der Verneinung von (1) und daher folgt (1) aus (2). Nun folgt wieder (2) aus der Definition. Denn wäre etwa $\vec{a}_1 = \sum_{i=2}^n \lambda_i \vec{a}_i$, so hätte man $(-1) \vec{a}_1 + \sum_{i=2}^n \lambda_i \vec{a}_i = \vec{0}$. Mit dem Ringschluß aber ist die Gleichwertigkeit aller drei Bedingungen gezeigt.

Vervollständigt wird das begriffliche Grundgerüst, indem wir betrachten, was man mit den Linearkombinationen eines (nicht notwendig linear unabhängigen) Systems von Vektoren alles darstellen kann:

DEFINITION 7. *Sei V ein Vektorraum über K , ferner $\vec{a}_i, 1 \leq i \leq n$, ein nicht notwendig linear unabhängiges System von Vektoren. Dann heißt die Menge aller Linearkombinationen $\sum_{i=1}^n \lambda_i \vec{a}_i$ mit $\lambda_i \in K, 1 \leq i \leq n$, der von den \vec{a}_i erzeugte Unterraum. Das System der \vec{a}_i heißt dann Erzeugendensystem für diesen Unterraum. Wird der ganze Raum V erzeugt, so handelt es sich um ein Erzeugendensystem für V über K .*

Fassen wir nunmehr die Eigenschaften der lineare Unabhängigkeit und eines Erzeugendensystems zusammen:

DEFINITION 8. *Wenn $\vec{a}_i, 1 \leq i \leq n$ ein linear unabhängiges Erzeugendensystem ist für V über K , so nennt man das System eine Basis für V über K . (Stets existiert eine solche - sogar sehr viele, für jeden Vektorraum $\neq \{\vec{0}\}$. Die Länge einer beliebigen Basis ist stets die Dimension des Raums - der Nullraum hat Dimension Null. Wieder gelingt die Verallgemeinerung auf unendliche Dimension.) Mit dem vorigen Satz folgt, dass jeder Vektor eindeutig als Linearkombination einer solchen Basis darstellbar ist.*

Beispiel: $\vec{e}_1, \dots, \vec{e}_n$ bildet eine Basis (sogenannte „kanonische Basis“) des \mathbb{R}^n . Das System $(1, 2), (2, 1)$ bildet auch eine Basis für \mathbb{R}^2 , dagegen ist $(1, 1), (2, 3), (1, 2)$ ein Erzeugendensystem für \mathbb{R}^2 , aber keine Basis, da nicht linear unabhängig.

Zur elementaren Rechentechnik:

Gegeben sei ein System $\vec{a}_i, 1 \leq i \leq k$, von Vektoren eines \mathbb{R}^n . Wie stellt man fest, ob es linear unabhängig ist? Wenn $k > n$, so ist nichts zu tun: Das System muss linear abhängig sein, da der Raum nur Dimension n besitzt. Wenn $k \leq n$, so kann man gemäß der Definition so vorgehen: Man bildet das lineare Gleichungssystem $\sum_{i=1}^k \lambda_i \vec{a}_i = \vec{0}$ und löst es. Hat die Lösungsmenge einen freien Parameter, so ist das System linear abhängig, sonst unabhängig. Hier ist ein praktischeres Verfahren: Man schreibt die Vektoren (sagen wir als Spaltenvektoren) auf als Rechteckschema. Dann macht man Spaltenumformungen gerade so wie die Zeilenumformungen bei linearen Gleichungssystemen: Eine Spalte (j) darf durch $\alpha(j) + \beta(k)$, mit einer andern Spalte (k) und $\alpha \neq 0$ (!) ersetzt werden. Man sorgt dabei systematisch für Nullen, und zwar so, dass man von einer Spalte zur nächsten

mindestens eine Null mehr hat. Wenn dann am Ende keine Nullspalte erzeugt wurde, so ist das System linear unabhängig, sonst nicht. (Hat man Zeilenvektoren, so läuft die Sache entsprechend mit Zeilenumformungen.) Beispiel: Betrachten wir das System im \mathbb{R}^4

$$\begin{pmatrix} 1 \\ -3 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \\ 2 \\ 4 \end{pmatrix}$$

und nennen wir die Spalten (I) , (II) , (III) . Ersetzen von (II) durch $2(II) - (I)$ und von (III) durch $(III) - (I)$ ergibt

$$\begin{pmatrix} 1 \\ -3 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 \\ 7 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ 4 \end{pmatrix}.$$

Nun könnte man die neue letzte Spalte (III') ersetzen durch $(III'') = (III') - 4(II')$ und sehen, dass kein Nullvektor entstanden ist. Das Verfahren ist beendet. Das System ist linear unabhängig. Aber schon nach dem ersten Schritt ist das klar: Die beiden Spalten rechts, (II') und (III') , sind offenbar linear unabhängig voneinander (keine Vielfachen voneinander), und $(I') = (I)$ ist offenbar wegen der Nullen in der dritten Komponente aus (II') , (III') nicht erzeugbar. Also ist das ganze System linear unabhängig. (Diese Überlegung hätte man allgemein durchzuführen, um das beschriebene Verfahren theoretisch zu rechtfertigen. Dazu wäre zu zeigen, dass sich durch die Spaltenumformungen nichts an der Eigenschaft der linearen Unabhängigkeit bzw. Abhängigkeit ändert. Beides ist nicht schwierig.)

Das Verfahren bringt noch etwas mehr: Sogar der von den Spaltenvektoren erzeugte Unterraum (im Beispiel des \mathbb{R}^4) ändert sich bei den beschriebenen Umformungen nicht. Auf diese Weise kann man etwa Richtungsvektoren durch einfachere (mit möglichst vielen Nullen) ersetzen, Beispiel: Für die Ebene E mit der Parameterdarstellung $\vec{x}_E(\lambda, \mu) = \lambda(1, 2, 2) + \mu(1, 2, 3)$, $\lambda, \mu \in \mathbb{R}$, kann sofort die einfachere Parameterdarstellung $\vec{y}_E(\alpha, \beta) = \alpha(1, 2, 2) + \beta(0, 0, 1)$, $\alpha, \beta \in \mathbb{R}$, gegeben werden, ersichtlich geht die Ebene durch die z -Achse. Schließlich wäre noch der erste Richtungsvektor ersetzbar durch $(1, 2, 0)$. Einfacher geht es dann wirklich nicht mehr.

5.3.2. *Lineare Gleichungssysteme und Matrizen.* Wir analysieren theoretisch, was ein lineares Gleichungssystem eigentlich mathematisch bedeutet. Daraus wird sich eine sehr nützliche Übersicht ergeben, die man immer heranziehen sollte, wenn man ein lineares Gleichungssystem einiger Komplexität vor sich hat. Betrachten wir dazu ein konkretes Beispiel:

$$\begin{aligned} 2x - 3y + z &= 1 \\ -3x + y &= 2 \\ 4x - 3y + 5z &= 0 \end{aligned}$$

Das können wir in drei Hauptstücke zergliedern: den Vektor auf der rechten Seite (die drei Zahlen zum Vektor zusammengefasst), den Vektor der Unbestimmten x, y, z und schließlich die Matrix der Koeffizienten bei den Unbestimmten, die wir im Beispiel entsprechend dem Schema des Gleichungssystems so aufschreiben:

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

Nun produziert auf der linken Seite dies Koeffizientenschema drei Zahlen aus einem beliebigen Vektor, wenn wir für x, y, z beliebige Zahlen einsetzen. Dies definieren wir als die abstrakte Operation „Matrix mal Vektor“,

zunächst im Beispiel:

$$\begin{pmatrix} 2 & -3 & 1 \\ -3 & 1 & 0 \\ 4 & -3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - 3y + z \\ -3x + y \\ 4x - 3y + 5z \end{pmatrix}.$$

Es entstehen also in den Komponenten gerade die linken Seiten des Gleichungssystems! Gewonnen werden sie, indem man den Spaltenvektor nacheinander auf die Matrixzeilen legt, übereinanderliegende Zahlen multipliziert und diese Produkte addiert. Jede Matrixzeile liefert damit einen Eintrag in den Resultatvektor. Nun die allgemeine Definition, dann ein paar Beispiele:

DEFINITION 9. Sei A eine $(m \times n)$ -Matrix (d.h. mit m Zeilen und n Spalten) mit reellen Koeffizienten. Den Koeffizienten in der i -ten Zeile und j -ten Spalte nennen wir a_{ij} . Wir schreiben dann auch $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ oder kurz $A = (a_{ij})$. Sei ferner \vec{x} ein beliebiger Spaltenvektor der Länge n mit reellen Komponenten, also $\vec{x} \in \mathbb{R}^n$ mit den Komponenten x_j , $1 \leq j \leq n$. Dann ist definiert:

$$A\vec{x} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}.$$

$A\vec{x}$ ist also ein Spaltenvektor der Länge m , und seine Komponenten lauten allgemein für $1 \leq i \leq m$: $\sum_{j=1}^n a_{ij}x_j$. Wenn wir mitverstehen, dass die Vektoren hier stets Spaltenvektoren sind, ferner i der Zeilenindex der Matrix, j der Spaltenindex, so können wir die Definition auch kurz so schreiben:

$$(a_{ij})(x_j) = \left(\sum_{j=1}^n a_{ij}x_j \right)$$

Beispiele:

$$\begin{pmatrix} 1 & 2 & -3 \\ 4 & 3 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -4 \\ 7 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ -2 & 3 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ -2x + 3y \\ 3x + 4y \end{pmatrix}.$$

Was hat man von dieser Trennung der Bestandteile eines linearen Gleichungssystems, über die starke Vereinfachung der symbolischen Schreibweise $A\vec{x} = \vec{b}$ hinaus? Der Vorteil liegt einfach darin, dass die Matrix als Abbildung aufzufassen ist, und zwar als lineare mit schönen einfachen Eigenschaften. Aus diesen ergibt sich gerade der angestrebte einfache Überblick. Wir definieren:

DEFINITION 10. Sei A eine reelle $(m \times n)$ -Matrix. Dann definiert A folgende (ebenfalls mit A bezeichnete) Abbildung mit Definitionsbereich \mathbb{R}^n und Wertebereich \mathbb{R}^m :

$$A: \mathbb{R}^n \rightarrow \mathbb{R}^m \\ \vec{x} \mapsto A\vec{x}.$$

Man liest die erste Zeile: Die Abbildung A geht von \mathbb{R}^n (sog. Urbildraum) nach \mathbb{R}^m (Werteraum). Die zweite Zeile: Einem Vektor $\vec{x} (\in \mathbb{R}^n)$ wird der Vektor $A\vec{x}$ (definiert zuvor als Matrix mal Vektor) zugeordnet.

Die Lösungsmenge von $A\vec{x} = \vec{b}$ ist also gerade die Menge aller Vektoren, welche durch A auf \vec{b} abgebildet werden. Man nennt diese Menge auch das Urbild von \vec{b} unter der Abbildung A .

Die schöne Eigenschaft dieser Abbildungen ist folgende:

SATZ 4. Die durch eine Matrix A definierte Abbildung ist stets linear, d.h. es gelten folgende Formeln:

$$\left. \begin{aligned} A(\vec{x} + \vec{y}) &= A\vec{x} + A\vec{y} \\ A(\lambda\vec{x}) &= \lambda A\vec{x} \end{aligned} \right\} \text{Linearität von } A.$$

Einsetzen in die Summenformel der Definition von Matrix mal Vektor ergibt dies unmittelbar: Setzen wir $A = (a_{ij})_{ij}$, $\vec{x} = (x_j)_j$, $\vec{y} = (y_j)_j$, so lautet die i -te Komponente von $A(\vec{x} + \vec{y})$ so:

$$\sum_j a_{ij}(x_j + y_j) = \sum_j a_{ij}x_j + \sum_j a_{ij}y_j.$$

Auf der rechten Seite steht aber die i -te Komponente von $A\vec{x} + A\vec{y}$. Die i -te Komponente von $A(\lambda\vec{x})$ lautet

$$\sum_j a_{ij}\lambda x_j = \lambda \sum_j a_{ij}x_j,$$

was wiederum die i -te Komponente von $\lambda A\vec{x}$ ist. Das ist schon der Beweis des Satzes, und dieser hat zwei wichtigen Folgen für lineare Gleichungssysteme:

SATZ 5. Ein homogenes lineares Gleichungssystem, d.h. der Form $A\vec{x} = \vec{0}$ mit einer $(m \times n)$ -Matrix A , hat einen Untervektorraum von \mathbb{R}^n als Lösungsmenge, den sogenannten Kern von A , kurz $\text{Kern}(A)$. $\text{Kern}(A)$ ist nun entweder $\{\vec{0}\}$ oder hat eine Basis $\vec{a}_1, \dots, \vec{a}_k$, mit $1 \leq k \leq n$. Im letzteren Fall ist der Kern also folgendermaßen zu parametrisieren: $\vec{x}(\lambda_1, \dots, \lambda_k) = \sum_{i=1}^k \lambda_i \vec{a}_i$, $\lambda_1, \dots, \lambda_k \in \mathbb{R}$. Die Zahl k der freien Parameter ist eindeutig bestimmt.

SATZ 6. Ein inhomogenes lineares Gleichungssystem $A\vec{x} = \vec{b}$, definitionsgemäß mit $\vec{b} \neq \vec{0}$, hat entweder keine Lösung (Lösungsmenge leer), oder aber eine einzige Lösung \vec{x}_0 - dies genau im Falle $\text{Kern}(A) = \{\vec{0}\}$, oder aber seine Lösung lautet in parametrisierter Form $\vec{x}(\lambda_1, \dots, \lambda_k) = \vec{x}_0 + \sum_{i=1}^k \lambda_i \vec{a}_i$, $\lambda_1, \dots, \lambda_k \in \mathbb{R}$, wobei \vec{x}_0 eine spezielle Lösung von $A\vec{x} = \vec{b}$ ist und $\vec{a}_1, \dots, \vec{a}_k$ eine Basis von $\text{Kern}(A)$. Die Lösungsmenge hat insbesondere für jeden Vektor \vec{b} dieselbe Zahl von freien Parametern. Insbesondere ist die Lösungsmenge leer oder aber ein affiner Unterraum von \mathbb{R}^n . Man nennt im letzten der drei Fälle gern den Ausdruck $\vec{x}_0 + \sum_{i=1}^k \lambda_i \vec{a}_i$ die allgemeine Lösung des Gleichungssystems. (Durch spezielle Werte der Koeffizienten erhält man eine beliebige spezielle Lösung.) Entsprechend ist $\sum_{i=1}^k \lambda_i \vec{a}_i$ die allgemeine Lösung der zugehörigen homogenen Gleichung.

So wichtig diese Aussagen in praktischer Hinsicht sind, so einfach sind sie zu beweisen, sehen wir einmal von der Existenz von Basen und der Eindeutigkeit der Dimension (k) vom Kern ab. Diese folgt aus Satz 2 und dem in Definition 7 miterwähnten Resultat, die wir nicht zeigen.

Zu Satz 5: Wenn \vec{x} und \vec{y} in $\text{Kern}(A)$ sind, also $A\vec{x} = A\vec{y} = \vec{0}$, so $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y} = \vec{0} + \vec{0} = \vec{0}$. Also ist $\vec{x} + \vec{y} \in \text{Kern}(A)$. Wenn $\vec{x} \in \text{Kern}(A)$, also $A\vec{x} = \vec{0}$, dann ist $A(\lambda\vec{x}) = \lambda A\vec{x} = \lambda\vec{0} = \vec{0}$. Also $\lambda\vec{x} \in \text{Kern}(A)$. Also ist $\text{Kern}(A)$ Unterraum.

Zu Satz 6: Nur für nicht leere Lösungsmengen ist etwas zu zeigen. Sei also \vec{x}_0 eine spezielle Lösung von $A\vec{x} = \vec{b}$. Wenn nun \vec{x}_0 die einzige Lösung ist, so behaupten wir: $\text{Kern}(A) = \{\vec{0}\}$. Sei nämlich $\vec{0} \neq \vec{y} \in \text{Kern}(A)$. Dann ist $A(\vec{x}_0 + \vec{y}) = A\vec{x}_0 + A\vec{y} = \vec{b} + \vec{0} = \vec{b}$. Also wäre $\vec{x}_0 + \vec{y} \neq \vec{x}_0$ eine weitere Lösung. Ist umgekehrt $\text{Kern}(A) = \{\vec{0}\}$, so kann keine Lösung außer \vec{x}_0 existieren. Sei nämlich $\vec{x}_1 \neq \vec{x}_0$ eine weitere. Dann gilt $A(\vec{x}_1 - \vec{x}_0) = A\vec{x}_1 - A\vec{x}_0 = \vec{b} - \vec{b} = \vec{0}$. Also hätten wir $\vec{0} \neq \vec{x}_1 - \vec{x}_0 \in \text{Kern}(A)$. Nun argumentiert man für den dritten Fall ganz ähnlich: Zunächst hat jede Lösung \vec{x}_1 von $A\vec{x} = \vec{b}$ eine Darstellung der Form $\vec{x}_0 + \vec{y}$, wobei $\vec{y} \in \text{Kern}(A)$. Denn wir haben soeben schon gezeigt, dass $\vec{x}_1 - \vec{x}_0 \in \text{Kern}(A)$ sein muss, setze nun $\vec{y} = \vec{x}_1 - \vec{x}_0$. Wir setzen unbewiesen die Existenz einer Basis $\vec{a}_1, \dots, \vec{a}_k$ mit $k \geq 1$ voraus. Damit ist jedes Element vom Kern als Linearkombination $\sum_{i=1}^k \lambda_i \vec{a}_i$ davon eindeutig darstellbar. Insgesamt haben wir dann: Jede weitere Lösung \vec{x}_1 ist in der Form $\vec{x}_0 + \sum_{i=1}^k \lambda_i \vec{a}_i$ darstellbar. (Diese Darstellung ist dann auch zwangsläufig eindeutig). Es verbleibt noch zu zeigen, dass umgekehrt jeder Vektor der Form $\vec{x}_0 + \sum_{i=1}^k \lambda_i \vec{a}_i$ eine Lösung ist. Da $\sum_{i=1}^k \lambda_i \vec{a}_i \in \text{Kern}(A)$, haben wir $A(\vec{x}_0 + \sum_{i=1}^k \lambda_i \vec{a}_i) = A\vec{x}_0 + A(\sum_{i=1}^k \lambda_i \vec{a}_i) = A\vec{x}_0 + \vec{0} = A\vec{x}_0 = \vec{b}$.

Es sei bemerkt, dass dieselben Sätze allgemeiner für jede lineare Abbildung von einem Vektorraum in einen anderen gelten, also nicht nur die Matrizen!

Systematisch fragen wir nicht nur nach Urbildern, sondern auch nach dem Bild der Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Das Bild eines einzelnen Vektors \vec{a} ist natürlich uninteressant, es ist als $A\vec{a}$ eindeutig auszurechnen. Aber welche Teilmenge von \mathbb{R}^m kommt heraus als Menge aller Bilder? Das ist sehr leicht zu sagen: Jeden Vektor $\vec{a} \in \mathbb{R}^n$ können wir in der Form $\sum_{j=1}^n \lambda_j \vec{e}_j$ darstellen, und wir beobachten: $A \sum_{j=1}^n \lambda_j \vec{e}_j = \sum_{j=1}^n \lambda_j A \vec{e}_j$, und $A \vec{e}_j$ ist nichts anderes als der j -te Spaltenvektor von A ! Jedes einzelne Bild ist eine Linearkombination der Spaltenvektoren von A . Die Summe zweier solcher Linearkombinationen ist wieder eine, ein Vielfaches einer solchen wieder eine. Also haben wir:

SATZ 7. *Sei A eine $(m \times n)$ -Matrix. Das Bild von A (kurz: $\text{Bild}(A)$), d.h. die Menge $\{A\vec{x} \mid \vec{x} \in \mathbb{R}^n\}$, ist ein Unterraum des \mathbb{R}^m , und es wird durch die Spaltenvektoren erzeugt. Daraus kann man mit Schaffen von Nullen ein maximales linear unabhängiges Teilsystem auswählen, das dann eine Basis des Bildes ist.*

Wir haben noch einen wichtigen Zusammenhang zwischen den Dimensionen, den wir nicht begründen, sondern nur in Beispielen illustrieren werden:

SATZ 8. *Dimension von $\text{Kern}(A)$ + Dimension von $\text{Bild}(A)$ = Dimension des Urbildraums.*

5.3.3. Die praktischen Folgerungen für lineare Gleichungssysteme. Eine erste ist bereits mit Satz 7 ausgesprochen: Jede Lösungsmenge eines beliebigen linearen Gleichungssystems lässt sich in linearer Parameterform darstellen. Das lehrte auch bereits die praktische Erfahrung mit dem Lösen solcher Systeme. Man beachte allerdings, dass noch so viele Beispiele keinen allgemeinen Beweis ersetzen! Darüber hinaus wissen wir: Wenn man ein System $A\vec{x} = \vec{b}$ schon gelöst hat (gleichgültig, ob $\vec{b} = \vec{0}$ oder nicht) und dabei k freie Parameter in der Lösungsmenge hatte, dann hat irgendein weiteres System $A\vec{x} = \vec{c}$ mit derselben Matrix, aber einem neuen Bildvektor, entweder gar keine Lösung oder aber *dieselbe* Zahl freier Parameter.

Wir fragen nunmehr danach, wie man einer Matrix schnell ansehen kann, ob die zugehörigen Systeme $A\vec{x} = \vec{b}$ stets (d.h. für jede Wahl von \vec{b}) genau eine Lösung / stets höchstens eine Lösung / stets mindestens eine Lösung haben. Folgender Satz gibt Auskunft darüber (beachten Sie zunächst die Anwendung mehr als einen Beweis):

SATZ 9.

- (1) $A\vec{x} = \vec{b}$ hat für **jedes** \vec{b} **mindestens eine** Lösung genau dann, wenn das System der **Zeilenvektoren** von A linear unabhängig ist.
- (2) $A\vec{x} = \vec{b}$ hat für **jedes** \vec{b} **höchstens eine** Lösung genau dann, wenn das System der **Spaltenvektoren** von A linear unabhängig ist.
- (3) $A\vec{x} = \vec{b}$ hat für **jedes** \vec{b} **genau eine** Lösung genau dann, wenn das System der Zeilenvektoren und das System der Spaltenvektoren von A linear unabhängig sind. Das ist jedoch bei quadratischen Matrizen bereits der Fall, wenn nur eine dieser Unabh.-Bedingungen erfüllt ist - die andere folgt dann.

Dieser Satz hat ein paar besonders übersichtliche Folgerungen (Eigenschaften, die unmittelbar an der bloßen Zahl der Zeilen und Spalten abzulesen sind):

1. Ist eine Matrix nicht quadratisch, so gibt es Vektoren \vec{b}_0 , so dass $A\vec{x} = \vec{b}_0$ unlösbar ist und auch solche, für die es mehr als eine Lösung gibt.
2. Ist $m > n$, die Zeilenzahl größer als die Spaltenzahl, so gibt es Vektoren \vec{b}_0 , für die $A\vec{x} = \vec{b}_0$ unlösbar ist.
3. Ist $n > m$, so gibt es Vektoren \vec{b}_0 , für die $A\vec{x} = \vec{b}_0$ mehr als eine Lösung hat.

Diese Folgerungen rühren einfach daher, dass es eben im \mathbb{R}^k keine längeren linear unabhängigen Systeme als von Länge k geben kann. Zum Satz selbst. Zur Bedingung dafür, dass es stets *höchstens eine* Lösung gibt, brauchen wir nur zu wiederholen, dass $A \sum_{j=1}^n \lambda_j \vec{e}_j = \sum_{j=1}^n \lambda_j A \vec{e}_j$ ist. Hätten wir zwei verschiedene Vektoren $\vec{x} = \sum_{j=1}^n \lambda_j \vec{e}_j \neq \sum_{j=1}^n \mu_j \vec{e}_j = \vec{y}$ mit $A \vec{x} = A \vec{y}$, so wären die Linearkombinationen verschieden, also für mindestens einen Index j_0 wäre $\lambda_{j_0} \neq \mu_{j_0}$, während $\sum_{j=1}^n \lambda_j A \vec{e}_j = \sum_{j=1}^n \mu_j A \vec{e}_j$. Das bedeutet: Die Spaltenvektoren wären linear abhängig. Umgekehrt folgt aus deren linearer Abhängigkeit, dass genau der beschriebene Fall eintritt. Damit ist die zweite Aussage komplett begründet. Zur ersten Aussage beobachten wir: Lineare Abhängigkeit der Zeilenvektoren bedeutet eine Beziehung der Form $\vec{a}_{i_0} = \sum_{i \neq i_0}^m \lambda_i \vec{a}_i$ (die Zeilenvektoren mit \vec{a}_i bezeichnet). Dann bilden wir folgenden speziellen Vektor \vec{b} (Spaltenvektor): Komponente zum Index i_0 sei 1, alle anderen Komponenten Null. Das Gleichungssystem $A \vec{x} = \vec{b}$ ist damit unlösbar, weil mit einer Lösung $A \vec{x}_1$ automatisch die links herauskommende Zahl in der Komponente i_0 Null sein müsste, auf der rechten Seite aber eine Eins verlangt ist. Umgekehrt wollen wir nur kurz bemerken: Wenn die Zeilenvektoren linear unabhängig sind, dann kann das Additionsverfahren keine linke Seite Null hervorbringen, es entsteht niemals ein Widerspruch im System, d.h. wir finden stets *mindestens eine* Lösung. Die dritte Aussage des Satzes fügt nun lediglich beide ersten zusammen mit „und“, ausgenommen die Zusatzbemerkung, dass bei quadratischen Systemen aus der linearen Unabhängigkeit der Spalten die der Zeilen folgt und umgekehrt. (Dasselbe gilt dann natürlich auch für die lineare Abhängigkeit.) Aber auch diese Zusatzbemerkung lässt sich relativ leicht zeigen.

Beispiele:

$$A_1 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

ist quadratisch. Da die Zeilen linear unabhängig sind, hat jedes Gleichungssystem mit dieser Matrix genau eine Lösung. Der Kern hat Dimension 0, das Bild Dimension 2.

$$A_2 = \begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix}$$

ist ebenfalls quadratisch, aber die Zeilenvektoren (und *folglich* auch die Spaltenvektoren) sind linear abhängig. Es gibt also zugehörige Systeme, die viele Lösungen haben, und solche, welche gar keine haben. Das weiß man ohne jede Rechnung! $A_1 \vec{x} = \vec{b}_0 = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$ hat keine Lösung, $A_1 \vec{x} = \vec{0}$ hat eine Lösungsmenge mit einem freien Parameter, der Kern hat also Dimension 1, und das Bild hat entsprechend Dimension 1. Es wird vom ersten Spaltenvektor erzeugt.

$$A_3 = \begin{pmatrix} 1 & 2 & 3 \\ -2 & 1 & 3 \end{pmatrix}$$

hat zwei linear unabhängige Zeilenvektoren. Also hat jedes zugehörige System *mindestens eine* Lösung. Das Bild muss zweidimensional sein, da etwa die ersten beiden Spalten linear unabhängig sind, die drei Spalten aber linear abhängig sein müssen als Vektoren im \mathbb{R}^2 . Daher wird das Bild etwa von den ersten beiden Spaltenvektoren erzeugt. Weitere Folgerung mit Dimensionssatz: Jede Lösungsmenge eines Systems der Form $A_3 \vec{x} = \vec{b}$ hat genau einen freien Parameter. (Mit $n > m$ war eindeutige Lösbarkeit bereits ausgeschlossen.) Das alles ohne jede Rechnung!

$$A_4 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 4 & 8 \end{pmatrix}$$

hat linear abhängige Zeilenvektoren, aber auch Spaltenvektoren. Das Bild wird erzeugt vom ersten Spaltenvektor und ist damit schmal im dreidimensionalen Werterraum. Also ist der Kern eindimensional. Für konkrete Systeme haben wir Unlösbarkeit zu erwarten, im Falle der Lösbarkeit gibt es jedoch eine unendliche eindimensionale

Lösungsmenge.

$$A_5 = \begin{pmatrix} 1 & 2 & 2 \\ -2 & 3 & 1 \\ -1 & 5 & 3 \end{pmatrix}$$

ist quadratisch. Aber die zugehörigen Systeme sind niemals eindeutig lösbar, entweder gibt es eine einparametrische Lösungsmenge, oder aber die leere Lösungsmenge. Das liegt daran, dass die Zeilenvektoren linear abhängig sind (bzw. die Spalten ebenfalls). Das Bild ist offenbar zweidimensional, daher der Kern und mit ihm jede nicht leere Lösungsmenge zugehöriger Systeme eindimensional. Der Spaltenvektor mit den Komponenten 0,0,1 ist jedenfalls nicht in $Bild(A)$.

$$A_6 = \begin{pmatrix} 1 & 2 & 2 \\ -2 & 3 & 1 \\ -1 & 5 & 6 \end{pmatrix}$$

dagegen ist nicht nur quadratisch, sondern mit Spaltenumformungen sieht man auch schnell, dass das System der Spaltenvektoren (und damit der Zeilenvektoren) linear unabhängig ist. Jedes mit dieser Matrix gebildete lineare Gleichungssystem ist also eindeutig lösbar.

Das praktische Fazit lautet: Hat man ein lineares Gleichungssystem vor sich, so achte man auf Zeilenzahl und Spaltenzahl sowie die lineare Unabhängigkeit bzw. Abhängigkeit der Zeilen und Spalten. Dann weiß man mit den angesprochenen Resultaten schon viel über das Lösungsverhalten, was dabei möglich und unmöglich ist, und zwar ohne bzw. mit sehr wenig Rechnung. Das alles, wie auch das praktische Lösungsverfahren selbst, ist natürlich nur bei recht kleinen Gleichungssystemen (wozu aber zweifellos (5×5) noch gehört) anwendbar. Für mehr helfen Computeralgebraprogramme und numerische Verfahren (es gibt durchaus sehr praktische Probleme, bei denen $m = n = 10.000$ ist!). Für deren Einsatz benötigt man jedoch auf jeden Fall die hier dargestellten Grundkenntnisse, dazu einiges mehr.

6. Betrag, Skalarprodukt, Vektorprodukt, Spatprodukt

6.1. Einführung von Betrag und Skalarprodukt. Wir haben die linearen Vektorraumoperationen einigermaßen kennengelernt und einiges damit schon angefangen - warum genügt das nicht? Einfach: Wir können noch fast gar nichts mit Längen und Winkeln rechnen. Lediglich „parallel - nicht parallel“ können wir von Vektoren sagen, andere Winkel nicht einmal rechnerisch unterscheiden. Auch können wir Längen nur bei parallelen Vektoren vergleichen: $2\vec{a}$ ist eben doppelt so lang wie \vec{a} , aber wie viel mal so lang wie \vec{b} ist \vec{a} , wenn beide nicht parallel sind? Dazu müssen wir eine weitergehende Struktur auf unsere Standard-Vektorräume bringen, die eines Euklidischen Raumes. Dazu setzen wir zunächst einmal voraus, dass alle Koordinatensysteme in diesem Abschnitt kartesisch seien. Nur dann entsprechen unsere Koordinaten-Rechnungen mit Längen und Winkeln denen der geometrischen Wirklichkeit, die wir beschreiben wollen. Wir beginnen mit der Länge von Vektoren, auch Betrag genannt. Das ist nicht nur das Einfachste, sondern wird das weitere auch begründen helfen.

DEFINITION 11. *Folgende Betragsfunktion auf \mathbb{R}^n wird definiert:*

$$\| \cdot \| : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sqrt{\sum_{i=1}^n x_i^2}$$

Motivierung: Die Länge eines Vektors kann bei kartesischen Koordinaten offenbar über Pythagoras gerade so ausgerechnet werden, wie man in Dimensionen 2,3 sieht. Für höhere Dimensionen genügt es, folgenden Eigenschaften, die man *allgemein beweisen* kann, anzusehen, dass sie offenbar das Wesentliche erfassen, was man von einem Längenbegriff erwartet:

SATZ 10. Die so definierte Betragsfunktion hat folgende Eigenschaften (allgemeingültige Formeln):

- (1) $|\vec{x}| \geq 0$
- (2) $|\vec{x}| = 0$ genau dann, wenn $\vec{x} = \vec{0}$
- (3) $|\lambda \vec{x}| = |\lambda| |\vec{x}|$, $|\lambda|$ ist der bei den reellen Zahlen.
- (4) $|\vec{x} + \vec{y}| \leq |\vec{x}| + |\vec{y}|$ (Dreiecksungleichung)

Die ersten drei sieht man sofort aus der Definition, die vierte beweist man zweckmäßig erst über das Skalarprodukt, das wir später einführen.

Praktische Beispiele:

Wie weit ist der Punkt P mit $\vec{x}_P^K = (1, 2, 3)$ (so in einem kartesischen System K dargestellt) vom Koordinatenursprung entfernt? Antwort: $\sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$. (Das schreiben wir ohne die unsäglichen „LE“ oder „Längeneinheiten“! Haben die Koordinaten inhaltlich die Interpretation „so viele Meter“, dann sind es eben so viele Meter.) Wie weit ist derselbe Punkt von Q entfernt, $\vec{x}_Q = (2, 3, 1)$? Antwort: Der Abstand ist $|\vec{x}_Q - \vec{x}_P| = \sqrt{5}$. Offenbar kann man jeden Vektor $\neq \vec{0}$ durch seinen Betrag dividieren, um gemäß (3) einen Vektor der Länge 1 von derselben Richtung zu produzieren.

Mit dem Betrag und dem anschaulichen Pythagorasatz kommt man sofort auf das nunmehr einzuführende Skalarprodukt: Genau dann, wenn \vec{a}, \vec{b} senkrecht aufeinander stehen, gilt die Gleichung $|\vec{b} - \vec{a}|^2 = |\vec{a}|^2 + |\vec{b}|^2$. Die linke Seite ist aber $|\vec{a}|^2 + |\vec{b}|^2 - 2 \sum_{i=1}^n a_i b_i$, wobei wir die Komponenten der Vektoren wie üblich bezeichnet haben. Also stehen \vec{a}, \vec{b} genau dann senkrecht aufeinander, wenn $\sum_{i=1}^n a_i b_i = 0$, und die nähere mathematische Analyse dieses Ausdrucks lohnt sich. Wir merken gleich noch an, dass wir sofort bemerken können, dass der Ausdruck einen Wert > 0 hat, wenn die Vektoren einen spitzen Winkel bilden, einen Wert < 0 , wenn die Vektoren einen stumpfen Winkel bilden. Schließlich wird uns die genaue Quantifikation von Winkeln damit gelingen.

DEFINITION 12. Das Euklidische Skalarprodukt auf \mathbb{R}^n ist folgende Abbildung (dabei bezeichnet $\mathbb{R}^n \times \mathbb{R}^n$ die Menge aller Paare von Vektoren aus \mathbb{R}^n):

$$\begin{aligned} \cdot : \quad \mathbb{R}^n \times \mathbb{R}^n &\rightarrow \mathbb{R} \\ \left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right) &\mapsto \sum_{i=1}^n x_i y_i \end{aligned}$$

Das Skalarprodukt ist also eine Abbildung mit zwei Eingabeschlitzen: In beide kommt je ein Vektor (gleicher Dimension!), heraus kommt ein Skalar, eine Zahl - man rechne also nur keinen Vektor aus, was häufig durch Vergessen der Summenbildung vorkommt. Daher „Skalarprodukt“! Das Resultat schreibt man gern $\vec{x} \vec{y}$, ohne Punkt. Auch \vec{x}^2 für $\vec{x} \vec{x}$. Auch von dieser Abbildung lassen sich leicht die wesentlichen algebraischen Eigenschaften zusammenfassen:

SATZ 11. Das Euklidische Skalarprodukt hat folgende wesentlichen algebraischen Eigenschaften (wiederum allgemeingültige Formeln):

- (1) $\left. \begin{aligned} \vec{x} (\vec{y} + \vec{z}) &= \vec{x} \vec{y} + \vec{x} \vec{z} \\ \vec{x} (\lambda \vec{y}) &= \lambda (\vec{x} \vec{y}) \end{aligned} \right\} \text{Linearität im zweiten Schlitz}$
- (2) $\vec{x} \vec{y} = \vec{y} \vec{x}$ (Symmetrie)
- (3) $\vec{x} \vec{x} = 0$ genau dann, wenn $\vec{x} = 0$ (positive Definitheit)

Die Symmetrie bewirkt mit (1) die Linearität auch im ersten Schlitz. Beide zusammen nennt man **Bilinearität**.

Die Eigenschaften sind unmittelbar über die Definition zu sehen und beruhen auf den entsprechenden Rechengesetzen (vor allem. Distributivgesetz) in den reellen Zahlen. Man beachte den Zusammenhang mit dem Betrag: $\vec{x}^2 = |\vec{x}|^2$. Die Gesetze bewirken, dass man im wesentlichen auch mit den nunmehr mit allen Operationen zu bildenden Rechenausdrücken *im wesentlichen* wie beim Zahlenrechnen umgehen kann. Aber es gibt eine wesentliche Ausnahme davon - Quelle vieler Anfängerfehler: *Ein Rechenausdruck wie $\vec{x} \vec{y} \vec{z}$ ist völlig sinnlos*. Hier braucht man die Klammern; denn $(\vec{x} \vec{y}) \vec{z}$ ist im allgemeinen verschieden von $\vec{x} (\vec{y} \vec{z})$! Man beachte, dass hier schon kategorial ein „Assoziativgesetz“ kein solches wäre - man hat es an einer Stelle mit der Multiplikation einer Zahl mit einem Vektor, an der anderen mit dem Skalarprodukt zu tun. Und vor allem wäre eine solche Regel falsch. Aus diesem Grunde verwendet man auch überhaupt keine höheren Exponenten als 2 im Zusammenhang mit dem Skalarprodukt! Schließlich vermeide man streng, durch Vektoren zu teilen - Umgang mit Beträgen und Skalarprodukten verführen dazu. Beispiele: Folgender Ausdruck ist für $\vec{a} \neq \vec{0}$ sinnvoll und die anschließende Rechnung korrekt - man beachte, dass man nicht durch Vektoren teilen oder kürzen darf, wohl aber durch Beträge, die verschieden von Null sind!

$$\left| \frac{\vec{a} \vec{b}}{\vec{a}^2} \vec{a} \right| = \left| \frac{\vec{a} \vec{b}}{|\vec{a}|^2} \right| |\vec{a}| = \frac{|\vec{a} \vec{b}|}{|\vec{a}|^2} |\vec{a}| = \frac{|\vec{a} \vec{b}|}{|\vec{a}|}.$$

Hier ist eine wichtige logische Falle für Anfänger zu erwähnen: Manche sagen hier so etwas wie: „Warum so umständlich - wenn ich durch den Vektor kürze, so kommt doch dasselbe heraus, nur viel einfacher!“ Dazu ist schlicht zu bemerken, dass das Unfug ist; denn selbstverständlich kann man aus Falschem etwas Wahres herleiten, man kann daraus sogar *alles* herleiten, jede falsche wie jede wahre Aussage. Es kommt daher sehr wohl auf *korrekte* Begründungen an. Machen wir das am selben Beispiel ganz klar: Sollte man mit dem Kürzungsvorschlag dann nicht konsequent zwei mal durch den Vektor kürzen dürfen? Dann hätte man aber als Vereinfachung den Ausdruck $\left| \frac{\vec{a} \vec{b}}{\vec{a}} \right|$. Und dies Resultat ist völlig falsch, da obenstehender Ausdruck mit \vec{a} senkrecht zu \vec{b} Null ergibt, auch mit $\left| \vec{b} \right| \neq 0$. Noch ein Beispiel: $(\vec{a} \vec{b})^2$ ist nicht etwa dasselbe wie $\vec{a}^2 \vec{b}^2$, beide Ausdrücke sind aber offensichtlich sinnvoll. Zu Beginn sollte man alle Ausdrücke sehr bedächtig lesen und analysieren wie „Vektor mal Vektor gibt eine Zahl, diese mal Vektor einen Vektor, der Betrag davon wieder eine Zahl“ usw.

6.2. Zwei grundlegende Anwendungen des Skalarproduktes.

6.2.1. *Erste Anwendung: Senkrechte Projektion.* Der Vektor \vec{b} soll zerlegt werden in eine (Vektor-) Komponente parallel zu \vec{a} und eine senkrecht zu \vec{a} . Wir zeigen, dass dies stets eindeutig möglich ist und nennen die Komponente von \vec{b} parallel zu \vec{a} die senkrechte Projektion von \vec{b} auf \vec{a} . Sie kann durch das Skalarprodukt folgendermaßen ausgedrückt werden:

$$\left(\text{Senkrechte Projektion von } \vec{b} \text{ auf } \vec{a}\right) = \frac{\vec{a} \vec{b}}{a^2} \vec{a}.$$

Diesen Ausdruck findet man durch den Ansatz $\lambda \vec{a} + \vec{l} = \vec{b}$, $\vec{l} \vec{a} = 0$. Skalares Anmultiplizieren an die erste (Vektor-) Gleichung ergibt $\lambda = \vec{a} \vec{b} / a^2$. Dies ist eine wichtige Rechentechnik, sie macht aus Vektorgleichungen Zahlgleichungen, und insbesondere fallen unliebsame Unbekannte heraus, wenn man entsprechende Bedingungen in Form von Beziehungen „... senkrecht auf ...“ hat. In unserem Falle fiel daher \vec{l} , der unbekannte Vektor, heraus. Nunmehr ist es leicht, auch diesen auszurechnen:

$$\text{Lotvektor } \vec{l} \text{ bei senkrechter Projektion von } \vec{b} \text{ auf } \vec{a}: \vec{l} = \vec{b} - \frac{\vec{a} \vec{b}}{a^2} \vec{a}.$$

6.2.2. *Cosinus des Winkels zwischen zwei Vektoren.* Die zweite Anwendung ergibt sich als Folgerung der Formel für die senkrechte Projektion:

$$\frac{\vec{a} \vec{b}}{|\vec{a}| |\vec{b}|} \text{ ist der Cosinus des Winkels zwischen } \vec{a}, \vec{b} \neq \vec{0}.$$

Dies ergibt unmittelbar den Cosinussatz, vektoriell formuliert: Sei φ der Winkel zwischen den Vektoren $\vec{a}, \vec{b} \neq \vec{0}$. Dann gilt:

$$|\vec{a} - \vec{b}|^2 = |\vec{a}|^2 + |\vec{b}|^2 - 2 \cos(\varphi) |\vec{a}| |\vec{b}|.$$

Das ist nämlich nur eine winzige Umformulierung des oben bereits festgestellten Sachverhaltes $|\vec{a} - \vec{b}|^2 = |\vec{a}|^2 + |\vec{b}|^2 - 2 \vec{a} \vec{b}$, da $\cos(\varphi) |\vec{a}| |\vec{b}| = \vec{a} \vec{b}$ gilt.

6.3. Weitere Anwendungen.

6.3.1. *Die Cauchy-Schwarzsche Ungleichung und die Dreiecksungleichung für Beträge, dazu eine schöne Formel für den Flächeninhalt eines Parallelogramms.* Wir beweisen zunächst folgende fundamentale Ungleichung:

SATZ 12 (Cauchy-Schwarzsche Ungleichung):

$$(1) \quad |\vec{a} \vec{b}| \leq |\vec{a}| |\vec{b}|.$$

Zusatz: Das Gleichheitszeichen trifft nur im dem Falle zu, dass beide Vektoren linear abhängig sind (wozu auch der Fall gehört, dass mindestens einer der Nullvektor ist).

Beweis: Für $\vec{a} = \vec{0}$ die Ungleichung mit Gleichheitszeichen klar. Für $\vec{a} \neq \vec{0}$ haben wir die obenstehende Zerlegung

$$\vec{b} = \frac{\vec{a} \vec{b}}{a^2} \vec{a} + \vec{l}, \quad \vec{l} \text{ senkrecht zu } \vec{a}.$$

Das ergibt (auch für den Sonderfall $\vec{l} = \vec{0}$) die Pythagorasbeziehung - man erinnere sich, dass stets $\vec{x}^2 = |\vec{x}|^2$,

$$\left(\frac{\vec{a} \cdot \vec{b}}{a^2} \vec{a}\right)^2 + \vec{l}^2 = \vec{b}^2, \text{ also (2) } (\vec{a} \cdot \vec{b})^2 + \vec{l}^2 a^2 = a^2 \vec{b}^2.$$

Da nun $\vec{l}^2 a^2 \geq 0$, haben wir $(\vec{a} \cdot \vec{b})^2 \leq a^2 \vec{b}^2$. Wurzelziehen ergibt die zu beweisende Ungleichung (1).

Mehr noch: Das Gleichheitszeichen resultiert genau in dem Falle, dass $\vec{l} = \vec{0}$, d.h. aber, dass \vec{a} parallel zu \vec{b} ist. Schließlich beobachten wir, dass $\vec{l}^2 a^2$ nichts anderes als das Quadrat des Flächeninhaltes von dem durch \vec{a} , \vec{b} aufgespannten Parallelogramm ist. Damit bedeutet (2) zugleich:

(3) Das von \vec{a} , \vec{b} aufgespannte Parallelogramm hat den Flächeninhalt

$$a^2 \vec{b}^2 - (\vec{a} \cdot \vec{b})^2.$$

Beachten Sie hier noch einmal, wie falsch es wäre, $a^2 \vec{b}^2$ mit $(\vec{a} \cdot \vec{b})^2$ zu identifizieren. Dann hätte jedes Parallelogramm Flächeninhalt Null!

6.3.2. *Die Normalenform für Ebenen im E^3 und Verallgemeinerungen.* Im dreidimensionalen Raum kann man offenbar eine Ebene E bestimmen durch folgende beiden Bedingungen: E steht senkrecht auf $\vec{n} \neq \vec{0}$, und P liegt auf E . Man nennt \vec{n} dann *einen* (nicht: *den*, weil er nicht eindeutig bestimmt ist) Normalenvektor für E . Wir werden sehen, dass dies unmittelbar die uns schon bekannte Gleichungsform für Ebenen ergibt. Umgekehrt können wir dann diese Gleichungsform *unmittelbar* geometrisch deuten, was wir bisher nicht konnten. Setzen wir also voraus, dass E die oben formulierten Eigenschaften habe. Aus $\vec{n} \perp E$ folgt unmittelbar, dass jeder Vektor $\vec{x}_Q - \vec{x}_P$ *genau dann* senkrecht zu \vec{n} steht, wenn $Q \in E$. Das bedeutet aber, dass die Gleichung $\vec{n} \cdot \vec{x} = \vec{n} \cdot \vec{x}_P$ als Lösungsmenge *genau* die Menge der Ortsvektoren von allen Punkten auf E besitzt. Somit haben wir für E die Bestimmungsgleichung (noch einmal: $\vec{n} \neq \vec{0}$!): Eine Normalenform für die Ebene E senkrecht zu \vec{n} , welche durch P geht, lautet

$$\vec{n} \cdot \vec{x} = \vec{n} \cdot \vec{x}_P.$$

In Koordinatenform ergibt das
$$\begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = d, \text{ mit } d = \vec{n} \cdot \vec{x}_P.$$

$$\text{Ausgerechnete Form dazu ist } n_1 x + n_2 y + n_3 z = d.$$

Umgekehrt können wir nun bei vorgelegter Gleichungsform für eine Ebene wie etwa $2x - 3y + 4z = 2$ sofort sagen, dass der Vektor $(2, -3, 4)$ senkrecht auf E steht, also ein Normalenvektor dafür ist.

Die angekündigte Verallgemeinerung besteht nun darin: Eine Bestimmungsgleichung der Form ($k \geq 2$) $\sum_{i=1}^k n_i x_i = d$ mit Unbestimmten x_i und Koeffizienten n_i , die nicht alle Null sind, hat einen $n - 1$ -dimensionalen affinen Unterraum des \mathbb{R}^n als Lösungsmenge, und einen solchen nennt man Hyperebene des \mathbb{R}^k . Der Vektor $\vec{n} = (n_1, \dots, n_k)$ steht dann wiederum senkrecht auf dieser Hyperebene und ist also Normale dazu, und die Richtung der Hyperebene ist durch diesen Vektor eindeutig bestimmt. (Oder auch wieder in geometrischer koordinatenfreier Form: $\vec{n} \cdot \vec{x} = d$, mit k -dimensionalen Vektoren, die nun auch aus V^k sein dürfen.) Illustrierende Beispiele: Im zweidimensionalen Raum haben wir die schon bekannte Normalenform für *Geraden*: $ax + by = d$, a, b nicht beide Null. (a, b) steht dann senkrecht auf der betreffenden Geraden. Für den vierdimensionalen Raum haben wir dreidimensionale Hyperebenen. Es wurde schon klar, dass man nicht etwa für Geraden im Dreidimensionalen eine solche Form erwarten kann, sondern dann ein System von zwei solchen Gleichungen braucht, was sich natürlich in höheren Dimensionen stark verallgemeinert.

6.4. Das Vektorprodukt (nur im V^3 bzw. \mathbb{R}^3). Wenn wir eine Ebene E etwa in Parameterform erhalten haben und wollen (etwa für Schnittaufgaben) eine Normalenform daraus gewinnen, so benötigen wir einen Vektor, der zu den gegebenen beiden Richtungsvektoren (welche *parallel* zu E liegen) senkrecht steht. Das könnten wir als ein lineares Gleichungssystem formulieren und finden, dass die Lösungsmenge einen freien Parameter hat. Aus der Lösungsmenge wäre dann ein Element $\neq \vec{0}$ brauchbar. Viel bequemer ist es jedoch, einen Normalenvektor als Vektorprodukt der Richtungsvektoren auszuwählen. Aber das ist keineswegs der einzige Nutzen des Vektorprodukts, viel größere Bedeutung hat es in der Physik für die Berechnung von Drehmomenten oder Lorentzkräften, Geschwindigkeitsvektoren bei starren rotierenden Körpern, usw. Bei diesen Anwendungen benötigt man die weiteren Eigenschaften des Vektorproduktes, welche dies eindeutig machen (im Gegensatz zu einem Normalenvektor einer Ebene).

DEFINITION 13 (geometrische Charakterisierung des Vektorprodukts). *Das Vektorprodukt zweier Vektoren $\vec{a}, \vec{b} \in V^3$ ist der eindeutig bestimmte Vektor $\vec{a} \times \vec{b}$ mit folgenden Eigenschaften:*

- (1) $\vec{a} \times \vec{b}$ steht senkrecht auf \vec{a} und auf \vec{b} .
- (2) $|\vec{a} \times \vec{b}|$ ist der Flächeninhalt des von beiden Vektoren erzeugten Parallelogramms.
- (3) Die Vektoren $\vec{a}, \vec{b}, \vec{a} \times \vec{b}$ ergeben in dieser Reihenfolge ein Rechtssystem.

DEFINITION 14 (algebraische Charakterisierung des Vektorprodukts). *Das Vektorprodukt \times ist eine Abbildung auf $V^3 \times V^3 \rightarrow V^3$ bzw. $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ mit folgenden Eigenschaften (dabei sei $\vec{e}_i, 1 \leq i \leq 3$, für V^3 ein kartesisches System bzw. für \mathbb{R}^3 die kanonische Basis von Einheitsvektoren, welche definitiv als solches bezeichnet wird, also paarweise senkrecht, alle gleich lang und in dieser Reihenfolge ein Rechtssystem bildend):*

- (1)
$$\left. \begin{aligned} \vec{a} \times (\vec{b} + \vec{c}) &= \vec{a} \times \vec{b} + \vec{a} \times \vec{c} \\ \vec{a} \times (\lambda \vec{b}) &= \lambda (\vec{a} \times \vec{b}) \end{aligned} \right\} \text{Linearität im 2. Schlitz}$$
- (2) $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$ Antikommutativität
- (3) $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3, \vec{e}_2 \times \vec{e}_3 = \vec{e}_1, \vec{e}_3 \times \vec{e}_1 = \vec{e}_2$. (Normierung.)

Aus der Eigenschaft (2) folgt mit (1), dass die Abbildung auch im 1. Schlitz linear ist, also bilinear wie das Skalarprodukt. Aus der Eigenschaft (3) folgt mit (2), dass $\vec{e}_2 \times \vec{e}_1 = -\vec{e}_3$ usw. Ferner folgen diese Eigenschaften aus den geometrischen elementargeometrisch (nicht formal zu beweisen, da E^3 hier nur anschaulicher Raum ist). Nun werden wir aus den algebraischen Eigenschaften zeigen, wie das Vektorprodukt eindeutig auszurechnen ist. Damit zeigen wir, dass die algebraischen Eigenschaften das Vektorprodukt eindeutig charakterisieren, wichtiger jedoch ist damit bewerkstelligte Herleitung der folgenden merkwürdigen Rechenformel für das Vektorprodukt: Jeder Vektor aus V^3 (bzw. \mathbb{R}^3) lässt sich eindeutig als Linearkombination der \vec{e}_i darstellen, da diese eine Basis bilden. Wir setzen also zwei Vektoren so an und rechnen *allein mittels* der algebraischen Eigenschaften das Vektorprodukt aus:

$$\begin{aligned} & (a_1 \vec{e}_1 + a_2 \vec{e}_2 + a_3 \vec{e}_3) \times (b_1 \vec{e}_1 + b_2 \vec{e}_2 + b_3 \vec{e}_3) = \\ & a_1 b_2 \vec{e}_1 \times \vec{e}_2 + a_1 b_3 \vec{e}_1 \times \vec{e}_3 + a_2 b_1 \vec{e}_2 \times \vec{e}_1 + a_2 b_3 \vec{e}_2 \times \vec{e}_3 \\ & \quad + a_3 b_1 \vec{e}_3 \times \vec{e}_1 + a_3 b_2 \vec{e}_3 \times \vec{e}_2 \\ & = (a_1 b_2 - a_2 b_1) \vec{e}_3 - (a_1 b_3 - a_3 b_1) \vec{e}_2 + (a_2 b_3 - a_3 b_2) \vec{e}_1. \end{aligned}$$

Im ersten Schritt haben wir das aus (1) folgende Distributivgesetz verwandt sowie die Faktoren herausgezogen. Ferner wurden dabei alle Glieder der Form $\dots \vec{e}_i \times \vec{e}_i$ weggelassen, da aus (2) folgt: $\vec{a} \times \vec{a} = \vec{0}$. Im zweiten Schritt wurde dann (3) benutzt und mit linearen Operationen zusammengefasst. Die Endform gibt die

Koordinatendarstellung für das Vektorprodukt aus denen der eingehenden Vektoren. Man beachte das Schema, insbesondere das Minuszeichen für die zweite Komponente unter Beibehalten des Indexschemas. Man kann als Koordinaten-Rechenregel auch so zusammenfassen: Man schreibe die beiden Koordinaten-Spaltenvektoren nebeneinander. Nun rechne man die i -te Komponente aus als Determinante der (2×2) -Matrix, die durch Streichen der i -ten Zeile entsteht. *Allerdings* bringe man bei der Berechnung der zweiten Komponente noch einen Faktor -1 an. Beispiel:

$$\begin{pmatrix} -2 \\ 3 \\ 4 \end{pmatrix} \times \begin{pmatrix} -1 \\ -3 \\ 5 \end{pmatrix} = \begin{pmatrix} 27 \\ 6 \\ 9 \end{pmatrix}.$$

Stets kontrolliere man ein ausgerechnetes Vektorprodukt daraufhin, ob es senkrecht zu den Ausgangsvektoren steht. Ist das der Fall, so kann man sicher sein, dass das so ausgerechnete Resultat stimmt.

6.4.1. *Konkrete Abstands- und Winkelprobleme.* Wir stellen kurz das Praktische zusammen: Den Abstand zwischen einem Punkt und einer Geraden können wir durch Minimalisieren des Ausdrucks $|x_P - \vec{x}_g(\lambda)|$ oder auch mittels eines geeigneten Lotvektors bekommen. Den Abstand zwischen zwei parallelen Ebenen berechnet man leicht über Normalenformen $\vec{n} \cdot \vec{x} = d_1$, $\vec{n} \cdot \vec{x} = d_2$ (sollten die Normalenvektoren nicht gleich sein, so bringe man zuvor einen geeigneten Faktor an einer oder beiden Gleichungen an) als $|d_1 - d_2| / |\vec{n}|$. Den Abstand zwischen zwei windschiefen Geraden (man braucht allerdings nur linear unabhängige Richtungsvektoren \vec{a} , \vec{b} für beide Geraden) kann man (mit Aufpunktvektoren \vec{x}_P , \vec{x}_Q für die beiden Geraden) berechnen durch $\left| (\vec{x}_P - \vec{x}_Q) \cdot \vec{a} \times \vec{b} \right| / \left| \vec{a} \times \vec{b} \right|$.

Winkelfragen: Entscheidend ist immer wieder, dass man eine solche Frage als Frage nach dem Winkel zwischen zwei Vektoren umformuliert. Standardbeispiele: Der Winkel zwischen zwei Ebenen ist als Winkel zwischen ihren Normalenvektoren auszurechnen. Man beachte, dass es zwei Winkel gibt, die sich zu 180 Grad ergänzen. Der Winkel zwischen einer Geraden und einer Ebene ist als Arcussinus des Cosinuswertes (ja, so!) vom Winkel zwischen Geraden-Richtungsvektor und Ebenen-Normalenvektor auszurechnen. Das kommt korrekt heraus, gleichgültig, wie diese Vektoren gewählt wurden. Wiederum gibt es zwei Winkel, die einander zu 180 Grad ergänzen.

6.5. Das Spatprodukt und die Determinante. Ein Spat ist eine Verallgemeinerung des Parallelogramms auf drei oder mehr Dimensionen. Wir behandeln die Frage des orientierten Volumens nicht allgemein für jede Dimension (was allerdings mittels der Determinante gelingt, deren Rechenvorschrift wir auch allgemein vorstellen), sondern konzentrieren uns auf den dreidimensionalen Fall. Es ist ein Körper, der mittels dreier Vektoren zu parametrisieren ist durch $\vec{x}(\alpha, \beta, \gamma) = \alpha \vec{a} + \beta \vec{b} + \gamma \vec{c}$, $\alpha, \beta, \gamma \in [0, 1]$. Im Falle, dass die drei Vektoren kein linear unabhängiges System bilden, entsteht damit ein sog. *ausgeartetes* Spat, dessen (dreidimensionales!) Volumen dann Null ist. Wir definieren nun die folgende Abbildung, welche in allen drei Schlitzen linear ist:

DEFINITION 15 (Spatprodukt). $[\vec{a}, \vec{b}, \vec{c}] = \vec{a} \cdot (\vec{b} \times \vec{c})$ (mit Skalarprodukt und Vektorprodukt gebildet, daher nur im dreidimensionalen Fall!)

Bemerkung: Man kann es ruhig der Bezeichnung mit dem Rechenausdruck $\vec{a} \cdot (\vec{b} \times \vec{c})$ bewenden lassen.

SATZ 13. *Das Spatprodukt ist in allen drei Eingabeschlitzen linear. Das Spatprodukt $[\vec{a}, \vec{b}, \vec{c}]$ liefert das dreidimensionale Volumen des von \vec{a} , \vec{b} , \vec{c} aufgespannten Spates mit Orientierungsvorzeichen, d.h. für ein Linkssystem steht ein negatives Vorzeichen vor dem Volumen. Insbesondere gilt $[\vec{a}, \vec{b}, \vec{c}] = 0$ genau dann, wenn das System der drei Vektoren linear unabhängig ist. Bei zyklischer Vertauschung (jeweils kommt der Letzte nach vorn) der Eingabevektoren bleibt das Spatprodukt unverändert, bei antizyklischer Vertauschung (antizyklisch sind alle nicht zyklischen Vertauschungen) ändert es sein Vorzeichen. Ferner gilt für eine kartesische Basis $\vec{e}_1, \vec{e}_2, \vec{e}_3 : [\vec{e}_1, \vec{e}_2, \vec{e}_3] = 1$.*

Die genannten *algebraischen* Eigenschaften allein bedeuten, dass es sich um eine alternierende Trilinearform handelt. Die Linearität in jedem Schlitz ergibt sich unmittelbar aus der Bilinearität einerseits des Skalarproduktes, andererseits des Vektorproduktes. Also praktisch: Distributives Rechnen und Herausziehen von Faktoren. Das Alternieren (Verhalten bei Vertauschungen) schließt insbesondere ein, dass stets z.B. $[\vec{a}, \vec{a}, \vec{b}] = 0$ usw. immer, wenn zwei der drei Vektoren identisch sind. (Tatsächlich kann man dies zu Definition von „alternierend“ machen und mittels der Trilinearität auf die oben angegebenen Vertauschungseigenschaften schließen.) Weiter nutze man stets das Vertauschungsverhalten für zusammenfassende Rechnungen wie

$$\vec{a} (\vec{b} \times \vec{c}) + 2 \vec{b} (\vec{a} \times \vec{c}) = \vec{a} (\vec{b} \times \vec{c}) - 2 \vec{a} (\vec{b} \times \vec{c}) = -\vec{a} (\vec{b} \times \vec{c}).$$

Zusammenhang mit der Determinante von (3×3) - Matrizen: Man hat in Komponenten ausgeschrieben:

$$\begin{aligned} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \left(\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \times \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \right) &= \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \begin{pmatrix} b_2 c_3 - b_3 c_2 \\ -(b_1 c_3 - b_3 c_1) \\ b_1 c_2 - b_2 c_1 \end{pmatrix} \\ &= a_1 (b_2 c_3 - b_3 c_2) + (-a_2) (b_1 c_3 - b_3 c_1) + a_3 (b_1 c_2 - b_2 c_1) . \end{aligned}$$

Das ist aber gerade die Berechnung der Determinante von

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

durch sogenannte Entwicklung nach der ersten Spalte: Nach Vorzeichenschema kommt $+a_1, -a_2, +a_3$, jeweils mal Unterdeterminante, welche nach Streichen der zu a_k gehörigen Spalte und Zeile entsteht. Eine (2×2) -Determinante berechnet sich einfach als Produkt der Hauptdiagonalen minus produkt der Nebendiagonalen, also

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

In der Koordinatenform sind also Determinante von (3×3) - Matrizen und Spatprodukt *dasselbe* und haben daher auch dieselbe Deutung. Das Wesentliche ist „multilinear alternierend“, und das kann man auf alle Dimensionen übertragen. Daher gibt es die Determinante in allen Dimensionen k , und sie hat stets die Eigenschaft, das k - dimensionale Volumen von Spaten mit Orientierungsvorzeichen zu liefern.

Die komplexen Zahlen

Dies Kapitel ist weitaus kürzer als die beiden anderen, gewichtigeren Hauptkapitel. Es bietet sich indessen auch als Überleitung von der Vektorrechnung zur Infinitesimalrechnung an, weil einerseits die komplexen Zahlen *auch und insbesondere* Vektoren sind, gewisse Elemente der Vektorrechnung also das Verständnis befördern und zugleich wiederholt werden. Andererseits kommen in natürlicher Weise Elemente der reellen hinein, insbesondere trigonometrische Funktionen, auch bereits die komplexe Exponentialfunktion. Die komplexen Zahlen bieten daher willkommenen Anlass, mit diesen Funktionen elementar umzugehen und reichen damit in den ersten Teil des Kapitels über die Funktionen hinein, der ein gewisses Grundverständnis sowie Grundkenntnisse wichtiger Funktionen bereitstellen soll, noch vor dem Differenzieren und Integrieren.

1. Der Körper \mathbb{C} der komplexen Zahlen

2. Motivierung der komplexen Zahlen

Man geht von \mathbb{N} zu \mathbb{Z} , um stets zu eine Zahl die negative bilden und stets die Subtraktion zweier Zahlen ausführen zu können. Dann erweitert man \mathbb{Z} zu \mathbb{Q} , der Menge der rationalen Zahlen, um die analogen Eigenschaften für die Multiplikation zu bekommen, d.h. für eine von Null verschiedene Zahl a stets die Inverse bezüglich der Multiplikation zu bekommen und jede Zahl durch a dividieren zu können. Zwei Umstände sind verantwortlich für die Nützlichkeit dieser Erweiterungen, einmal die praktische Anwendbarkeit der dadurch ermöglichten Operationen, zum andern die Erhaltung der Rechenstrukturen bzw. Rechengesetze bei der jeweiligen Erweiterung. \mathbb{Q} erscheint nunmehr für das Rechnen irgendwie perfekt und ist es auch in seiner Weise. Was man sinngemäß „bürgerliches Rechnen“ nennen könnte, kommt damit vollständig aus. Warum diese äußerst schwierige Übergang zu \mathbb{R} ? Das ist eine gewaltige Erweiterung, die motiviert wurde durch das Bestreben, einen Grenzwert (Limes) für alle solche Zahlenfolgen zu finden, die sich zu einem Punkt zusammenziehen - und in \mathbb{Q} nur auf ein Loch stoßen würden. (Das gilt für fast alle!) Ursprünglich hatte man die Idee des Kontinuums, eben ohne Löcher, so dass man bei stetiger Veränderung einer abhängigen Veränderlichen stets so etwas wie ein Maximum bzw. Minimum finde, wenn man den Definitionsbereich auf ein endliches Intervall beschränkt. Und die (beweisbar!) ökonomischste Weise, diese Eigenschaft herzustellen, ist die Einführung der ganzen Menge \mathbb{R} , eine Entwicklung, die im 17. Jahrhundert begann und erst Ende des 19. Jahrhunderts vollendet war. Dass \mathbb{R} auch im algebraischen Sinne bereits recht „vollständig“ ist, erkennt man daran, dass jedes reelle Polynom sich bereits (bis auf einen konstanten Faktor) über \mathbb{R} in lauter Linearfaktoren der Form $(x - a)$ und quadratische Faktoren zerlegen lässt. Insbesondere besitzt jedes reelle Polynom, das einen Vorzeichenwechsel bei seinen Werten hat, auch mindestens eine reelle Nullstelle. Z.B. weiß man das sofort für $x^{11} - 30x^{10} + 5x^8 - 3x^3 + 2x^2 + 1$, ohne eine solche exakt ausrechnen zu können. Insbesondere kann man Wurzeln wie $\sqrt{2}$, $\sqrt{3}$ usw. bilden in \mathbb{R} bilden, nicht aber in \mathbb{Q} , was ereits im Altertum bekannt und bewiesen(!) war. Aber \mathbb{R} ist doch immer noch nicht algebraisch vollständig in dem Sinne, dass Polynome *alle* Nullstellen haben, die sie haben sollten, so dass man sie vollständig in Linearfaktoren zerlegen kann. Z.B. hat $x^2 + 1 = 0$ keine Lösung in \mathbb{R} , so dass man darin nicht $\sqrt{-1}$ bilden kann. Lauter quadratische Gleichungen ohne reelle Nullstelle entstehen mit den quadratischen Polynomen, bei denen man nur Werte > 0 oder nur solche < 0 hat. Diese Polynome gerade lassen sich über \mathbb{R} nicht mehr in Linearfaktoren zerlegen. Bewusst wurde das Bedürfnis beim formalen Rechnen mit der Lösungsformel für quadratische Gleichungen, und man sprach von „imaginären“ (bloß gedachten, denen

nichts Wirkliches entsprach) Zahlen wie $2 + 3\sqrt{-1}$. Dabei merkte man schnell, dass allein $\sqrt{-1}$ nötig war: Jede *formale* Lösung einer quadratischen Gleichung ohne reelle Lösung hat die Gestalt $a \pm b\sqrt{-1}$, Beispiel: $2 + \frac{1}{2}\sqrt{-5} = 2 + \frac{5}{2}\sqrt{-1}$. Führt man die imaginäre Zahl j (sie heißt traditionell in der Mathematik i , doch ist dieser Buchstabe für die Stromstärke in der Technik verbraucht, so dass die Techniker „ j “ vorziehen) ein mit der Forderung $j^2 = -1$, so kann man wie gewöhnlich mit diesen formalen Lösungen rechnen, sie addieren und multiplizieren, gerade so, wie man eben auch mit Ausdrücken wie $2 + \sqrt{3}$ exakt umgeht, nur eben beachtend, dass wegen $j^2 = -1$ konsequent folgt: $(a + jb)(c + jd) = ac - bd + j(ad + bc)$, Endform, wieder ein Ausdruck der Gestalt $\alpha + \beta j$, mit reellen Zahlen α, β , wenn a, b, c, d reell waren. Zwei Fragen stellten sich mathematisch konsequent: Kann man die reellen Zahlen so erweitern, dass eine solche Zahl j mit $j^2 = -1$ in der erweiterten Menge ist, die Addition und Multiplikation (einschließlich der Inversenbildungen) darin so durchführbar sind, dass alle wesentlichen Rechengesetze (d.h. die Körperaxiome, siehe Kap. 1) weiter gelten? Die zweite Frage: Kann man auf gesicherter Basis ein Modell widerspruchsfrei konstruieren, also die komplexen Zahlen ihres „bloß formalen“ oder gar nebulös-„imaginären“ Charakters entheben? Beide Fragen wurden im 19. Jahrhundert positiv beantwortet, unter wesentlicher Mitwirkung von Gauß. Mehr noch wurde auch schon bewiesen (Liouville), dass der Körper der komplexen Zahlen wirklich algebraisch vollständig (man sagt: „algebraisch abgeschlossen“) ist in dem Sinne, dass jedes nichtkonstante Polynom mit komplexen (!) Koeffizienten tatsächlich über \mathbb{C} in Linearfaktoren zerfällt. Mittlerweile kennt man dafür mehrere äußerst verschiedene Beweise, doch ist keiner so einfach, dass wir ihn hier bringen könnten. Hingegen ist die Konstruktion von \mathbb{C} mit der Verifikation der Rechengesetze für Körper sehr leicht (wenn einmal die Gaußsche Idee da ist), sie folgt nun.

3. Konstruktion von \mathbb{C}

Wir haben zunächst die Menge zu definieren, anschließend die algebraischen Operationen Addition und Multiplikation, dann ist einzusehen, dass die Körperaxiome folgen. Bei der Gelegenheit lernen wir die Bedeutung der komplexen Konjugation für das Rechnen mit komplexen Zahlen kennen. Hinzu kommt die sehr wichtige geometrische Interpretation, die wir im Rahmen der Vektorrechnung leicht durchführen können.

DEFINITION 16. *Der Körper \mathbb{C} der komplexen Zahlen ist die Menge \mathbb{R}^2 der reellen Zahlenpaare versehen mit folgender Addition und Multiplikation:*

Bezeichnung. Das Element $(a, b) \in \mathbb{R}^2$ wird geschrieben: $z = a + jb$, wenn man es als komplexe Zahl auffasst.

Dabei heisst a Realteil von z , $a = \operatorname{Re}(z)$, b Imaginärteil von z , $b = \operatorname{Im}(z)$.

Addition: $(a + jb) + (c + jd) = (a + c) + j(b + d)$, also wie im \mathbb{R}^2 .

Multiplikation: $(a + jb)(c + jd) = (ac - bd) + j(ad + bc)$, neu!

Bemerkung: Die Multiplikation ist neu, sie ist weder das Skalarprodukt, noch hat sie das Geringste mit dem Vektorprodukt zu tun (schon aus Dimensionsgründen). Sie gehört nicht zur Vektorraumstruktur des \mathbb{R}^2 , sondern eben zur Körperstruktur \mathbb{C} . Daher ist es günstig, die Elemente in der Form $a + jb$ und nicht (a, b) zu schreiben: Wir rechnen dann, wie wir es von Ausdrücken wie $a + b\sqrt{2}$ gewöhnt sind, beachten nur, dass $j^2 = -1$ - beachten Sie, dass dies durch Einsetzen in die Definition der Multiplikation folgt, dass aber auch gerade in diesem Sinne die Multiplikation definiert ist. Insbesondere ist an der Schreibweise praktisch, dass man naheliegend $a + 0 \cdot j = a$ schreibt, während man bei Paarschreibweise eine Komponente Null mitzuschreiben hätte. Nun gilt der grundlegende (und sehr einfach zu zeigende)

SATZ 14. *Mit dieser Definition bilden die komplexen Zahlen einen Körper, erfüllen also die Körperaxiome (vgl. Kap. 1). Sie bilden die gesuchte Erweiterung von \mathbb{R} , weil man einerseits jede reelle Zahl a als $a + j \cdot 0$ in \mathbb{C} wiederfindet, andererseits offenbar mit den Zahlen 1 und j auch jede Zahl $a + jb$, $a, b \in \mathbb{R}$, in \mathbb{C} sein muss. \mathbb{C} ist also minimale Erweiterung von \mathbb{R} , so dass man eine Wurzel der Gleichung $x^2 + 1 = 0$ findet.*

Zusatz: Die komplexen Zahlen erlauben keine Anordnung mehr wie \mathbb{R} , so dass die Rechengesetze für angeordnete Körper gelten könnten. Wohl aber kann man als **Betrag** einer komplexen Zahl $z = a + jb$ gerade den Betrag von \mathbb{R}^2 übernehmen - das ist stets eine *reelle* Zahl ≥ 0 :

DEFINITION 17 (Betrag einer komplexen Zahl). $|a + jb| = \sqrt{a^2 + b^2}$.

Nun zur Begründung des Satzes: Direkt mit den Definitionen für Addition und Multiplikation ist einzusehen, dass sie in der Menge von \mathbb{C} stets eindeutig ausführbar sind und die Resultate wieder in dieser Menge liegen. Der restliche Beweis ist bis auf eine Stelle durch simples Einsetzen der Definitionen und Nutzung der Rechengesetze für \mathbb{R} zu führen. Hier die interessante Stelle: Wir übersehen sofort, dass $0 + j \cdot 0$ die Null und $1 + j \cdot 0$ die Eins sind, und dass $-(a + jb) = -a + j(-b)$ als Inverses bezüglich der Addition fungiert. (Bildung wie die von $-\vec{x}$ in der Vektorrechnung.) Einzig die Existenz der Inversen bezüglich der Multiplikation bzw. die Ausführbarkeit der Division durch alle Zahlen $\neq 0$ ist nicht unmittelbar zu sehen. Wir überlegen dazu, wie $1/z$ aussehen müsste, indem wir die Körper-Rechengesetze anwenden. Das liefert aber nur die Idee, ist kein Beweis. Für den müssen wir den Spieß umdrehen: Wir haben zu prüfen, dass der gewonnene Ausdruck das Verlangte leistet. Wir setzen $z = a + jb$ und rechnen:

$$\frac{1}{a + jb} = \frac{a - jb}{(a + jb)(a - jb)} = \frac{a - jb}{a^2 + b^2} = \frac{a}{a^2 + b^2} + j \frac{-b}{a^2 + b^2}.$$

Das diente nur zum Auffinden: Wir dürfen nicht voraussetzen, dass $1/z$ überhaupt bildbar ist. Jetzt die Argumentation: Wir behaupten, dass für $a + jb = z \in \mathbb{C}$ mit $z \neq 0$ (d.h. mindestens eine der Zahlen a, b verschieden von Null) stets die Zahl

$$1/z = z^{-1} := \frac{a}{a^2 + b^2} + j \frac{-b}{a^2 + b^2}$$

eindeutig definiert ist in \mathbb{C} (das ist unmittelbar mit dieser Definition klar) und das für ein multiplikativ Inverses Verlangte leistet. Das können wir sofort nachrechnen - und zwar setzen wir hier *nur* die Definition der komplexen Multiplikation voraus und verwenden sonst nur die Gesetze des Rechnens mit reellen Zahlen (!):

$$\begin{aligned} z^{-1}z &= \left(\frac{a}{a^2 + b^2} + j \frac{-b}{a^2 + b^2} \right) (a + jb) \\ &= \frac{a}{a^2 + b^2}a - \frac{-b}{a^2 + b^2}b + j \left(\frac{a}{a^2 + b^2}b + \frac{-b}{a^2 + b^2}a \right) \\ &= 1 + j \cdot 0, \text{ also die Eins wie verlangt.} \end{aligned}$$

Bemerkung: Mit dem Satz kann man stets wie üblich rechnen, d.h. Klammer- und Vorzeichenregeln, Distributivgesetze, Bruchrechnen etc. funktionieren alle wie gewohnt. Wendet man das fertige Resultat an, so kann man die obenstehende Rechnung viel leichter ausführen (im Beweis ging das eben nicht, weil das Resultat nicht anzuwenden, sondern erst zu beweisen war). Das sieht dann so aus:

$$\left(\frac{a}{a^2 + b^2} + j \frac{-b}{a^2 + b^2} \right) (a + jb) = \frac{(a - jb)(a + jb)}{a^2 + b^2} = 1.$$

Für unsere Rechnung war es entscheidend, den Nenner mit $(a - jb)$ zu erweitern, um j herauszubringen. (Den Trick kannten wir schon, analog bringen wir Wurzeln aus Nennern wie $a + b\sqrt{2}$) Diese Bildung ist von großer Wichtigkeit für \mathbb{C} , daher heben wir Definition und wesentliche Eigenschaften hervor:

DEFINITION 18 (komplexe Konjugation). Die Zahl $\bar{z} = a - jb$ ($a, b \in \mathbb{R}$) (lies: „ z quer“) heißt die *Konjugierte* zu $z = a + jb$. Die Abbildung $\mathbb{C} \rightarrow \mathbb{C}$, $z \mapsto \bar{z}$ heißt (komplexe) *Konjugation*.

SATZ 15. Die Konjugation ist vertauschbar mit $+$, $*$, d.h. $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$, und $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$, und sie ist ihre eigene Umkehrabbildung, d.h. $\overline{\bar{z}} = z$. Weiter lässt die Konjugation reelle Zahlen fest, d.h. $\overline{a + j \cdot 0} = a + j \cdot 0$. Zusammengefasst: Die Konjugation ist ein Körperisomorphismus von \mathbb{C} auf sich selbst, man sagt daher: ein *Körperautomorphismus* von \mathbb{C} , der \mathbb{R} fest lässt.

Die Begründung ist ganz leicht: Man rechnet einfach z.B. aus:

$$\begin{aligned} \overline{(a + jb)(c + jd)} &= \overline{(ac - bd) + j(bc + ad)} = (ac - bd) - j(bc + ad) \\ &= (a - jb)(c - jd) = \overline{a + jb} \cdot \overline{c + jd}. \end{aligned}$$

Analog für die Addition. Eine nützliche Folgerung kann man sofort ziehen:

FOLGERUNG 1. *Bei jedem Polynom mit reellen Koeffizienten treten die nicht reellen Nullstellen paarweise konjugiert auf. (Damit auch alle, weil die Konjugierte einer reellen Zahl diese selbst ist.)*

Der Beweis ist mit dem Satz sehr einfach (Voraussetzung: Die Koeffizienten α_i sind reell!). Sei $f(z) = \sum_{i=0}^n \alpha_i z^i = 0$. Dann

$$f(\bar{z}) = \sum_{i=0}^n \alpha_i \bar{z}^i = \sum_{i=0}^n \overline{\alpha_i z^i} = \overline{\sum_{i=0}^n \alpha_i z^i} = \overline{0} = 0.$$

Also ist \bar{z} Nullstelle von f , wenn z es ist.

4. Das anschauliche Verständnis von \mathbb{C}

Das anschauliche Verstehen von \mathbb{C} ist gut möglich und daher sehr hilfreich und wichtig. Klar ist zunächst, dass man \mathbb{C} wie \mathbb{R}^2 anschaulich besitzt: Man fasst die komplexen Zahlen als Ortsvektoren in der Ebene auf. Dazu identifiziert man \vec{e}_1 mit der Zahl 1 und \vec{e}_2 mit der Zahl j und lässt diese Vektoren ein kartesisches Koordinatensystem bilden. Die Zahl $z = a + jb$ hat dann als (*reelle!*) kartesische Koordinaten ihren Realteil und ihren Imaginärteil.

DEFINITION 19. *Die kartesischen Koordinaten von $z = a + jb$ sind $\operatorname{Re}(z) = a$ und $\operatorname{Im}(z) = b$. Beide sind reelle Zahlen! Die Koordinatenachsen nennt man reelle und imaginäre Achse.*

Vorsicht: Der Imaginärteil von $z = a + jb$ ist $\operatorname{Im}(z) = b \in \mathbb{R}$, nicht etwa jb , was für $b \neq 0$ in \mathbb{C} ist, aber gerade nicht in \mathbb{R} . Entsprechend ist $|a + jb| = \sqrt{a^2 + b^2}$, also $|z|^2 = \operatorname{Re}^2(z) + \operatorname{Im}^2(z) = z \cdot \bar{z}$. (Nicht etwa kommt so etwas wie $a^2 + (jb)^2 = a^2 - b^2$, was negativ sein kann und woraus dann gar keine reelle Quadratwurzel zu ziehen ist.)

Entsprechend kann man auch die Addition und Subtraktion von komplexen Zahlen verstehen wie bei den Vektorpfeilen im \mathbb{R}^2 . Dasselbe gilt für die Multiplikation einer komplexen Zahl mit einer *reellen*. So weit haben wir natürlich nur die Vektorraumstruktur von \mathbb{C} anschaulich verstanden wie zuvor die von \mathbb{R}^2 . Wie steht es mit der Multiplikation und Division? Dazu müssen wir einen anderen anschaulichen Zugang nutzen. Offenbar kann man eine komplexe Zahl auch charakterisieren durch ihre Länge und ihren Winkel, gemessen von der positiven reellen Achse an entgegen dem Uhrzeigersinn. (Bei der Zahl 0 ist der Winkel beliebig, nicht eindeutig.) Diese beiden Zahlen, Länge und Winkel, heißen die Polarkoordinaten von z . Wir wissen bereits, dass wir jeden Vektor aus \mathbb{R}^2 schreiben können als $r(\cos(\varphi), \sin(\varphi))$, als komplexe Zahl also in der Form $r(\cos(\varphi) + j \sin(\varphi))$. Daher definieren wir:

DEFINITION 20. *Die Polarkoordinaten der komplexen Zahl $z = r(\cos(\varphi) + j \sin(\varphi))$, $r \geq 0$, $r \in \mathbb{R}$, $0 \leq \varphi < 2\pi$, sind r und φ . Wichtige systematisch berechnete (s.u.) und abkürzende Schreibweise:*

$$e^{j\varphi} := \cos(\varphi) + j \sin(\varphi).$$

Daher nennt man den Ausdruck $re^{j\varphi}$ auch Polardarstellung von $z = re^{j\varphi}$. Dabei muss stets $r \geq 0$ sein, dagegen erlaubt man sich auch negative Werte für φ - es ist dann entsprechend mit dem Uhrzeigersinn um den Betrag des Winkels zu drehen.

SATZ 16 (Koordinatenumwandlung). *Hat z die Polarkoordinaten r, φ , so lauten die kartesischen Koordinaten von z : $\operatorname{Re}(z) = r \cos(\varphi)$, $\operatorname{Im}(z) = r \sin(\varphi)$. Ist $z = a + jb$ mit kartesischen Koordinaten $a, b \in \mathbb{R}$ gegeben, so sind die Polarkoordinaten an folgender Darstellung jeweils abzulesen:*

$$z = \begin{cases} \sqrt{a^2 + b^2} e^{j \arctan(b/a)}, & \text{wenn } a > 0, \\ \sqrt{a^2 + b^2} e^{j(\pi + \arctan(b/a))}, & \text{wenn } a < 0, \\ b e^{j\pi/2}, & \text{wenn } a = 0, b > 0, \\ |b| e^{-j\pi/2}, & \text{wenn } a = 0, b < 0. \end{cases}$$

(Hinzu kommt noch $0e^{j \cdot 0}$ z.B. für $z = 0$, was nicht eindeutig und ziemlich uninteressant ist.)

Begründung des zweiten Teils: Den Betrag kennen wir schon, lediglich der Winkel zu gegebenen kartesischen Koordinaten ist fraglich. Für $a \neq 0$ ist offenbar $\tan(\varphi) = b/a$. Nun müssen wir ein wenig aufpassen: Die Tangensfunktion ist nur im Bereich $] -\pi/2, \pi/2[$ umkehrbar. Mit der Setzung $\varphi = \arctan(b/a)$ erhalten wir nur diese Winkel. Diese Berechnung taugt also gerade für den Fall $a > 0$. Haben wir $a < 0$, so bilden wir zu $z = a + jb$ die negative Zahl $-a - jb$, deren Winkel korrekt $\arctan(b/a)$ ist. Dazu haben wir offenbar gerade π zu addieren, um auf den Winkel für z zu kommen.

Bemerkung über exakte Darstellungen: Eine exakte Darstellung mit einfachen kartesischen Koordinaten führt im allgemeinen auf eine komplizierte in Polarkoordinaten, die nicht exakt ausrechenbar ist, wohl angebbbar im Sinne der obenstehenden Formeln. Umgekehrt gehört zu einer einfachen Polardarstellung im allgemeinen keine einfache kartesische. Man sollte jedoch für die einfachen Standardwinkel die exakten Sinus- und Cosinuswerte kennen:

Winkel	Sinuswert	Cosinuswert
0	0	1
$\pi/6$	$\frac{1}{2} = \frac{1}{2}\sqrt{1}$	$\frac{1}{2}\sqrt{3}$
$\pi/4$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{2}$
$\pi/3$	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}$

Dazu achte man auf die Vorzeichen. Beispiel: $-2-2j = 2\sqrt{2}e^{-3\pi/4}$. Umgekehrt: $3e^{7\pi/6} = 3(-1/2 - j \cdot \sqrt{2}/2) = -3/2 - j \cdot \sqrt{2}/2$.

Mittels der Polarkoordinaten gelingt folgende anschauliche Auffassung der Multiplikation: Eine Zahl z wird graphisch mit z_0 multipliziert, indem man die Zahl z mit $|z_0|$ streckt und das Resultat um den Winkel von z_0 weiterdreht entgegen dem Uhrzeigersinn. (Bei negativer Winkelangabe für z_0 bedeutet das eine Drehung um den absoluten Winkel mit dem Uhrzeigersinn.) Die Begründung folgt im nächsten Teil:

5. Multiplikation usw. in Polarkoordinaten

Grundlegend ist folgender

$$e^{j\varphi} e^{j\psi} = e^{j(\varphi+\psi)}.$$

Beweis: Schreiben wir die Gleichung aus mit der Definition von $e^{j\alpha}$, so haben wir

$$(\cos(\varphi) + j \sin(\varphi)) (\cos(\psi) + j \sin(\psi)) = \cos(\varphi + \psi) + j \sin(\varphi + \psi).$$

Andererseits ergibt Ausmultiplizieren der linken Seite

$$\cos(\varphi) \cos(\psi) - \sin(\varphi) \sin(\psi) + j (\sin(\varphi) \cos(\psi) + \cos(\varphi) \sin(\psi)).$$

Wegen der linearen Unabhängigkeit von $1, j$ über \mathbb{R} können wir Koeffizientenvergleich anstellen, haben also gleichwertig:

$$\begin{aligned} \cos(\varphi + \psi) &= \cos(\varphi) \cos(\psi) - \sin(\varphi) \sin(\psi) \text{ und} \\ \sin(\varphi + \psi) &= \sin(\varphi) \cos(\psi) + \cos(\varphi) \sin(\psi). \end{aligned}$$

Das sind nun genau die Additionstheoreme von Cosinus und Sinus. Wir können sie z.B. beweisen, indem wir einfach das Matrizenprodukt der Drehmatrizen ausrechnen:

$$\begin{aligned} & \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \begin{pmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{pmatrix} \\ = & \begin{pmatrix} \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi) & -\sin(\varphi)\cos(\psi) - \cos(\varphi)\sin(\psi) \\ \sin(\varphi)\cos(\psi) + \cos(\varphi)\sin(\psi) & \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi) \end{pmatrix}. \end{aligned}$$

Andererseits wissen wir, dass die Hintereinanderschaltung der Drehungen mit den Winkeln ψ, φ die Drehung mit $\varphi + \psi$ ergibt, also ist das Produkt der Matrizen auch gleich

$$\begin{pmatrix} \cos(\varphi + \psi) & -\sin(\varphi + \psi) \\ \sin(\varphi + \psi) & \cos(\varphi + \psi) \end{pmatrix}.$$

Koeffizientenvergleich nur in der ersten Spalte der entstehenden Matrixgleichung ergibt die benötigten Additionstheoreme.

FOLGERUNG 2. Seien $re^{j\varphi}, se^{j\psi}$ Polardarstellungen zweier komplexer Zahlen, dann gilt:

$$\begin{aligned} re^{j\varphi} se^{j\psi} &= rse^{j(\varphi+\psi)} \\ \frac{re^{j\varphi}}{se^{j\psi}} &= \frac{r}{s}e^{j(\varphi-\psi)}, \quad s \neq 0 \\ (re^{j\varphi})^n &= r^n e^{jn\varphi}, \quad n \in \mathbb{N}. \end{aligned}$$

Ferner erhält man folgendermaßen alle n -ten Wurzeln von $z = re^{j\varphi}$:

$$\text{Die } n\text{-ten Wurzeln von } z = re^{j\varphi} \text{ sind } r^{1/n}e^{j\left(\frac{\varphi+k\cdot 2\pi}{n}\right)}, \quad k = 0, 1, \dots, n-1.$$

Mit den Additionstheoremen ist nicht mehr viel zu tun, nur die Beträge sind abzutrennen und der üblichen Rechnung in \mathbb{R} zuzuführen. Für die Wurzeln beachte man, dass die Ausdrücke

$$e^{j(\varphi+k\cdot 2\pi)} \quad \text{für } k = 0, \dots, n-1$$

verschiedene Rechenausdrücke derselben komplexen Zahl sind, dass jedoch Teilen des Winkels durch n lauter verschiedene Winkel unterhalb eines Vollwinkels ergeben. Beispiel: Alle 5. Wurzeln von 1 lauten $e^{2kj\pi/5}$, $k = 0, 1, 2, 3, 4$. Anschaulich bilden alle n -ten Wurzeln einer Zahl z ein gleichmäßiges Schiffssteuerrad mit n Speichen.

6. Lösen von einfachen Gleichungen in \mathbb{C}

Die einfachen Gleichungen, die wir lösen können, sind lineare (auch Gleichungssysteme) und quadratische. Dies ist auch in \mathbb{C} wichtig. Zu den linearen Gleichungen und auch Gleichungssystemen bemerken wir nur kurz, dass alles Systematische, was wir in Kapitel 1 darüber sagten, voll und ganz erhalten bleibt, also alle Resultate über Lösungsverfahren, Struktur der Lösungsmengen, Existenz und Eindeutigkeit von Lösungen. Grundlegend war dafür die Vektorraumstruktur von \mathbb{R}^n über \mathbb{R} . Nun hat aber \mathbb{C}^n über \mathbb{C} dieselbe Struktur. Für das Lösungsverfahren insbesondere benutzten wir lediglich die Grundrechenarten im Zahlenkörper. Die verhalten sich jedoch in \mathbb{C} ebenso. Lediglich wird es etwas mühsamer, die benötigten Multiplikationen auszuführen, um etwa zwei Zeilen eines Gleichungssystems miteinander zu kombinieren. Man merke sich nur: Alle Zahlen sind nun komplexe, ebenso müssen die Parameter \mathbb{C} statt \mathbb{R} durchlaufen, wenn wir eine unendliche Lösungsmenge eines linearen Gleichungssystems mit Koeffizienten in \mathbb{C} bekommen. Auch die Matrixeinträge bei Matrixendarstellung sind nunmehr komplex. Wir illustrieren das an einem Beispiel:

$$\begin{aligned} \frac{1}{2-j}z_1 + 3jz_2 - (1+j)z_3 &= 0 \\ 2z_1 - (1+j)z_2 + jz_3 &= 0 \end{aligned}.$$

Gesucht ist die Lösungsmenge in \mathbb{C}^3 in den Unbestimmten z_1, z_2, z_3 . Zunächst bringen wir die erste Zeile in ordentliche Form, indem wir sie mit $2 - j$ multiplizieren, das ergibt

$$\begin{aligned} z_1 + (3 + 6j)z_2 - (3 + j)z_3 &= 0 \\ 2z_1 - (1 + j)z_2 + jz_3 &= 0 \end{aligned}$$

Hinauswerfen von z_1 ist nun das Einfachste, Resultat mit $2(I) - (II)$ ist

$$(7 + 13j)z_2 - (6 + 3j)z_3 = 0$$

Daraus erhalten wir mit freiem Parameter z_2 :

$$\begin{aligned} z_3 &= \frac{1}{3} \cdot \frac{7 + 13j}{2 + j} z_2 = \left(\frac{9}{5} + \frac{19}{15}j \right) z_2. \\ z_1 &= \left(\frac{17}{15} - \frac{2}{5}j \right) z_2. \quad (\text{Einsetzen in die erste Gleichung oben.}) \end{aligned}$$

Die Lösungsmenge lautet also in parametrisierter Form

$$L(z) = z \left(\frac{17}{15} - \frac{2}{5}j, 1, \frac{9}{5} + \frac{19}{15}j \right), \quad z \in \mathbb{C}.$$

Parameterwerte in \mathbb{C} , ebenso die Einträge des konstanten Vektors. Sonst ist nichts verändert!

Zwei wichtige Hinweise zum Rechnen mit komplexen Zahlen:

- Man vermeide unbedingt falsch lineares Rechnen mit Brüchen - die Gefahr ist erfahrungsgemäß groß bei Anfängern, die Gleichung $1/z + 1/u = 1/v$ mit komplexen z, u, v umzuformen in $z + u = v$, was natürlich genau so falsch ist wie im Reellen.
- Bei Multiplikationen und Divisionen klammere man unbedingt so viel rein Reelles aus wie möglich. Sonst belastet man die komplexen Operationen noch unnötig mit dauerndem Hineinmultiplizieren eventuell gewichtiger reeller Faktoren. (Ausbleiben der entsprechenden Vereinfachung führt in aller Regel zum „Versanden“ der Rechnung.)

Ein Beispiel aus der komplexen Rechnung mit Widerständen (s.u.) - R, ω, C, L sind dabei *reellwertige äußere Parameter*: Schaltet man eine Reihe von Ohmschem Widerstand R und kapazitivem Widerstand $1/j\omega C$ ($\omega =$ Kreisfrequenz des Wechselstroms) parallel mit einem induktiven Widerstand $j\omega L$, so ergibt sich die Addition der Leitwerte (Kehrwerte der Widerstandswerte, die hier komplex sind), man hat

$$\begin{aligned} \frac{1}{Z} &= \frac{1}{j\omega L} + \frac{1}{R + \frac{1}{j\omega C}}, \quad \text{und das ergibt korrekt} \\ Z &= \frac{\left(R + \frac{1}{j\omega C} \right) j\omega L}{R + j \left(\omega L - \frac{1}{\omega C} \right)}; \quad (\text{Bruchrechnung!}) \end{aligned}$$

Nun wird man jedenfalls den reellen Faktor vor das Ganze schreiben (Regel: „Alles Reelle ausklammern, so weit wie möglich!“) und den verbleibenden Bruch mit ωC erweitern (Doppelbrüche zu beseitigen), dann den Nenner $j-$ frei machen:

$$\begin{aligned} Z &= \omega L \frac{1 + jR\omega C}{R\omega C + j(\omega^2 CL - 1)} = \omega L \frac{(1 + jR\omega C)(R\omega C + j(1 - \omega^2 LC))}{R^2\omega^2 C^2 + (\omega^2 CL - 1)^2} \\ &= \omega L \frac{R\omega^3 C^2 L + j(R^2\omega^2 C^2 + 1 - \omega^2 CL)}{R^2\omega^2 C^2 + (\omega^2 CL - 1)^2}. \end{aligned}$$

Es ist nicht tunlich, das zweite Quadrat im Nenner auszumultiplizieren, das macht den Nenner nur unübersichtlicher. Ein wichtiges Hilfsmittel ist hier die Einheitenkontrolle: ωL hat die Einheit eines Widerstandes, also sollte der Bruch einheitenfrei sein bzw. Einheit Eins haben. Man prüfe nach: Jeder Summand im Zähler hat Einheit Eins, ebenso jeder Summand des Nenners. Träfe man Summanden verschiedener Einheiten an, so hätte

man vermutlich irgendeinen Faktor falsch, z.B. irgendwo eine falsche Potenz, wenn nicht Schlimmeres. Hat man nun weitere Rechnungen mit Z auszuführen, so lässt man tunlich den hier ausgeklammerten reellen Faktor ωL - außen vor und multipliziert ihn niemals hinein. Dazu ein weiteres einfaches Beispiel: Seien a, b, c reellwertige Parameter, dann hat man folgende Berechnung einer kartesischen Endform:

$$\frac{1+j}{\frac{a}{b^2c-c^2b}(2-3j)} = \frac{b^2c-c^2b}{a} \frac{1+j}{2-3j} = \frac{b^2c-c^2b}{13a} (-1+5j).$$

In \mathbb{C} können wir nicht nur quadratische Gleichungen mit reellen Koeffizienten lösen, sondern auch alle solchen mit komplexen Koeffizienten. Das verspricht der oben angeführte Satz von der Vollständigkeit des Körpers der komplexen Zahlen. Wir wollen es aber konkret einsehen und die Rechnung kennenlernen. (Nebenbei haben wir dann neben der Existenz von n verschiedenen Wurzeln der Gleichung $z^n = z_0$ eine zweite konkrete Illustration jenes Satzes durch einen Spezialfall.) Betrachten wir eine beliebige quadratische Gleichung $\alpha z^2 + \beta z + \gamma = 0$ mit $\alpha \neq 0$, $\alpha, \beta, \gamma \in \mathbb{C}$, so können wir wie im Reellen zunächst auf Normalform $z^2 + pz + q = 0$ bringen und haben dann dieselbe Lösungsformel (das alles hing nur am Rechnen in einem beliebigen Zahlkörper!):

$$z_{1,2} = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}.$$

Da nun aber im allgemeinen unter der Wurzel eine komplexe Zahl ist, stellt sich die Frage nach einer ordentlichen Berechnung einer Quadratwurzel einer komplexen Zahl. Bei einer in Polarform gegebenen Zahl ist das natürlich sehr leicht. Aber aus unserer Rechnung werden wir hier aber ein kartesische Form vorfinden und eine ebensolche exakt berechnen wollen. Wir geben also eine beliebige Zahl $z = a + jb$ vor (im Beispiel denken wir an die ausgerechnete Zahl unter der Wurzel) und suchen eine komplexe Zahl u mit $u^2 = z$. Wir setzen $u = x + jy$ kartesisch mit unbekanntem Koordinaten an und haben die Bedingungen (die erste aus Berechnung von u^2 und Gleichsetzen mit z , die zweite über $|u|^2 = |z|$):

$$\begin{aligned} x^2 - y^2 + 2jxy &= a + jb, \\ x^2 + y^2 &= \sqrt{a^2 + b^2}. \end{aligned}$$

Koeffizientenvergleich in der ersten Gleichung ergibt insbesondere $x^2 - y^2 = a$, Addition zur zweiten Gleichung bzw. Subtraktion von der zweiten Gleichung führen zu

$$\begin{aligned} x &= \pm \frac{\sqrt{2}}{2} \sqrt{a + \sqrt{a^2 + b^2}}, \quad y = \pm \frac{\sqrt{2}}{2} \sqrt{-a + \sqrt{a^2 + b^2}}, \\ 2xy &= b \text{ kennzeichnet die erlaubten Vorzeichenverteilungen.} \end{aligned}$$

Man beachte, dass alle Zahlen hier reell sind, und dass die Wurzeln stets im Reellen existieren! Nunmehr haben wir noch die Bedingung aus der ersten Gleichung zu erfüllen, dass $2xy = b$. Also müssen in unseren Lösungsformeln für x, y die Vorzeichen so verteilt werden: Im Falle $b < 0$ verschieden, im Falle $b > 0$ gleich. Es gibt damit auch nur zwei und nicht etwa vier Lösungen, also Zahlen $u \in \mathbb{C}$, die Quadratwurzeln von z sind. Man beachte, dass die Formeln auch korrekt sind für den Fall $b = 0$: Wenn zusätzlich $a > 0$, so haben wir $y = 0$, $x = \pm\sqrt{a}$, und bei $a < 0$ ergibt sich $x = 0, y = \pm\sqrt{-a}$. Für $a = b = 0$ gibt es nur die Lösung $x = y = 0$.

Reelle Funktionen

Dies Kapitel besteht aus drei Abschnitten. Im ersten lernt man die wichtigsten speziellen Funktionen kennen und miteinander verbinden, insbesondere elementare Überlegungen zur Beziehung zwischen Rechenausdrücken und Graphen herzustellen. Die andern Teile handeln dann vom Differenzieren und Integrieren, mit wichtigen Anwendungen.

1. Reelle Funktionen und ihre Graphen

1.1. Grundbegriffe zur Beschreibung wichtiger Phänomene bei Funktionen und ihren Graphen. Die einfachsten Funktionen sind die linearen mit Geraden als Funktionsgraphen, diese und quadratische Funktionen haben wir bereits elementar kennengelernt. Die nächsten Verallgemeinerungen sind die Polynome. Ihre Graphen gehen glatt auf und ab, um dann ins Unendliche zu verschwinden. Es ist im allgemeinen dabei nicht mehr möglich, Nullstellen und Extrema geschlossen auszurechnen, man bedient sich der Näherungslösungen in komplizierten Fällen. Die nächste algebraische Klasse sind die gebrochen rationalen Funktionen, ihre Ausdrücke (Funktionsterme) sind Brüche von Polynomen. Sie können ein neue Phänomene zeigen: (Gerade) Asymptoten parallel zur x -Achse sowie Pole - das sind (gerade) Asymptoten parallel zur y - Achse. Unter einer **Asymptote** einer Kurve in einer bestimmten Umgebung versteht man generell eine andere Kurve, die sich in dieser Umgebung beliebig der ersten nähert. Zum Beispiel ist der Graph von $g(x) = x^2$ eine (relativ einfache!) Asymptote des Graphen von $f(x) = \sqrt{x^4 + 1}$ für $x \rightarrow \pm\infty$. Nun zu den einfachsten Asymptoten, die Geraden sind: Eine Asymptote für den Graphen von f parallel zur x - Achse entsteht, wenn $f(x)$ für $x \rightarrow \infty$ bzw. $x \rightarrow -\infty$ gegen einen endlichen Grenzwert geht. Ein **Pol** (Asymptote parallel zur y - Achse) von f liegt an der Stelle x_0 vor, wenn die Funktionswerte $f(x)$ gegen ∞ oder $-\infty$ gehen, falls man sich mit den x - Werten beliebig der Stelle x_0 nähert (einseitig oder beidseitig). Zum Beispiel hat $f(x) = 1/x$ einen Pol bei $x_0 = 0$, und dies können wir sofort verallgemeinern: Wenn $f(x) = g(x)/h(x)$ und $h(x) \rightarrow 0$ für $x \rightarrow x_0$, während der Zähler gegen einen Wert verschieden von Null geht, so hat f an der Stelle x_0 einen Pol. Noch allgemeiner: Wenn Zähler und Nenner beide nach Null gehen für $x \rightarrow x_0$, der Zähler aber langsamer als der Nenner, so entsteht ein Pol. Um zu quantifizieren, was „langsamer“ ist, wird man später die Ableitung heranziehen. Aber es sollte zunächst klar sein, dass eine höhere Potenz bei Polynomen entscheidet: $(x^3 + 2x^2)/(x^4 + x)$ hat sicher in $x_0 = 0$ einen Pol. Ebenso geht diese Funktion sicher nach Null für $x \rightarrow \pm\infty$. Der Graph einer gebrochen rationalen Funktionen, welche Pole besitzen, zerfällt in mehrere Äste, die wiederum ganz glatt sind - man spreche dann nicht von mehreren Graphen.

Weiter in der Klassifikation der Funktionen: Es gibt äußerst wichtige Funktionen, welche als Grundbausteine für die Naturbeschreibung fungieren, die aber komplizierter sind als Polynome und gebrochen-rationale Funktionen. Das sind die im nächsten Abschnitt zu besprechenden trigonometrischen Funktionen sowie Exponential- und Logarithmusfunktionen. Sie haben rechnerisch exakte Reihendarstellungen der Form $f(x) = \sum_{i=0}^{\infty} \alpha_i (x - x_0)^i$, und sie können schneller oder langsamer wachsen als jedes Polynom, auch unendlich viele Nullstellen haben usw. Zunächst jedoch zu den Grundbegriffen. Zunächst sei betont:

DEFINITION 21. *Eine reelle Funktion (im einfachsten Sinne) ist eine Abbildung $A \rightarrow B$, mit $A, B \subseteq \mathbb{R}$ (also Mengen reeller Zahlen) und einer eindeutigen Zuordnungsvorschrift $x \mapsto f(x)$, meist als Gleichung $f(x) = \dots$ (Ausdruck in x) geschrieben. A heißt Definitionsbereich und B der Wertebereich der Funktion. Oft gibt man nur die Gleichung und meint dann den maximalen reellen Definitionsbereich und als Wertebereich \mathbb{R} .*

Definitions- und Wertebereichsangaben werden wichtig, sobald man von der Umkehrbarkeit einer Funktion spricht:

DEFINITION 22. $f : A \rightarrow B$ heißt umkehrbar, wenn es zu jedem $b \in B$ genau ein $a \in A$ gibt mit der Eigenschaft $f(a) = b$. Dann definiert nämlich $f^{-1} : B \rightarrow A$, $b \mapsto$ (dasjenige $a \in A$, so dass $f(a) = b$), eine Funktion, die sogenannte Umkehrfunktion oder inverse Funktion von f .

Man beachte: f^{-1} ist keineswegs dasselbe wie $1/f$, die man reziproke Funktion von f nennt, d.h. $f^{-1}(x)$ ist nicht etwa $1/f(x)$. Beispiel: $f :]0, \infty[\rightarrow]0, \infty[$, $f(x) = x^2$, ist umkehrbar, mit der Quadratwurzelfunktion auf $]0, \infty[$ als Umkehrfunktion. Aber die Quadratfunktion g mit Definitionsbereich \mathbb{R} ist nicht umkehrbar, da man dann zwei Lösungen in x der Gleichung $g(x) = b$ mit $b > 0$ hat. Ebenso ist auch $h :]0, \infty[\rightarrow \mathbb{R}$, $h(x) = x^2$, nicht umkehrbar, weil es Zahlen $b \in \mathbb{R}$ gibt, nämlich alle negativen < 0 , für welche die Gleichung $h(x) = b$ gar keine Lösung hat. Es macht also gar keinen Sinn, die Umkehrbarkeit einer Funktion auch nur zu diskutieren, wenn Definitions- und Wertebereich nicht genau angegeben sind. Wir merken uns noch, dass man den Graphen von f^{-1} aus dem von f durch Spiegelung an der Geraden $y = x$ erhält.

Wird eine Funktion nur durch einen Rechenausdruck angegeben, so mache man sich stets klar, an welchen Stellen dieser Ausdruck undefiniert sein könnte, und betrachte den maximalen reellen Definitionsbereich. Die typischen Quellen sind: Ein Nenner wird Null, eine Quadratwurzel wäre aus einer negativen Zahl zu ziehen, oder auch eingebaute Pole in einer Grundfunktion wie Logarithmus. Man beachte, dass ein Rechenausdruck bereits dann nicht definiert ist, wenn nur einer seiner Teilausdrücke nicht definiert ist. Z.B. ist $1/\sin(x) + 1/(x+1) + \sqrt{10-x^2}$ undefiniert bei allen Nullstellen von Sinus sowie bei $x = -1$, schließlich für alle x mit $|x| > \sqrt{10}$.

Man beachte weiter, dass über das Verhalten einer Funktion in der Umgebung einer Stelle, an der sie nicht definiert ist, oder auch an der Grenze eines Bereiches von Undefiniertheit, im allgemeinen etwas Genaueres beobachtet und gesagt werden sollte. Es ist ein notorischer Anfängerfehler, zu verkennen, dass z.B. „ f hat in x_0 einen Pol“ viel mehr sagt als „ f ist an der Stelle x_0 nicht definiert“. Betrachten wir dazu ein Beispiel: $f(x) = \sin(x)/x$ ist für $x = 0$ nicht definiert. Aber sie hat dort keinen Pol, vielmehr läuft der Graph glatt durch diese Stelle mit y -Wert 1, man kann die „Definitionslücke“ also problemlos schließen (s.u. im Abschnitt über die trigonometrischen Funktionen) durch die Definition:

$$g(x) = \begin{cases} \frac{\sin(x)}{x}, & \text{wenn } x \neq 0 \\ 1 & \text{für } x = 0. \end{cases}$$

Dagegen funktioniert so etwas nicht für $h(x) = \sin(x)/x^2$. h hat einen Pol bei $x = 0$. Wieder anders verhält sich $f(x) = \sqrt{x}$ bei $x = 0$, an der Grenze ihres Definitionsbereiches.

Elementarste Grundphänomene sind weiter das qualitative Steigungsverhalten einer Funktion sowie die Standard-Symmetrien:

DEFINITION 23. f heißt monoton wachsend in einem Bereich D , wenn für alle $a, b \in D$ gilt: Wenn $a \leq b$, so $f(a) \leq f(b)$. („Streng monoton wachsend“: Statt \leq steht beide Male $<$ in der Bedingung.) Monotonies Fallen einer Funktion f bedeutet entsprechend, dass stets $f(a) \geq f(b)$, wenn $a \leq b$, in einem anzugebenden Bereich.

Beispiele: $f(x) = 1/x$ ist streng monoton fallend im Bereich $]0, \infty[$, nicht aber im gesamten Definitionsbereich aller Zahlen $\neq 0$. Die Sinusfunktion ist streng monoton steigend im Bereich $[-\pi/2, \pi/2]$ und streng monoton fallend im Bereich $[\pi/2, 3\pi/2]$.

DEFINITION 24. f sei in einem Bereich D symmetrisch um Null definiert. Dann heißt f eine **gerade Funktion**, wenn für alle $x \in D$ gilt: $f(x) = f(-x)$. f heißt eine **ungerade Funktion**, wenn für alle $x \in D$ gilt: $f(x) = -f(-x)$.

Beispiele: $f(x) = x^n$ ist für gerades natürliches n eine gerade Funktion, für ungerades eine ungerade Funktion.

Wir werden uns bei der Diskussion von Funktionsgraphen zweckmäßig so verhalten, dass wir die Eigenschaften eines gewissen Katalogs von Grundfunktionen auswendig kennen und dann auf die Eigenschaften (einfacher)

Zusammensetzungen schließen, d.h. solcher Funktionen, die aus den Rechenausdrücken der Grundfunktionen aufgebaut sind. Das genaue Verständnis der Zusammensetzungen wird auch bei Ableitung und Integral eine wichtige Hilfe sein. Stets wird der Aufbau sein: Zunächst die Grundfunktionen, dann die Zusammensetzungen. Wir werden die Zusammensetzungen systematisch nach den Grundfunktionen besprechen, jedoch bereits bei den Grundfunktionen exemplarisch kennenlernen. Eine besonders einfache Art der Zusammensetzung besteht darin, einen Grundausdruck einfach nur linear zu transformieren. Dann ist insbesondere das graphische Verhalten sofort komplett ablesbar. Daher nehmen wir dies vorweg:

1.2. Lineare Transformationen von Funktionsgraphen, und die zugehörigen Veränderungen der Rechenausdrücke. Betrachten wir ein Beispiel: $g(x) = (x - 1)^2$, $x \in \mathbb{R}$. Die zugehörige Grundfunktion ist $f(x) = x^2$. Wir haben $g(3) = f(3 - 1) = f(2)$, usw., also entsteht der Graph von g aus dem Graphen von f durch Verschieben parallel zu x - Achse um 1 nach rechts. (Nicht nach links, was oft aus dem Auftreten von „ $x - 1$ “ geschlossen wird!) Entsprechend hat g sein Minimum bei $x = 1$. Oder $h(x) = (2x)^2$. Der Faktor 2 bewirkt, dass $h(5) = f(10)$ usw., die Entwicklung läuft also im Zeitraffer ab. Das bedeutet: Den Graphen von h erhält man aus dem von f , indem man diesen um den Faktor 2 längs der x - Achse *staucht* (nicht: *streckt*, „wegen $2x$ “!). dabei bleibt die y - Achse fest, als Streckungszentrum. Nun können wir die beiden Schritte auch hintereinander ausführen, müssen dabei nur streng auf die Reihenfolge achten. Der Graph von $k(x) = (2x - 1)^2$ entsteht aus dem von f , indem man diesen *zuerst* um 1 nach rechts schiebt, *dann* längs der x - Achse *staucht*. Überprüfen wir das an der Nullstelle: Die Nullstelle von f sollte nach beiden Operationen um 1 nach rechts gewandert und dann im Wert halbiert sein, also bei $x = 1/2$ liegen, und das ist auch so. Entsprechend den Bemerkungen zu den einzelnen Operationen hätten wir den Rechenausdruck x^2 zuerst so zu verändern, dass wir für x einsetzen: $x - 1$, im zweiten Schritt hätten wir für x einzusetzen: $2x$. Auf diese Weise entsteht gerade $k(x)$. Dagegen ist die Reihenfolge in der Berechnung von $2x - 1$ aus x gerade umgekehrt! Man überprüfe, dass die umgekehrte Reihenfolge der Operationen zum Ausdruck $(2(x - 1))^2$ gehört. Entsprechende graphische Operationen sind hinsichtlich der y - Achse möglich, wobei die Deutung einfacher ist, da insbesondere keine Reihenfolgenveränderung im Spiel ist: $g(x) = 2f(x)$ bedeutet eine Streckung längs der y - Achse mit Faktor 2, $h(x) = 2f(x) + 3$ eine Verschiebung um 3 nach oben *nach* der Streckung. Das Naive ist hier also einmal das Richtige. Bemerkenswert sollte noch, dass Einsetzen von $-x$ für x eine Spiegelung an der y - Achse bedeutet, die Bildung von $g(x) = -f(x)$ eine Spiegelung an der x - Achse. Ferner bedeutet Streckung mit Faktor $1/2$ dasselbe wie Stauchung mit 2, Stauchung mit $1/2$ dasselbe wie Streckung mit 2. Fassen wir zusammen:

DEFINITION 25. g ist entsteht durch eine lineare Transformation aus f , wenn der Rechenausdruck von g auf folgende Gestalt gebracht werden kann:

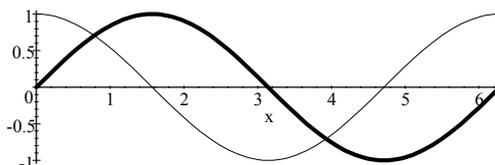
$$g(x) = \alpha f(\gamma x + \delta) + \beta.$$

Resultat dazu: Wenn $\alpha, \gamma \neq 0$, so entsteht der Graph von g aus dem von f durch die angegebenen Operationen von Strecken, Stauchen, Spiegeln. Beschriftet man die Achsen entsprechend den Koordinatentransformationen, so sieht der Graph von g wie der von f aus (das wird man gern so machen, indem man eine Achse verschiebt und die Einheiten richtig wählt, nicht so gern bei Spiegelungen die Achsenrichtungen umkehren). Ist einer der Faktoren Null, so entsteht eine banale konstante Funktion, welche die Information von f nicht mehr enthält.

1.3. Die trigonometrischen Funktionen ∇ . Aus den Definitionen von \sin, \cos am Einheitskreis folgen sofort elementargeometrisch einige wichtige Formeln:

$$\begin{aligned} \sin^2(x) + \cos^2(x) &= 1 \text{ (Pythagorasbeziehung)} \\ \sin(-x) &= -\sin(x), \quad \cos(-x) = \cos(x) \text{ (sin gerade, cos ungerade)} \\ \sin(x + 2\pi) &= \sin(x), \quad \cos(x + 2\pi) = \cos(x) \text{ (periodisch, Periodendauer } 2\pi) \\ \sin(x + \pi) &= -\sin(x), \quad \cos(x + \pi) = -\cos(x) \text{ (Antiperiode } \pi) \\ \sin(x + \pi/2) &= \cos(x), \quad \cos(x - \pi/2) = \sin(x), \text{ Folgerungen:} \\ \sin(\pi/2 - x) &= \cos(x), \quad \cos(\pi/2 - x) = \sin(x). \end{aligned}$$

Die Letzteren ergeben sich so aus den vorigen: $\cos(\pi/2 - x) = \cos(x - \pi/2)$, mit der Symmetrie. $\sin(\pi/2 - x) = -\sin(x - \pi/2) = -(-\sin(x + \pi/2)) = \cos(x)$. Hier sind noch einmal die Graphen im Bereich $[0, 2\pi]$, fett: Sinus, dünn: Cosinus. Man erkennt, dass der Graph von Sinus durch Verschiebung des Cosinus-Graphen um $\pi/2$ nach rechts entsteht, d.h. $\sin(x) = \cos(x - \pi/2)$. Es ist so skaliert, dass die richtigen Steigungen graphisch entstehen:

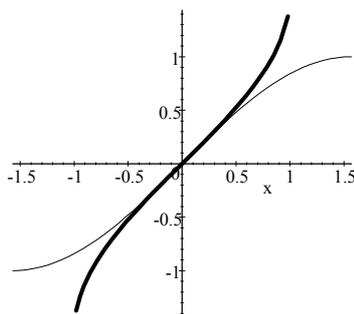


Die Nullstellen von \sin sind: $k\pi$, $k \in \mathbb{Z}$, von \cos : $\pi/2 + k\pi$, $k \in \mathbb{Z}$. Die Maxima von \sin liegen an den Stellen $\pi/2 + 2k\pi$, $k \in \mathbb{Z}$, die Minima bei $3\pi/2 + 2k\pi$, $k \in \mathbb{Z}$. (Für \cos hat man diese entsprechend um $\pi/2$ nach links zu schieben, also $\pi/2$ abzuziehen.) Entsprechend hat z.B. die Funktion $g(x) = -2\sin(3x + 2)$ ihre Maxima bei $(3\pi/2 + 2k\pi - 2)/3 = \pi/2 - 2/3 + 2k\pi/3$, $k \in \mathbb{Z}$. Das bekommt man ohne Rechnung so: Ausgehend von den *Minima* bei Sinus (wegen des negativen Vorzeichens) hat man erst um zwei nach rechts zu schieben, dann mit Faktor 3 zu stauchen (so entsteht der Graph zu $-\sin(3x + 2)$ aus dem Graphen zu $-\sin(x)$), also passiert dasselbe mit der Lage der Extremwerte. Der Vorfaktor 2 ändert nichts daran.

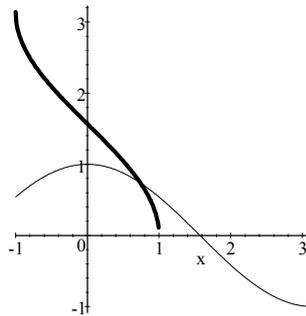
Zu den Formeln gehören die Additionstheoreme, die wir bereits bei den komplexen Zahlen kennengelernt haben. Wegen der Periodizität sind \sin, \cos natürlich nicht global umkehrbar, wohl aber bei geeigneter Einschränkung. Die Umkehrfunktionen heißen Arcusfunktionen, weil sie die Bogenlänge zum entsprechenden trigonometrischen Wert liefern; dazu hat die Tangensfunktion eine wichtige Bedeutung:

DEFINITION 26. : \arcsin ist die Umkehrfunktion zu $\sin|_{[-\pi/2, \pi/2]} [-\pi/2, \pi/2] \rightarrow [-1, 1]$, d.h. der Einschränkung von \sin auf $[-\pi/2, \pi/2]$, mit Fixierung des Wertebereichs auf die tatsächlich angenommenen Werte. \arcsin gibt also Winkel zwischen $-\pi/2$ und $\pi/2$. Analog schränkt man für \arccos den Definitionsbereich von \cos auf $[0, \pi]$ ein. Man definiert $\tan(x) = \sin(x)/\cos(x)$, für alle Werte von x , die nicht Nullstellen von \cos sind. Die Umkehrung der Einschränkung von \tan auf $]-\pi/2, \pi/2[$ nennt man \arctan . Man hat $\arctan: \mathbb{R} \rightarrow]-\pi/2, \pi/2[$.

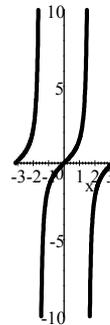
Hier sind die Graphen: Arcussinus (fett) und Sinus (dünn):



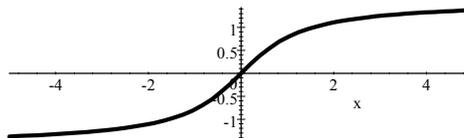
Arcuscosinus (fett) und Cosinus (dünn):



Tangens: (korrekt mit maßstäblicher Steigung, daher so gestreckt):



Arcustangens (wiederum maßstäblich für korrekte Steigung):



Die wesentliche Eigenschaft von Tangens: Sie ist ungerade und periodisch mit Periodendauer π . Nullstellen sind die von Sinus. Arcustangens ist ebenfalls ungerade und hat horizontale Asymptoten bei $y = \pm\pi/2$.

Für Tangens gilt ebenfalls ein Additionstheorem, das man mit Bruchrechnung sofort herleitet aus denen für Sinus und Cosinus:

$$\tan(x + y) = \frac{\tan(x) + \tan(y)}{1 - \tan(x)\tan(y)}$$

Den Graphen für Tangens überlegt man elementar so für den Bereich $[0, \pi/2[$: Dort haben Zählerfunktion \sin und Nennerfunktion \cos positive Werte, also auch der Tangens-Bruch. Weiter steigt die Sinusfunktion streng monoton, es fällt die Cosinusfunktion, auch streng monoton. Also steigt \tan in diesem Bereich monoton. Schließlich geht der Nenner $\cos(x)$ für $x \rightarrow \pi/2$, $x < \pi/2$, nach Null, der Zähler gegen 1. Das bedeutet einen

Pol, Steigen der Funktionswerte nach $+\infty$. Nun ist \tan offenbar ungerade und hat Periode π (wieso?). Damit ist das ganze graphische Verhalten geklärt, natürlich bis auf feinere Fragen der Steigung. Aber dazu benötigen wir nur folgenden wesentlichen Baustein, mit dem das genaue Steigungsverhalten von Sinus an der Stelle $x = 0$ geklärt ist - man beachte, dass gerade dies nur stimmt, wenn man mit dem Bogenmaß arbeitet - mit den unverbesserlichen „Grad“ ist das völlig falsch:

SATZ 17. *Ein wichtiger Grenzwert:*

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = 1$$

Diese Aussage lässt sich tatsächlich elementar begründen. Zunächst wollen wir jedoch erklären, was die Aussage mit dem „Limes“ (Grenzwert) genau bedeutet. Sie sagt: Bilden wir *irgendeine* Folge $(x_i)_{i \in \mathbb{N}}$, alle $x_i \neq 0$, die gegen Null geht, so geht die Folge $(\sin(x_i)/x_i)_{i \in \mathbb{N}}$ nach 1. Dabei ist noch zu klären, was es bedeutet, dass eine Folge von Zahlen $(x_i)_i$ gegen einen Grenzwert a geht, welche Beziehung man analog $\lim_{i \rightarrow \infty} (x_i) = a$ schreibt: Das ist *definitionsgemäß* genau dann der Fall, wenn für jede noch so kleine Umgebung von a der Gestalt $]a - \varepsilon, a + \varepsilon[$, $\varepsilon > 0$, gilt: Von einem Index i_0 ab, d.h. für alle $i \geq i_0$, gilt $x_i \in]a - \varepsilon, a + \varepsilon[$ bzw. $|x_i - a| < \varepsilon$. Insbesondere haben wir bei einer an der Stelle a stetigen Funktion *definitionsgemäß*, dass $\lim_{x \rightarrow a} f(x) = f(a)$. Für den behaupteten Grenzwert betrachten wir zunächst im Einheitskreis den Fall $x > 0$, aber $x < \pi/2$ - es interessieren für die Grenzwertfrage nur kleine Werte von x , wir müssen allerdings dann noch den Fall $x < 0$ betrachten. Wir schätzen wir ab, dass das rechtwinklige Dreieck mit den Kathetenlängen $\cos(x)$, $\sin(x)$ einen Flächeninhalt höchstens so groß hat wie das Kreisflächensegment mit Winkel x , dies wieder höchstens so groß ist wie das den Bogen von außen erfassende Dreieck mit den Kathetenlängen 1, $\tan(x)$. Also (alle verglichenen Zahlen > 0):

$$\begin{aligned} \frac{1}{2} \sin(x) \cos(x) &\leq \frac{x}{2\pi} \pi \leq \frac{1}{2} \tan(x), \text{ also} \\ \frac{1}{\sin(x) \cos(x)} &\geq \frac{1}{x} \geq \frac{\cos(x)}{\sin(x)}, \text{ es folgt} \\ \frac{1}{\cos(x)} &\geq \frac{\sin(x)}{x} \geq \cos(x). \end{aligned}$$

Für $x < 0$ gilt aber die letzte Ungleichung auch, wegen der Symmetrien der beteiligten Funktionen! Nun ist \cos eine stetige Funktion an der Stelle $x_0 = 0$, ebenso $1/\cos$ an dieser Stelle stetig. Also $\lim_{x \rightarrow 0} \cos(x) = \cos(0) = 1$, $\lim_{x \rightarrow 0} (1/\cos(x)) = 1$. Da nun die Werte $\sin(x)/x$ zwischen zwei eingeschlossen sind, die beide gegen 1 gehen, stimmt der behauptete Grenzwert.

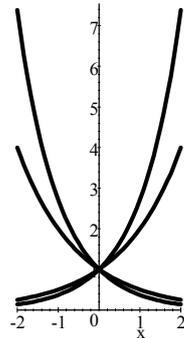
1.4. Exponential- und Logarithmusfunktionen, Potenzfunktionen. Wir führen den zweiten wichtigen Typ von Grundfunktionen ein, die sich nicht als gebrochen rationale Funktionen und nicht einmal als algebraische Funktionen darstellen (dabei lässt man noch Wurzelfunktionen im allgemeinsten Sinne als Bausteine zu, nicht nur Ausdrücke x^n mit ganzem Exponenten). Für sie benötigen wir den Rechenausdruck a^x mit beliebigen $a > 0$, $x \in \mathbb{R}$. Dazu stelle man sich vor, dass jede reelle Zahl Grenzwert einer Folge rationaler Zahlen ist. Für rationale Zahlen $q, r \in \mathbb{Q}$, $q > 0$, kennen wir bereits q^r , z.B. ist $(2/3)^{-3/5} = 1/\sqrt[5]{(2/3)^3} = \sqrt[5]{27/8}$. Sei nun (q_i) , alle $q_i > 0$, eine Folge rationaler Zahlen, welche gegen q konvergiert, also Grenzwert q hat, $q > 0$. Sei weiter (r_i) eine Folge rationaler Zahlen, die gegen r konvergiert. Dann hat die Folge $(q_i^{r_i})$ einen Grenzwert (der ist stets eindeutig, wenn er existiert), und dieser Grenzwert wird als Wert des Ausdrucks q^r definiert. Wir wollen hier nur diese Vorstellung erzeugen, ohne die Existenz des benötigten Grenzwerts zu zeigen. (Stattdessen verweisen wir darauf, dass wir im Rahmen unseres Aufbaus auch später die Integrale heranziehen könnten, um erst einmal \ln zu definieren, mit der Umkehrfunktion \exp dann den Exponentialausdruck a^x , $a > 0$ definieren könnten.)

DEFINITION 27. Die Exponentialfunktionen sind die folgenden Funktionen \exp_a , $a \in \mathbb{R}$, $a > 0$, $a \neq 1$:

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

Speziell gibt es eine einzige Zahl, die Eulerzahl e , für welche die Funktion \exp_e die Eigenschaft hat: $\exp'_e = \exp_e$. Diese Funktion nennt man \exp ohne Index. Ihr Rechenausdruck ist e^x .

Man hüte sich vor dem folgenschweren Anfängerfehler, a^x mit x^a zu verwechseln (letzterer Ausdruck definiert die Potenzfunktionen, s.u.) Hier sind die Graphen einiger Exponentialfunktionen, und zwar zu den Basen $e, 2, 1/e, 1/2$ - generell beachte man: $a^{-x} = (1/a)^x$:



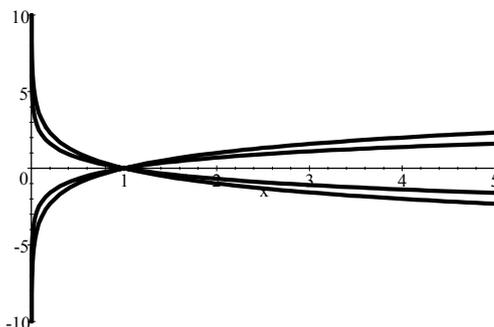
Die wesentlichen Eigenschaften für Basen $a > 1$: Alle Werte sind > 0 . Sie wachsen *sehr schnell* gegen Unendlich für $x \rightarrow \infty$, entsprechend fallen sie schnell Null für $x \rightarrow -\infty$; denn $a^{-x} = 1/a^x$. Entsprechend symmetrische Aussagen hat man für $0 < a < 1$. Stets ist $a^0 = 1$, $a > 0$. Nur die Kurve zur (wegen dieser Einfachheit *natürlich* genannten) Basis $e \approx 2.71$ hat Steigung 1 bei $x = 0$. e ist nicht rational und nicht einmal Lösung einer Polynomgleichung mit rationalen Koeffizienten (wie z.B. $\sqrt{2}$). Man nennt daher e eine transzendente Zahl. (π ist auch eine solche Zahl.) Die strenge Monotonie der Exponentialfunktionen lässt ihre Umkehrfunktionen definieren, die Logarithmusfunktionen:

DEFINITION 28. Die Logarithmusfunktionen zu $a > 0$, $a \neq 1$ sind die Funktionen $\log_a = (\exp_a)^{-1}$ (nicht mit $1/\exp_a$ zu verwechseln!), konkret zu definieren durch

$$\log_a :]0, \infty[\rightarrow \mathbb{R} \\ x \mapsto \text{die Lösung } y \\ \text{der Gleichung } a^y = x.$$

Speziell bezeichnet man \log_e kurz mit \ln . (In Computerprogrammen heißt diese Funktion allerdings meistens „log“.)

Man mache sich zunächst einmal ein paar einfachste Beispiele klar wie $\log_2(32) = 5$, $\log_2(1/32) = -5$, $\log_{1/2}(32) = -5$, $\log_{1/2}(1/32) = 5$, $\log_{10}(10000) = 4$, $\log_{10}(0.0001) = -4$. Wie bereits bekannt ergeben sich die Graphen der Umkehrfunktionen durch Spiegelung an der Geraden $y = x$. Das macht sofort den Pol der Logarithmusfunktionen bei $x = 0$ und das *sehr langsame* Wachsen nach Unendlich in den Fällen $a > 1$ klar (für $0 < a < 1$ entsprechendes Fallen - es gilt $\log_{1/a} = -\log_a$). Dennoch seien die Graphen noch einmal gezeigt, und zwar für dieselben Exponenten $e, 2, 1/e, 1/2$ der oben gezeigten Exponentialfunktionen (bei den Logarithmusfunktionen bedeutet kleinerer Basisbetrag größere Steigung):



Es folgen die wichtigen Grundformel für Exponential- und Logarithmusfunktion - wir schreiben sie bequem nur für die natürliche Basis auf, sie gelten jedoch für alle Basen. Benötigt man einmal eine andere Basis, so kann man die zugehörige Funktion mittels der letzten Formeln sofort auf natürliche umschreiben:

$$\begin{array}{ll}
 \text{Seien } x, y \in \mathbb{R}, a > 0. & \text{Seien } a, b > 0, y \in \mathbb{R}. \\
 e^0 = 1 & \ln(1) = 0 \\
 e^{-x} = \frac{1}{e^x} & \ln\left(\frac{1}{a}\right) = -\ln(a) \\
 e^{x+y} = e^x \cdot e^y & \ln(ab) = \ln(a) + \ln(b) \\
 (e^x)^y = e^{x \cdot y} & \ln(a^y) = y \cdot \ln(a) \\
 a^x = e^{x \ln(a)} \quad (a \neq 1) & \log_a(b) = \frac{\ln(b)}{\ln(a)} \quad (a \neq 1)
 \end{array}$$

Es sollte beobachtet werden, dass die einander gegenüber stehenden Formeln jeweils dasselbe aussagen, nur einmal mit exp und einmal mit ln formuliert. Dazu ist lediglich die Umkehrfunktionseigenschaft heranzuziehen, welche in Formeln besagt: $\ln(e^x) = x$ ($x \in \mathbb{R}$), $e^{\ln(x)} = x$ ($x > 0$). Betrachten wir als Beispiel das dritte Formelpaar und setzen die linke Version voraus. Seien nun $a, b > 0$, dann setzen wir $x = \ln(a)$, $y = \ln(b)$, also haben wir $\ln(a) + \ln(b) = x + y$. Weiter gilt nach Setzung von x, y : $a = e^x, b = e^y$, also $\ln(ab) = \ln(e^x e^y) = \ln(e^{x+y}) = x + y$, mit Einsatz der linksseitigen Formel. Daher $\ln(ab) = \ln(a + b)$, und die rechtsseitige ist bewiesen. Die linksseitigen Formeln sind nichts anderes als die bekannten Rechengesetze für das Potenzieren, wir müssen allerdings voraussetzen, dass sie tatsächlich beim Übergang von rationalen Basen und Exponenten auf beliebige reelle (Basen positiv) bestehenbleiben, also bei den oben beschriebenen Grenzprozessen. Nun wollen wir allerdings hier - auch im Folgenden - generell ohne Beweis voraussetzen, dass die Grenzwertbildung mit den Grundrechenarten vertauschbar ist. Andererseits sei verwiesen darauf, dass die Integraldefinition der Logarithmusfunktion ln ohne weiteres den Beweis der obenstehenden Logarithmusformeln (vgl. weiter unten die Grundintegrale) erlaubt, aus denen man dann diejenigen für die Exponentialfunktion wie beschrieben herleiten kann.

Eine unmittelbare Anwendung der Exponentialfunktionen in der Naturbeschreibung: Exponentielles Wachstum. Das Charakteristikum der Exponentialfunktionen ist dies: Bei festen Differenzen der Werte der unabhängigen Variablen gibt es in der abhängigen stets eine Vervielfachung mit einem festen Faktor $> 0, \neq 1$. Das können wir sofort nachrechnen; sei $a > 0, a \neq 1$, dann gilt: $f(t + \Delta) = a^{t+\Delta} = a^t \cdot a^\Delta = f(t) \cdot a^\Delta$. Addiert man also Δ zu beliebigem Wert von t , so multipliziert sich der Funktionswert mit a^Δ , und zwar unabhängig davon, von welcher Stelle t man ausgeht (!). Umgekehrt kann man stets zu vorgegebenem Vervielfachungsfaktor k fragen, wie groß Δ sein muss, damit Addition von Δ zur Vervielfachung mit k führt: $a^\Delta = k$ ist aufzulösen nach Δ : $\Delta = \ln(k)/\ln(a)$. Für den Fall, dass $a > 1$ und $k < 1$ (oder umgekehrt), bedeutet dies natürlich einen negativen Wert von Δ . Schließlich sollten wir noch sehen, wie man bei Vorgabe der Standard-Daten für ein exponentielles Wachstum zu einer Funktionsbeschreibung kommt mittels linearer Transformation einer Exponentialfunktion: Eine zeitabhängige Größe $q(t)$ vervielfache sich mit Faktor k in der Zeitspanne d , und zur Zeit t_0 habe sie den Wert q_0 . Wie sieht $q(t)$ zu beliebiger Zeit t aus (vorausgesetzt, es handle sich um

eine „ordentliche differenzierbare Funktion“)? Wir modifizieren den Grunda Ausdruck k^t (k ist offenbar passende Basis für den Faktor der Vervielfachung) und schreiben dann auf natürliche Basis um:

$$q(t) = q_0 k^{(t-t_0)/d} = q_0 e^{\ln(k)(t-t_0)/d}.$$

Selbstverständlich beschreibt dies für $d > 0$ und $0 < k < 1$ ein exponentielles Fallen, das im Begriff „exponentielles Wachstum“ mitgemeint sein sollte. Beispiele: Radioaktiver Teilchenzerfall, Amplitudenverhalten bei gedämpfter Schwingung, Aufladevorgang bei einem Kondensator.

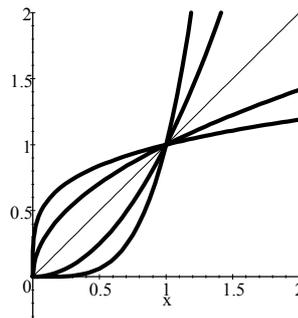
Abschließend erwähnen wir noch als dritten Block von komplizierteren Grundfunktionen die Potenzfunktionen:

DEFINITION 29. Sei $a \in \mathbb{R}$, $a > 0$. Die Potenzfunktion zur Basis a ist dann die Funktion

$$p_a : \begin{array}{ll} [0, \infty[& \rightarrow [0, \infty[\\ x & \mapsto x^a \end{array}$$

(p_a ist kein allgemein üblicher Name, einen solchen gibt es nicht.) Man beachte, dass die unabhängige Variable in der Basis steht und gerade nicht im Exponenten. Man beachte, dass jede Zahl $a > 0$ als Exponent in Frage kommt, also nicht etwa nur die natürlichen. Insbesondere sind die Wurzelfunktionen mit Ausdrücken wie $\sqrt[3]{x} = x^{1/3}$ usw. dabei.

Die Eigenschaften sind klar: Alle diese Funktionen sind streng monoton steigend, alle sind umkehrbar, $p_{1/a}$ ist die Umkehrfunktion von p_a , weil $(x^a)^{1/a} = x^{a \cdot 1/a} = x^1 = x$. Beachtlich ist die unendliche Steigung in $x = 0$ bei den Funktionen p_a mit $0 < a < 1$. Hier sind die Graphen für $a = 1/4, 1/2, 1, 2, 4$:



Es sollte klar sein, dass sich unter diesen Funktionen mit $a = 1/4$ die stärkste Steigung nahe Null ergibt, die schwächste im Bereich $x > 1$, mit $a = 4$ die schwächste Steigung nahe Null, die stärkste für $x > 1$.

1.5. Der rekursive Aufbau eines geeigneten Vorrats von Funktionen aus Grundfunktionen.

Wir können bemerken, dass bereits einige wichtige Grundphänomene der Natur durch lineare Transformationen von unseren Grundfunktionen beschrieben werden können: Freier Fall durch $s(t) = gt^2/2$, harmonische Schwingungen durch $y(t) = \alpha + \beta \sin(\omega t + \varphi)$, exponentielles Wachsen oder Fallen durch $q(t) = q_0 e^{(t-t_0)/d}$. Aber damit kommt man nicht aus. Wichtige Vorgänge wie gedämpfte Schwingung verlangen komplexere Zusammensetzungen wie $y(t) = e^{-kt} \beta \sin(\omega t + \varphi)$. Die wichtigste Funktion der Statistik ist die Dichte der Standard-Normalverteilung, die lautet $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$. Wir werden uns daher zweckmäßig klarmachen, was die benötigten Funktionsbildungsprozesse bzw. Verknüpfungen von Funktionen zu neuen Funktionen sind. Einen wichtigen Prozess werden wir hier nicht ausdrücklich besprechen, sondern nur auf seine Bedeutung hinweisen: Aus unendlich vielen Funktionen f_i , $i \in \mathbb{N}$, kann man unter geeigneten Konvergenzbedingungen die neue Funktion $g(x) = \sum_{i=0}^{\infty} f_i(x)$ (unendliche Reihe) machen, so gelingt beispielsweise die Darstellung einer „Sägezahnfunktion“ durch Bilden von $\sum_{n=1}^{\infty} \sin(nx)/n$. Weitere wichtige Funktionen lassen sich über Integrale definieren. Bevor wir auf die Verknüpfungen von Funktionen zu sprechen kommen, wollen wir noch einmal

zusammenstellen, was wir hier als Vorrat von Grundfunktionen betrachten wollen. (Es sei nur bemerkt, dass man diesen Bereich bei Bedarf jederzeit aufstocken kann, zum Beispiel mittels spezieller in Formelsammlungen aufgenommenen Funktionen).

1.5.1. *Ein für unsere Zwecke ausreichender Vorrat von Grundfunktionen.*

DEFINITION 30. *Unsere Grundfunktionen sind im engeren Sinne nur: Die Funktion mit konstantem Wert 1, also $f(x) = 1, x \in \mathbb{R}$, dazu die Identität, $id(x) = x, x \in \mathbb{R}$, schließlich \sin und \exp . Im weiteren Sinne werden wir alle konstanten Funktionen, alle besprochenen trigonometrischen Funktionen (\sin, \cos, \tan und Arcusfunktionen), alle Potenzfunktionen sowie \exp und \ln als Grundfunktionen betrachten.*

Die praktische Nutzenanwendung ist diese: Man kennt (d.h.: hat im Kopf!) die wesentlichen Eigenschaften der Grundfunktionen im weiteren Sinne (Graphen, später auch Ableitung und Integral). Weiter kennt man Regeln, nach denen (für Aspekte des graphischen Verhaltens, für Ableitungen und Integrale) aus den Eigenschaften der Bestandteil-Funktionen auf die einer zusammengesetzten Funktion zu schließen ist. Man produziert also aus dem Wissen über die Grundfunktionen mit der Handhabung der Zusammensetzungsregeln all das, was man schließlich benötigt.

1.5.2. *Die grundlegenden Operationen, welche aus Funktionen neue Funktionen machen.*

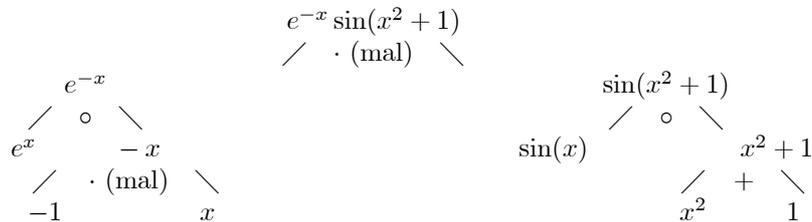
DEFINITION 31. *Die elementaren Verknüpfungen von Funktionen zu neuen sind die folgenden:*

Aus $f : D \rightarrow \mathbb{R}, \alpha \in \mathbb{R} :$	$\alpha f : D \rightarrow \mathbb{R}$ $x \mapsto \alpha f(x)$	Multiplikation einer Funktion mit einer Zahl
Aus $f, g : D \rightarrow \mathbb{R} :$	$f \pm g : D \rightarrow \mathbb{R}$ $x \mapsto f(x) \pm g(x)$	Addition (Subtraktion) von zwei Funktionen
Aus $f, g : D \rightarrow \mathbb{R} :$	$f \cdot g : D \rightarrow \mathbb{R}$ $x \mapsto f(x) \cdot g(x)$	Multiplikation zweier Funktionen
Aus $f, g : D \rightarrow \mathbb{R} :$	Sei $E = \{x \in D \mid g(x) \neq 0\}$ $\frac{f}{g} : E \rightarrow \mathbb{R}$ $x \mapsto f(x)/g(x)$	Quotient zweier Funktionen
Aus $f : D \rightarrow E \subseteq \mathbb{R},$ $g : E \rightarrow \mathbb{R} :$	$g \circ f : E \rightarrow \mathbb{R}$ $x \mapsto g(f(x))$	Hintereinander- schaltung von Funktionen
Aus $f : D \rightarrow E \subseteq \mathbb{R},$ f umkehrbar	$f^{-1} : E \rightarrow D$ $x \mapsto \text{das } y \in D$ mit $f(y) = x$	Umkehrfunktion einer umkehrbaren Funktion

Die schwierigste Verbindung ist die der Hintereinanderschaltung oder Verkettung $g \circ f$, lies „ g hinter f “. Der entsprechende Ausdruck wird gebildet, indem man den Rechenausdruck für f in den von g einsetzt, also z.B. für $f(x) = x + x^2, g(x) = x \cdot \cos(x) : g(f(x)) = (x + x^2) \cdot \cos(x + x^2)$. Insbesondere ist auf etwa benötigte Klammern zu achten. Dagegen wäre im Beispiel $f(g(x)) = x \cos(x) + (x \cos(x))^2$. Man beachte und nutze Schreibweisen wie $\ln^2(x) = \ln(x) \cdot \ln(x) = (\ln(x))^2$ usw.

1.5.3. *Das richtige strukturelle Lesen von Rechenausdrücken.* Rechenausdrücke für Funktionen sollten niemals „von rechts nach links“ gelesen werden, sondern stets im Sinne ihres rekursiven Aufbaus aus den Ausdrücken von Grundfunktionen wie im folgenden Beispiel rekursiv abgebaut werden, bis man bei Ausdrücken von Grundfunktionen angelangt ist. Es sei betont, dass es keinen Sinn macht, Rechenausdrücke mit \circ zu verknüpfen, gemeint ist im folgenden Beispiel an der entsprechenden Stelle, dass die Funktion \sin hinter die Funktion f

geschaltet wird, deren Rechenausdruck $x^2 + 1$ ist. Wir arbeiten lediglich mit den Rechenausdrücken allein, um die Einführung einer langen Reihe von Namen für die zugehörigen Funktionen zu sparen.



Wir haben im Beispiel Wert darauf gelegt, zu zeigen, dass man tatsächlich bis zu unseren Grundfunktionen zurückkommt. Man sollte bemerken, wie man im Beispiel auch bis zu den Grundfunktionen im engeren Sinne noch gelangen kann. In der Praxis wird man sich stets mit einem Abbau begnügen, der gerade so weit geht, dass man von den Bestandteilen, auf die man reduziert hat, die gewünschten Eigenschaften kennt, den Graphen also oder die Ableitung etwa. Dann nutzt man die Zusammensetzungsregeln, um den Baum hochzuwandern.

1.5.4. *Das Zusammensetzen von Graphen.* Im vorigen Beispiel würde man sich den Graphen zu $e^{-x} \sin(x^2 + 1)$ etwa so klarmachen: Aus den Graphen zu $x^2 + 1$ und \sin setzt man den zu $\sin(x^2 + 1)$ zusammen: Das ist eine gerade Funktion, die eine zu beiden Seiten immer schneller werdende Schwingung darstellt, zwischen Werten $-1, 1$. Nun setzt man aus diesem Graphen und dem zu e^{-x} , den wir schon kennen, den Graphen zum gesamten Ausdruck $e^{-x} \sin(x^2 + 1)$ zusammen, indem man sich die Werte multipliziert denkt: Da e^{-x} stets positiv ist, wird das Vorzeichen überall vom Anteil $\sin(x^2 + 1)$ bestimmt. Der Faktor e^{-x} moduliert offenbar die Amplitude: Sie geht steil nach unendlich für $x \rightarrow -\infty$, schnell nach Null für $x \rightarrow \infty$. Die Kurven zu $\pm e^{-x}$ treten als Hüllkurven auf. Beim Ableiten werden wir gerade so rekursiv vorgehen: Aus den Ableitungen zu $\sin(x)$, $x^2 + 1$ werden wir die zu $\sin(x^2 + 1)$ zusammenbasteln, aus dieser und der zu e^{-x} dann die Ableitung von $e^{-x} \sin(x^2 + 1)$.

Bemerkung: Man sollte die elementaren Eigenschaften von Funktionen wie Symmetrie, Monotonie usw. so gut verstanden haben, dass man gewisse Schlüsse direkt ziehen kann der Form: „Wenn f auf D monoton steigend und g auf $f(D) = \{f(x) \mid x \in D\}$ monoton fallend ist, so ist $g \circ f$ auf D monoton fallend“. Dagegen macht es keinen Sinn, hundert Regeln dieser Form auswendig zu lernen. Man „sieht so etwas“ routinemäßig, sobald man die Graphen einiger 50 nicht so banaler Rechenausdrücke einmal überlegt hat. Schon gar nicht erwarte man, dass ein geschlossenes Schema anzugeben sei, das man nur schematisch befolgen müsse, um jeden Funktionsgraphen zur Gänze richtig zu erkennen. Ein solches Schema kann man nur für lächerliche Klassen von Funktionen geben, etwa „alle Polynomfunktionen vom Grade höchstens drei“. Unter „Kurvendiskussion“ wollen wir daher sogleich korrigierend verstehen: Das intelligente Zusammensetzen komplexerer Funktionsgraphen aus einfacheren. Damit kommt man bei flexiblem Einsatz von Beobachtungen mit der Konzentration auf folgende Fragen, die man als leitende stets stelle, bereits sehr weit!

Leitende Fragen zur Kurvendiskussion:

Definitionsbereich? Standardsymmetrie? Welche Werte kommen vor?
Vorzeichen der Werte (stückweise)?
Monotonie (stückweise)?
Lokale Dominanz eines Teilausdrucks (lokal bei $\pm\infty$ ist das die Frage nach dem Verhalten für $x \rightarrow \infty$, $x \rightarrow -\infty$, auch das Entwickeln eines Pols fällt darunter).

Zur Dominanz kann man sich ein paar praktisch sehr nützliche Regeln merken:

Dominanzregeln:

Für $x \rightarrow \infty$ dominiert:
die Exponentialfunktion über alle Potenzfunktionen,
alle Potenzfunktionen über \ln ,
jede Potenzfunktion mit höherem Exponenten über
eine solche mit niedrigerem Exponenten, also:

$$(a > 0) : \lim_{x \rightarrow \infty} \frac{x^a}{e^x} = 0, \quad \lim_{x \rightarrow \infty} \frac{\ln(x)}{x^a} = 0, \quad \lim_{x \rightarrow \infty} \frac{x^a + \text{niedere}}{x^b + \text{niedere}} = 0 \text{ für } b > a.$$

Diese Regeln werden wir später im Rahmen eines wichtigen Typs von Anwendung der Ableitung (de L'Hospitalsche Regeln) rechtfertigen können. Man sieht bereits, dass sie nicht eben so selbstverständlich sind wie: „Wenn $f(x)$ und $g(x)$ beide nach ∞ gehen für $x \rightarrow \infty$, so geht $f(x) \cdot g(x)$ ebenfalls nach ∞ für $x \rightarrow \infty$ “. Weiter sollte man auch diese Regeln mit einer gewissen Flexibilität handhaben, die in ihnen steckt: Unmittelbar wörtlich ist hier keine Aussage über z.B. das Verhalten von xe^x für $x \rightarrow -\infty$ gemacht, aber man schließt leicht: $\lim_{x \rightarrow -\infty} xe^x = \lim_{x \rightarrow \infty} (-x)e^{-x} = -\lim_{x \rightarrow \infty} (x/e^x) = 0$. Ähnlich folgt der Grenzwert von $x \ln(x)$ für $x \rightarrow 0$.

Bemerkung: Man kann nicht erwarten, dass man *alle* Fragen zu einem Funktionsgraphen mit dem bisher Besprochenen beantworten kann. Dazu gehören insbesondere *quantitative* Fragen nach der genauen Lage von Extremwerten, aber auch *feinere* Fragen zum *qualitativen* Steigungsverhalten einer Funktion. Dazu werden wir dann die Ableitung einsetzen können. Zum Beispiel steigt $x^{2/3}$ langsamer als eine Gerade nach Unendlich, dagegen $(x^{2/3})^2 = x^{4/3}$ schneller als eine Gerade. Wie steht es mit $\ln^2(x)$? Die Ableitung gibt sofort Auskunft darüber: Langsamer als eine Gerade.

2. Ableitung (Differentiation) von Funktionen

2.1. Die Idee der Ableitung. Zwei verschiedene Ideen führen unmittelbar zu ein und demselben Begriff der Ableitung, die geometrische Vorstellung einer Tangente an einen Funktionsgraphen in einem vorgeetzten Punkte sowie die Absicht, in einem Punkt eine optimale Näherung einer nichtlinearen Funktion durch eine lineare vorzunehmen. Erstere Idee ist allgemeiner bekannt als Vorwissen, letztere ist hingegen die systematisch wichtigere.

- Die erstere Idee: Eine Funktion f sowie der Punkt $(x_0, f(x_0))$ seien vorgegeben, der Graph von f verlaufe glatt in diesem Punkt. Dann sollte es eine Tangente an den Graphen von f im Punkt $(x_0, f(x_0))$ geben, und wir fragen, wie man deren Steigung ausrechnen kann. Die geometrische Idee dazu ist es, einfach eine Folge von Punkten $(x_i, f(x_i))_{i \in \mathbb{N} \setminus \{0\}}$ zu nehmen, wobei $(x_i)_i$ gegen x_0 konvergiert - sonst ist die Folge beliebig, und dann eine Folge von Sekanten durch die zwei Punkte $(x_0, f(x_0)), (x_i, f(x_i))$ zu bestimmen, die geometrisch gegen die Tangente konvergieren sollte, so dass sich deren Steigung demnach als Grenzwert der Sekantensteigungen ergibt (zur Illustration vgl. die Abbildung, nachdem der Gebrauch von „ Δx “ eingeführt ist.)

Damit man überhaupt solch beliebige Sekantenfolgen bilden kann, so setzen wir voraus, dass f in einer beidseitigen Umgebung von x_0 definiert sei.

Da es auf die Folge (x_i) nicht ankommt, so sagt man, man lasse x gegen x_0 gehen und betrachte dafür den Grenzwert der Sekantensteigungen, also:

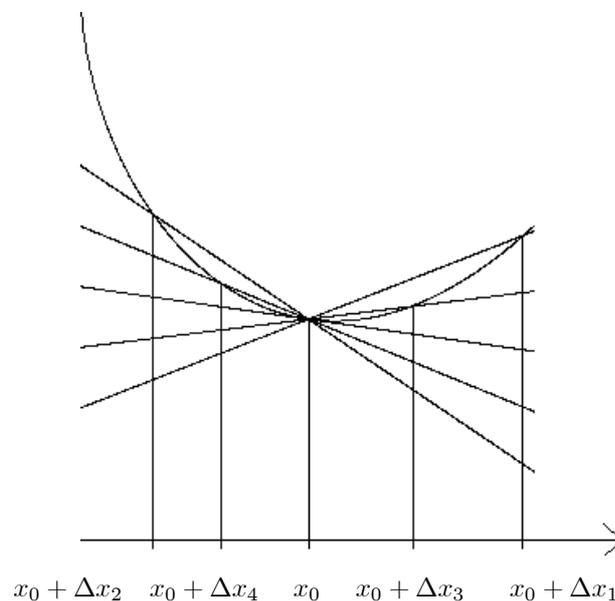
$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0}$$

Dies schreibt man lieber anders, da die Differenz $x - x_0$ gerade die systematisch wichtige Sache ist (nenne sie Δx - man lese das auf keinen Fall „Delta mal x“, sondern „Delta x“; das wird wie ein einziger Buchstabe gebraucht, eine untrennbare Einheit, gerade so wie bei Indizes! Insbesondere bedeutet Δx^2 , dass die Zahl Δx

quadriert wird, es bedarf dazu keiner Klammer wie $(\Delta x)^2$. Gemeint ist inhaltlich immer irgendeine kleine Differenz von x -Werten):

$$\lim_{\Delta x \rightarrow 0 \ (\Delta x \neq 0)} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Dieser Grenzwert ergibt nicht nur im Falle, dass der Graph von f in x_0 glatt verläuft, die richtige Tangentensteigung, sondern zusätzlich bildet seine Existenz das adäquate Kriterium von Glattheit, also: Der Grenzwert der Sekantensteigungen existiert genau dann, wenn der Graph von f in x_0 glatt ist. Hier sieht man, wie die Tangente als Grenzsekante entsteht (man stelle sich $\Delta x_1, \Delta x_2, \Delta x_3, \Delta x_4 \dots$ als Folge vor, die gegen Null geht - die Sekanten gehen dann gegen die Tangente):



Insgesamt landen wir bei folgender Definition:

DEFINITION 32. Eine Funktion f sei in einer Umgebung von x_0 definiert. f heißt differenzierbar in x_0 genau dann, wenn der Grenzwert

$$\lim_{\Delta x \rightarrow 0 \ (\Delta x \neq 0)} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

existiert. Dessen (eindeutig bestimmter) Wert heißt dann $f'(x_0)$, „Ableitung von f an der Stelle x_0 “. Die Ableitungsfunktion von f ist dann die Funktion $x \mapsto f'(x)$, definiert an allen Stellen x , wo f differenzierbar ist. Es sei noch einmal die Formel hervorgehoben:

$$(2.1) \quad f'(x_0) = \lim_{\Delta x \rightarrow 0 \ (\Delta x \neq 0)} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \quad (\text{wenn existent}).$$

Beispiel: Nehmen wir einmal für die Funktion $f(x) = x^2$ an der Stelle $x_0 = 3$ folgende konkrete Folge von Δx -Werten und geben dazu die Sekantensteigungen:

Δx	$\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x} = \frac{(3+\Delta x)^2-3^2}{\Delta x}$
0.1	6.1
0.01	6.01
0.001	6.001
0.0001	6.0001
0.00001	6.00001

Man wird hier *empirisch* bereits einsehen, dass $f'(x_0)$ in unserem Falle den Wert 6 haben wird. *Mathematisch* freilich muss man beweisen, dass das wirklich der Grenzwert ist. Nun, im Beispiel ist das wirklich leicht:

$$\frac{(3 + \Delta x)^2 - 3^2}{\Delta x} = \frac{6\Delta x + (\Delta x)^2}{\Delta x} = 6 + \Delta x \quad (\rightarrow 6 \text{ für } \Delta x \rightarrow 0, \text{ klar!})$$

(Hier immer $\Delta x \neq 0$ für die Folgenbildung!) Sogar allgemein für jedes x_0 anstelle von 3 ist das einfach:

$$\frac{(x_0 + \Delta x)^2 - 3^2}{\Delta x} = \frac{2x_0\Delta x + (\Delta x)^2}{\Delta x} = 2x_0 + \Delta x \quad (\rightarrow 2x_0 \text{ für } \Delta x \rightarrow 0)$$

Wir sehen damit, dass für unsere Quadratfunktion $f(x) = x^2$ gilt: $f'(x) = 2x$, für alle $x \in \mathbb{R}$.

- Die zweite Idee: Lokale lineare Approximation einer Funktion, Näherung 1. Ordnung

Diese Idee gibt mehr inhaltlichen Hintergrund, ihr Gesichtspunkt ist überhaupt ein sehr allgemeiner und bedeutender, für theoretische sowie für Anwendungszwecke. Schließlich werden wir sehen, dass diese Idee viel verallgemeinerungsfähiger ist, und zwar sowohl hinsichtlich der Dimensionen (\rightarrow Funktionen mehrerer unabhängiger Variablen) als auch hinsichtlich der Fortsetzung im Grade der Näherung, die bis hin zur exakten Reihendarstellung komplizierter Funktionen führt. Um gerade diese letztere Fortsetzung zu verdeutlichen, beginnen wir mit einer Vorstufe, die noch vor der Ableitung liegt, der Näherung 0. Ordnung.

Stellen wir uns vor, dass man eine komplizierte Funktion hat, sagen wir z.B. $\exp(x) = e^x$, und dass man an einer Stelle x_0 ihren Wert kenne, im Beispiel wäre das bei $x_0 = 0$ der Fall. Nun wollen wir wissen, was $\exp(x_0 + \Delta x)$ für kleine $|\Delta x|$ ist, und es genüge uns eine gute Näherung. Für eine in x_0 stetige Funktion (das bedeutet ja gerade: Unterscheidet sich x nur wenig von x_0 , so unterscheidet sich auch $f(x)$ nur wenig von $f(x_0)$) liegt es nahe, zu sagen: $f(x_0 + \Delta x) \approx f(x_0)$. Tatsächlich klappt das auch im Beispiel so ordentlich, wie man das erwarten kann:

Δx	$e^{\Delta x}$	absolute Differenz zum Näherungswert $e^0 = 1$
0.1	1.1052	0.1052
0.01	1.0101	0.0101
-0.01	0.99005	0.00995

Fassen wir zusammen:

Näherung 0. Ordnung einer in x_0 stetigen Funktion in einer Umgebung von x_0

$$f(x_0 + \Delta x) \approx f(x_0),$$

geschrieben als Gleichung mit Rest- oder Fehlerglied:

$$f(x_0 + \Delta x) = f(x_0) + R_{f,x_0}(\Delta x).$$

Die Näherung geschieht mit einem Polynom 0. Grades (d.h. einfach mit einer Konstanten), daher „0. Näherung“. Wenn die Funktion f stetig in x_0 ist, so erfüllt das Restglied folgende Bedingung:

Restgliedbedingung (oder Fehlerbedingung) 0. Ordnung:

$$|R(\Delta x)| \rightarrow 0 \text{ für } \Delta x \rightarrow 0.$$

(Hier darf auch die konstante Nullfolge als Δx -Folge genommen werden. (Konsequenz: $R(0) = 0$.) Gleichwertig könnte man auch Stetigkeit von R in 0 fordern (mit der Konsequenz $R(0) = 0$) und sich für die Δx -Folgen mit $\Delta x \neq 0$ begnügen.)

Umgekehrt: Fordert man diese in sich vernünftige Bedingung für den auftretenden Rest R bei *irgendeiner* Näherung 0. Ordnung um x_0 bei irgendeiner Funktion f ,

$$f(x_0 + \Delta x) = c + R(\Delta x),$$

so folgt daraus:

- (i) f ist stetig in x_0
- (ii) $c = f(x_0)$

Denn laut Restbedingung hat man für die konstante Folge $0, 0, \dots$ als Δx -Folge (oder gleich mit $R(0) = 0$): $f(x_0) = c$, da $R(0) = 0$, und für $|\Delta x|$ klein genug hat man $|R(\Delta x)|$ so klein, wie man will, also $f(x_0 + \Delta x)$ so nahe bei c , wie man will. Das bedeutet aber genau die Stetigkeit von f in x_0 .

Eine vernünftige Näherung (welche die entsprechende Restbedingung erfüllt) 0. Ordnung von f um x_0 existiert also genau dann, wenn f in x_0 stetig ist. Zusätzlich ist die Näherung eindeutig bestimmt, die Konstante muss gerade $f(x_0)$ sein.

Wir kommen nunmehr zur Näherung 1. Ordnung. Nach unserer Vorbereitung sollte der Ansatz klar sein: Wir wählen zur Näherung ein Polynom 1. Grades, also:

$$f(x_0 + \Delta x) = a + b \cdot \Delta x + R(\Delta x).$$

Das mag etwas befremden, da die unabhängige Variable x (links als $x_0 + \Delta x$ geschrieben) auch als unabhängige Variable der linearen Funktion rechts auftreten sollte, so dass man $\alpha + \beta(x_0 + \Delta x)$ erwartet. Aber $\alpha + \beta(x_0 + \Delta x) = (\alpha + \beta x_0) + \beta \Delta x$, also kann man das ohne weiteres als lineare Funktion in Δx umschreiben.

Wir fragen uns zunächst, wie man die Restbedingung zu formulieren hätte. Man sollte erwarten, dass die Bedingung 0. Ordnung zu verschärfen wäre: Mit einem Polynom 1. Grades sollte man eine bessere Näherung hinbekommen als mit einer Konstanten. Mit der Bedingung 0. Ordnung erhält man sofort wie oben, dass $a = f(x_0)$ werden muss, und das setzen wir sofort in unseren Ansatz ein:

$$f(x_0 + \Delta x) = f(x_0) + b \cdot \Delta x + R(\Delta x).$$

Nun sieht man, dass die Bedingung 0. Ordnung weiter nichts über b hergibt: Mit in x_0 stetigem f ist diese Bedingung für *jeden* Wert von b erfüllt, da der Ausdruck $b\Delta x$ nach Null geht für $\Delta x \rightarrow 0$. Man braucht also eine schärfere Bedingung, um sinnvoll einen Wert von b zu bestimmen. Nun haben die Zahlen Δx kleine Beträge, wir denken an beliebig kleine Umgebungen von x_0 . Daher liegt es nahe, dass man für die (aufgeblasenen) Reste $R(\Delta x)/\Delta x$ ($\Delta x \neq 0$) fordert, dass sie gegen 0 gehen für $\Delta x \rightarrow 0$. Das liegt natürlich auch geometrisch nahe; denn für eine gute lokale Näherung eines Funktionsgraphen durch eine Gerade wird man erwarten, dass sich bei beliebiger Vergrößerung der Stelle immer stärker das Bild einer Geraden herauskristallisiert. Die Vergrößerung bedeutet aber, dass man die Fehler entsprechend durch Division durch Δx vergrößert. Daher formulieren wir folgendermaßen:

DEFINITION 33. **Tangentenzerlegung einer Funktion um eine Stelle x_0 und Restbedingung 1. Ordnung**

$$f(x_0 + \Delta x) = f(x_0) + b \cdot \Delta x + R(\Delta x)$$

heißt Tangentenzerlegung von f an der Stelle x_0 $f(x_0 + \Delta x) = f(x_0) + b \cdot \Delta x$ genau dann, wenn die folgende Restbedingung 1. Ordnung erfüllt ist:

$$\lim_{\Delta x \rightarrow 0, \Delta x \neq 0} \frac{R(\Delta x)}{\Delta x} = 0.$$

Wir wollen nun sehen, dass damit die Zahl b eindeutig bestimmt ist, wenn diese Restbedingung überhaupt erfüllbar ist durch irgendeine Wahl von b . Dafür brauchen wir nur die obenstehende Gleichung gleichwertig für $\Delta x \neq 0$ folgendermaßen umzuformulieren:

$$\frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = b + \frac{R(\Delta x)}{\Delta x}.$$

Lassen wir nun Δx gegen Null gehen, so kommt mit der Restbedingung unsere alte Definition der 1. Ableitung gemäß der ersten Idee:

$$b = \lim_{\Delta x \rightarrow 0, \Delta x \neq 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}.$$

Also $b = f'(x_0)$. Und ein Grenzwert ist eindeutig bestimmt, wenn er denn existiert. Denn in beliebig kleiner vorgebarbarer Umgebung dieses Wertes müssen von einem Index an alle Folgenglieder liegen, was offensichtlich nicht für zwei verschiedene Werte gelten kann, da man um sie hinreichend kleine Kreise ziehen kann, die sich nicht überschneiden.

Wir haben damit folgendes

Resultat: f besitzt eine Tangentenzerlegung an der Stelle x_0 genau dann, wenn die erste Ableitung von f an der Stelle x_0 existiert, und die Tangentenzerlegung lautet damit

$$(2.2) \quad f(x_0 + \Delta x) = f(x_0) + f'(x_0) \cdot \Delta x + R(\Delta x).$$

Das Fehlerglied, der Rest R , erfüllt dann die Bedingung

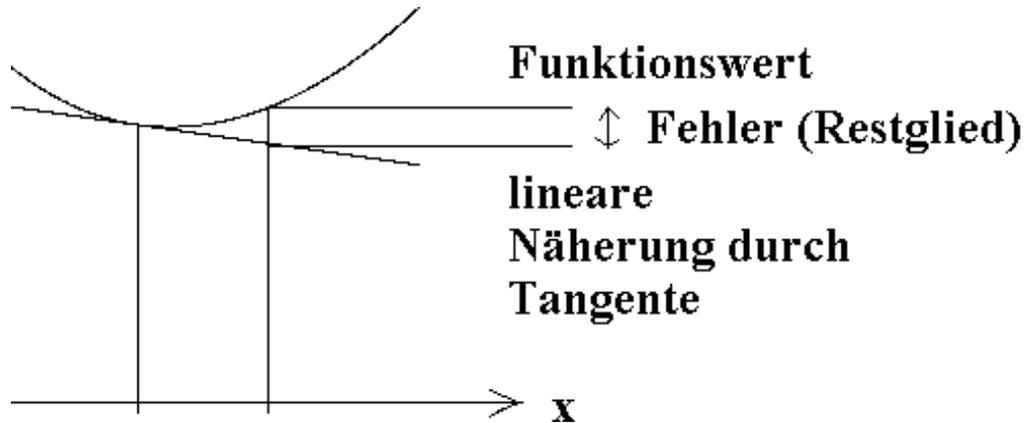
$$\lim_{\Delta x \rightarrow 0, \Delta x \neq 0} \frac{R(\Delta x)}{\Delta x} = 0, \text{ oder } \frac{R(\Delta x)}{\Delta x} \rightarrow 0 \text{ für } \Delta x \xrightarrow{\Delta x \neq 0} 0.$$

Die Näherung

$$(2.3) \quad f(x_0 + \Delta x) \approx f(x_0) + f'(x_0) \cdot \Delta x$$

heißt dann **Näherung 1. Ordnung**. (Sie existiert also nur im Falle der Differenzierbarkeit von f an der Stelle x_0 .)

Hier ist eine graphische Illustration dieser Idee:



x_0 $x_0 + \Delta x$
 Umgekehrt hat man auch: Wenn für eine Näherung durch ein Polynom 1. Grades der Gestalt

$$f(x_0 + \Delta x) = a + b \cdot \Delta x + R(\Delta x), \text{ für alle } \Delta x \text{ in einer Umgebung von Null,}$$

diese Restbedingung gilt und zusätzlich $R(0) = 0$, dann ist f in x_0 differenzierbar und insbesondere auch stetig, und es gilt $a = f(x_0)$, $b = f'(x_0)$. Denn mit der Gleichung für $\Delta x = 0$ hat man $f(x_0) = a + R(0) = a$ (nur dafür brauchen wir $R(0) = 0$). Nun folgert man wie oben geschehen, dass $f'(x_0)$ existiert und $b = f'(x_0)$. Die Stetigkeit folgt sofort aus der Differenzierbarkeit (ohne weitere Zusatzvoraussetzungen), da mit

$$\lim_{\Delta x \rightarrow 0, \Delta x \neq 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = f'(x_0)$$

insbesondere gilt:

$$\lim_{\Delta x \rightarrow 0, \Delta x \neq 0} (f(x_0 + \Delta x) - f(x_0)) = 0, \text{ oder } \lim_{\Delta x \rightarrow 0, \Delta x \neq 0} f(x_0 + \Delta x) = f(x_0),$$

was eine der Formulierungen der Stetigkeit von f in x_0 ist.

Beispiel:

Wir wollen nun einmal wiederum am Beispiel der Exponentialfunktion mit $x_0 = 0$ sehen, wie das funktioniert - wir haben ja bereits in die Definition der Zahl e hineingesteckt, dass $\exp'(0) = \exp(0) = 1$:

Die Tangentenerlegung für \exp an der Stelle $x_0 = 0$ lautet:

$$e^{0+\Delta x} = e^0 + e^0 \Delta x + R(\Delta x) = 1 + \Delta x + R(\Delta x),$$

die Näherung 1. Ordnung ist also (Weglassen des Restgliedes!):

$$e^{\Delta x} \approx 1 + \Delta x.$$

Das ergibt:

Δx	$e^{\Delta x}$	absolute Differenz zum Näherungswert $1 + \Delta x$
0.1	$e^{0.1} : 1.1052$	0.0052
0.01	$e^{0.01}$	0.000050167
-0.01	$e^{-0.01}$	0.000049834

Vergleichen Sie mit den Ergebnissen der Näherung 0. Ordnung: Diese ist viel besser, man hat etwa die doppelte Stellenzahl hinter dem Komma korrekt.

2.2. Rezept zum Auffinden der ersten Ableitung mittels Tangentenerlegung. Aus den Ausführungen der zweiten Idee zur Ableitung ergibt sich eine sehr praktische Methode, Rechenausdrücke für die erste Ableitung von Funktionen *herzuleiten, auch ohne sie zuvor zu kennen*. Das beruht darauf:

SATZ 18. *Hat man eine Zerlegung*

$$f(x_0 + \Delta x) = f(x_0) + b\Delta x + R(\Delta x)$$

und erfüllt R die Restbedingung

$$\lim_{\Delta x \rightarrow 0, \Delta x \neq 0} \frac{R(\Delta x)}{\Delta x} = 0,$$

so gilt stets:

$$b = f'(x_0).$$

Das ist meist viel leichter als die Berechnung des Grenzwertes der Differenzenquotienten, Beispiele:

$$(x + \Delta x)^4 = x^4 + 4x^3\Delta x + 6x^2\Delta x^2 + 4x\Delta x^3 + \Delta x^4,$$

und der Restterm

$$6x^2\Delta x^2 + 4x\Delta x^3 + \Delta x^4$$

hat offenbar die Eigenschaft, noch durch Δx geteilt gegen Null zu gehen, also kann man die Ableitung

$$f'(x) = 4x^3, \text{ für } f(x) = x^4,$$

unmittelbar ablesen als den Faktor in der Zerlegung vor Δx . Damit ist sowohl die Existenz der Ableitung (für beliebiges x) als auch ihre Gestalt gezeigt!

Als nächstes Beispiel zeigen wir, dass die Exponentialfunktion überall differenzierbar ist und überall $\exp'(x) = \exp(x)$ gilt. Wir setzen dabei lediglich $\exp'(0) = 1$ und die allgemeine Gleichung $e^{x+y} = e^x \cdot e^y$ voraus. Das ergibt, wenn wir für $e^{\Delta x}$ die Tangentenerlegung von \exp an der Stelle $x_0 = 0$ einsetzen (s.o.):

$$e^{x+\Delta x} = e^x \cdot e^{\Delta x} = e^x(1 + \Delta x + R(\Delta x)) = e^x + e^x\Delta x + R(\Delta x).$$

Nun ist der Restterm laut Definition der von der Tangentenerlegung um Null, also in Ordnung, und wir lesen einfach den Faktor bei Δx ab:

$$\exp'(x) = \exp(x), \text{ für alle } x \in \mathbb{R}.$$

2.3. Der einfache Aufbau aller Ableitungen mittels des rekursiven Aufbaus aller Funktionen.

Tatsächlich braucht man auch 2.2 nur für kompliziertere theoretische Überlegungen, wenn man es einmal auf die Grundfunktionen angewandt und zur Herleitung der Regeln für das Zusammensetzen der Ableitungen der Bestandteile zu den Ableitungen von Zusammensetzungen strapaziert hat. Aus letzterem ergeben sich dann einfache Ableitungsregeln, mit denen man komplizierte Ableitungen zusammenbasteln kann.

2.3.1. *Die Ableitungen der Grundfunktionen.* Zunächst führen wir eine Notation ein, die es erlaubt, von Ableitungen von Rechenausdrücken zu reden, ohne immer ein Symbol wie f für die jeweilige Funktion einzuführen:

$$\frac{d}{dx}f(x) := f'(x).$$

Das erinnert an den „Differentialquotienten“ $\frac{dy}{dx}$, nur, dass wir statt „ y “ stets „ $f(x)$ “ schreiben und außerdem $\frac{d}{dx}$ (lies: „d nach dx“) als Operator vor den Ausdruck $f(x)$ setzen.

Hier ist die praktische Tabelle mit den Ableitungen unserer Grundfunktionen:

$$(2.4) \quad \begin{aligned} \frac{d}{dx}x^a &= ax^{a-1}, \quad a \in \mathbb{R} \\ \frac{d}{dx}e^x &= e^x \\ \frac{d}{dx}\ln(x) &= \frac{1}{x} \end{aligned}$$

Die zweite Formel haben wir in 2.2 schon eingesehen. Die dritte werden wir mittels einer Regel (Ableitung von Umkehrfunktionen, 2.4) aus der zweiten herleiten, und ebenso werden wir (Kettenregel, 2.4) die erste Formel in voller Allgemeinheit aus der zweiten herleiten. Wir sollten nur noch herausstellen, wie allgemein die erste Grundformel ist:

Insbesondere hat man

$$\begin{aligned} \frac{d}{dx}x^0 &= 0 \\ \frac{d}{dx}x^n &= nx^{n-1} \quad (n \in \mathbb{N}, n \neq 0) \\ \frac{d}{dx}\sqrt{x} &= \frac{d}{dx}x^{\frac{1}{2}} = \frac{1}{2}x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}} \\ \frac{d}{dx}\sqrt[5]{x^4} &= \frac{d}{dx}x^{\frac{4}{5}} = \frac{4}{5}x^{-\frac{1}{5}} \\ \frac{d}{dx}x^{-\frac{3}{8}} &= -\frac{3}{8}x^{-\frac{11}{8}}. \end{aligned}$$

Natürlich sollte man es als geometrisch selbstverständlich ansehen, dass eine konstante Funktion Ableitung Konstante Null hat. (Genau genommen wäre der Rechenausdruck $0 \cdot x^{-1}$ an der Stelle Null nicht definiert, aber die Ableitung der Funktion $f(x) = x^0 = 1$ existiert natürlich auch bei $x = 0$, mit dem Wert 0.)

2.3.2. *Die Ableitungsregeln für Zusammensetzungen.* Für diese haben wir etwas mehr zu tun. Sie lauten stets genauer: Wenn die Ableitungen der Bestandteile (an den jeweils betreffenden Stellen, die aus der Formel ersichtlich sind) existieren, so auch die Ableitung der Zusammensetzungen, und letztere hat dann den angegebenen Wert.

Linearität des Ableitungsoperators:

$$(2.5) \quad \begin{aligned} \frac{d}{dx}(cf(x)) &= c \frac{d}{dx}f(x), \text{ kürzer } (cf)' = cf' \\ \frac{d}{dx}(f(x) + g(x)) &= \frac{d}{dx}f(x) + \frac{d}{dx}g(x), \text{ kürzer } (f + g)' = f' + g' \end{aligned}$$

Produkt- und Quotientenregel:

$$(2.6) \quad \begin{aligned} (f \cdot g)' &= f'g + g'f \\ \left(\frac{f}{g}\right)' &= \frac{f'g - g'f}{g^2}. \end{aligned}$$

Kettenregel (für Hintereinanderschaltungen):

$$(2.7) \quad \frac{d}{dx}g(f(x)) = g'(f(x)) \cdot f'(x), \text{ kürzer: } (g \circ f)' = (g' \circ f) \cdot f'.$$

Regel für Umkehrfunktionen:

$$(2.8) \quad (f^{-1})'(f(x)) = \frac{1}{f'(x)}, \text{ wenn } f'(x) \neq 0.$$

Bevor wir die komplizierteren der benötigten Herleitungen geben (nur diese ohnehin), wollen wir zunächst einige Anwendungsbeispiele anschauen, insbesondere zuerst auch den Rest zu den Grundfunktionen erledigen:

Beispiele:

- Die Grundregel

$$\frac{d}{dx}x^a = ax^{a-1}$$

sieht man so: Zunächst hat man für $x > 0$ (so dass man $\ln(x)$ bilden kann):

$$\frac{d}{dx}x^a = \frac{d}{dx}(e^{\ln(x)})^a = \frac{d}{dx}e^{a \cdot \ln(x)} = \frac{a}{x}e^{a \cdot \ln(x)} = \frac{a}{x}x^a = ax^{a-1}.$$

Nur Umformungen und einmal die Kettenregel waren beteiligt.

Tatsächlich bilden $x = 0$ und $x < 0$ Sonderfälle: Für allgemeine Exponenten a ist x^a überhaupt nicht definiert für $x \leq 0$, und dafür macht dann auch die Ableitung keinen Sinn. Für $x = 0$ ist zwar x^a mit $1 > a > 0$ definiert, aber das hat bei $x = 0$ unendliche Steigung und somit keine Ableitung. Genau dies gibt der hergeleitete Ausdruck auch wieder. Mit $a > 1$ hat man $0^a = 0$, die Ableitung existiert als einseitige mit dem richtigen Wert Null, den der Ausdruck wiederum angibt. Natürlich ist x^a zuweilen auch für negative Werte von x definiert, aber dann ergibt sich wiederum der richtige Wert, weil es sich um eine (positive oder negative) ungerade ganze Zahl a handeln muss. Kein Problem: Die Funktion $x \mapsto x^a$ ist dann ungerade, die Ableitung also gerade (kann man leicht herleiten!), und somit stimmt die Formel wieder, da $x \mapsto ax^{a-1}$ mit ungeradem a tatsächlich eine gerade Funktion ist, also ist die Ableitung überall korrekt, wenn sie es auf der positiven Seite war.

- Die Ableitung der Logarithmusfunktion

Die Umkehrfunktions-Regel gibt her:

$$(\exp^{-1})'(\exp(x)) = \ln'(e^x) = \frac{1}{\exp(x)} = \frac{1}{e^x}.$$

Dabei ist x beliebig, somit e^x eine beliebige Zahl > 0 . Folglich gilt:

$$\ln'(x) = \frac{1}{x} \text{ für alle } x > 0.$$

(Verwenden Sie zunächst einen Buchstaben a für e^x aus der darüber stehenden Formel, das gilt dann für alle $a > 0$, und nun denken Sie daran, dass der Buchstabe in Ausdrücken „für alle...“, „es gibt...“ keine Rolle spielt.)

- Einfache Beispiele zur Anwendung der Kettenregel:

$$\frac{d}{dx}(x^2 + 1)^{70} = 140x(x^2 + 1)^{69}.$$

$$\frac{d}{dx} \ln(x^2 + 1) = \frac{2x}{x^2 + 1}.$$

$$\frac{d}{dx} a^x = \frac{d}{dx} e^{x \cdot \ln(a)} = \ln(a) \cdot e^{x \cdot \ln(a)} = \ln(a) \cdot a^x, \quad a > 0.$$

$$\frac{d}{dx} \log_a(x) = \frac{d}{dx} \frac{\ln(x)}{\ln(a)} = \frac{1}{\ln(a)} \ln'(x) = \frac{1}{x \cdot \ln(a)}, \quad a > 0.$$

(Man beachte, dass man für den konstanten (!) Nenner $\ln(a)$ nicht die Quotientenregel benötigt (obgleich sie selbstverständlich auch das richtige Resultat brächte, nur viel zu umständlich), sondern einfach das Stehenbleiben eines konstanten Faktors - hier $\frac{1}{\ln(a)}$ - beim Ableiten (gemäß Linearität) benutzen kann.)

- Ein Beispiel zur iterierten Verwendung der Kettenregel für Mehrfachschachtelungen:

$$\begin{aligned} \frac{d}{dx} e^{\sqrt{x^2+1}} &= \exp'(\sqrt{x^2+1}) \cdot \frac{d}{dx} \sqrt{x^2+1} = \exp(\sqrt{x^2+1}) \cdot \frac{2x}{2\sqrt{x^2+1}} \\ &= \frac{x}{\sqrt{x^2+1}} e^{\sqrt{x^2+1}}. \end{aligned}$$

Man sieht also, wie das logisch funktioniert: Man wendet einmal die Kettenregel an, dabei sieht man, dass die innere Funktion wiederum eine Schachtelung ist, und hat die Kettenregel erneut zur Ableitung dieser inneren Funktion anzuwenden. Mit etwas Übung würde man den Ausdruck $\frac{d}{dx} \sqrt{x^2+1}$ nicht mehr erst hinschreiben, sondern gleich dessen Ergebnis.

Die Herleitung zweier ausgewählter Regeln: Kettenregel und Quotientenregel

Zur Herleitung der Kettenregel braucht man lediglich die Voraussetzungen, dass f an der Stelle x_0 , g an der Stelle $f(x_0)$ differenzierbar seien, in die Existenz von Tangenzenzerlegungen umzusetzen und anschließend das Rezept von 2.2 zu befolgen, d.h. den Ausdruck $(g \circ f)(x_0 + \Delta x)$ für die zu differenzierende Funktion zu schlachten. Man beachte, dass wir die komische Kettenregel-Formel auf diese Weise *finden* werden (nicht eine uns von Autoritäten eingeflöbte Aussage bestätigen).

Die Voraussetzungen lauten:

- (i) $f(x_0 + \Delta x) = f(x_0) + f'(x_0)\Delta x + R(\Delta x)$, mit erfüllter Restbed. für R .
- (ii) $g(f(x_0) + \Delta y) = g(f(x_0)) + g'(f(x_0))\Delta y + S(\Delta y)$, mit erfüllter Restbed. für S .

Es ist kein prinzipieller Unterschied zwischen Δx und Δy , die Gleichungen gelten für jeweils alle diese Zahlen, wobei die Restbedingungen etwas Interessantes nur für kleine derartige Zahlen aussagen. Jetzt zerlegen wir:

$$\begin{aligned}
g(f(x_0 + \Delta x)) &= (g \circ f)(x_0 + \Delta x), \text{ also das Gewünschte]} \\
&= g(f(x_0) + \underbrace{f'(x_0)\Delta x + R(\Delta x)}_{\Delta y \text{ nennen wir das}}) \text{ (für inneren Ausdruck (i) benutzt)} \\
&= g(f(x_0) + \Delta y) \\
&= g(f(x_0)) + g'(f(x_0))\Delta y + S(\Delta y) \text{ (benutze (ii))} \\
&= g(f(x_0)) + g'(f(x_0))(f'(x_0)\Delta x + R(\Delta x)) + S(\Delta y) \text{ (\Delta y eingesetzt)} \\
&= g(f(x_0)) + \underbrace{g'(f(x_0))f'(x_0)}_{\text{Faktor bei } \Delta x, \text{ gesuchte Ableitung!}} \Delta x + \underbrace{g'(f(x_0))R(\Delta x) + S(\Delta y)}_{\text{Das sollte der Restterm sein.}}
\end{aligned}$$

Wir lesen die Ableitung als Faktor bei Δx ab, nachdem wir uns überzeugt haben, dass der Restterm in Ordnung ist. Er lautet

$$T(\Delta x) = g'(f(x_0))R(\Delta x) + S(\Delta y)$$

Zu zeigen ist, dass

$$\frac{T(\Delta x)}{\Delta x} \rightarrow 0 \text{ für } \Delta x \rightarrow 0.$$

Nun ist $T(\Delta x)$ eine Summe, und es genügt, wenn wir von beiden Summanden diese Eigenschaft zeigen, erstens:

$$\frac{g'(f(x_0))R(\Delta x)}{\Delta x} \rightarrow 0 \text{ für } \Delta x \rightarrow 0.$$

Nach Voraussetzung (i) gilt

$$\frac{R(\Delta x)}{\Delta x} \rightarrow 0 \text{ für } \Delta x \rightarrow 0.$$

Dann gilt aber dasselbe auch, wenn mit der konstanten (endlichen) Zahl $g'(f(x_0))$ multipliziert wird, dass das noch gegen Null geht.

Zweitens ist zu zeigen, dass:

$$\frac{S(\Delta y)}{\Delta x} \rightarrow 0 \text{ für } \Delta x \rightarrow 0.$$

Das ist nun nicht direkt in den Voraussetzungen enthalten, sondern wir haben gemäß (ii) nur erst:

$$\frac{S(\Delta y)}{\Delta y} \rightarrow 0 \text{ für } \Delta y \rightarrow 0.$$

Wir haben jedoch einen konkreten Ausdruck von Δy als Funktion von Δx . Setzen wir voraus, dass für alle hinreichend kleinen $|\Delta x|$ mit $\Delta x \neq 0$ gilt, dass $\Delta y \neq 0$ ist, so haben wir für alle interessierenden (nämlich hinreichend kleinen) Δx :

$$\begin{aligned}
\frac{S(\Delta y)}{\Delta x} &= \frac{S(\Delta y)}{\Delta y} \cdot \frac{\Delta y}{\Delta x} \text{ (Erweitern mit } \Delta y \neq 0) \\
&= \frac{S(\Delta y)}{\Delta y} \cdot \frac{f'(x_0)\Delta x + R(\Delta x)}{\Delta x} \text{ (Für } \Delta y \text{ Ausdruck eingesetzt im Zähler).}
\end{aligned}$$

Von diesem Ausdruck ist nun einzusehen, dass er gegen Null geht, wenn nur Δx gegen Null geht. Nun, der erste Faktor geht nach Null, weil mit Δx auch Δy nach Null geht (definition von Δy), und weil mit Δy auch $\frac{S(\Delta y)}{\Delta y}$ nach Null geht, gemäß Voraussetzung (ii). Was ist mit dem zweiten Faktor? Er geht gegen den endlichen festen Wert $f'(x_0)$; denn

$$\frac{f'(x_0)\Delta x + R(\Delta x)}{\Delta x} = \frac{f'(x_0)\Delta x}{\Delta x} + \frac{R(\Delta x)}{\Delta x} = f'(x_0) + \frac{R(\Delta x)}{\Delta x},$$

und $\frac{R(\Delta x)}{\Delta x} \rightarrow 0$ für $\Delta x \rightarrow 0$ nach Voraussetzung (i).

Damit ist der ganze Restterm in Ordnung und die Kettenregel bewiesen, lediglich verbleibt zu überlegen, dass die Voraussetzung, dass $\Delta y \neq 0$ für hinreichend kleine $|\Delta x|$, $\Delta x \neq 0$, gilt, keinen Verlust bedeutet. Kurzum: Getrennt zu behandeln ist noch:

Der Fall $\Delta y = 0$ für beliebig kleine $|\Delta x|$.

Daraus folgt (zusammen mit unserer Voraussetzung, dass $f'(x_0)$ existiert): $f'(x_0) = 0$. Wir haben also zu zeigen, dass $(g \circ f)'(x_0)$ existiert und den Wert $0 = g'(f(x_0)) \cdot f'(x_0)$ hat. Wir nehmen eine beliebige Δx -Folge, die nach Null geht, sagen wir $(\Delta x_i)_{i \in \mathbb{N}}$, und haben einzusehen, dass dafür die Folge

$$\left(\frac{(g(f(x_0 + \Delta x_i)) - g(f(x_0)))}{\Delta x_i} \right)_{i \in \mathbb{N}} \quad (*)$$

ebenfalls gegen Null geht. Das ist aber leicht zu sehen: Betrachten wir die zur Folge $(\Delta x_i)_i$ gehörige Folge $(\Delta y_i)_i$, $\Delta y_i = f(x_0 + \Delta x_i) - f(x_0)$. Nehmen wir davon nur diejenigen Folgenglieder, die verschieden von Null sind, so können wir wie im ersten Fall schließen und erhalten, dass die zugehörige Teilfolge von (*) gegen Null geht. Durch die Δy_i , die Null sind, werden in (*) noch zusätzlich Nullen „eingeschossen“, und es bleibt dabei, dass (*) gegen Null geht. Also stimmt auch im betrachteten Fall die Kettenregel-Formel: Mit $f'(x_0) = 0$ ergibt auch $g'(f(x_0)) \cdot f'(x_0)$ den richtigen Ableitungswert für $g \circ f$ an der Stelle x_0 , nämlich Null.

Herleitung der Quotientenregel: Zunächst genügt es, die Regel speziell nur für Funktionen der Form

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

zu zeigen. (Den Rest besorge man mit der Produktregel, die man sich leicht bei Bedarf herleiten kann.) Außerdem können wir (ohne Verlust) noch weiter spezialisieren auf den Fall, dass $g(x) > 0$ im relevanten Intervall. - Sonst gehe man über zu $-g$ und benutze die Linearität der Ableitung. Wir bilden nun

$$h(x) = \ln(f(x)) = -\ln(g(x))$$

und leiten dies mittels der schon bereitliegenden Kettenregel ab:

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

Das ergibt

$$f'(x) = -\frac{g'(x) \cdot f(x)}{g(x)} = -\frac{g'(x)}{g^2(x)}.$$

Damit ist gezeigt, dass $\frac{d}{dx} \frac{1}{g(x)} = -\frac{g'(x)}{g^2(x)}$, und mit einer Anwendung der Produktregel kommt man nun auf die Formel $\frac{d}{dx} \frac{f(x)}{g(x)} = \frac{f'(x)g(x) - g'(x)f(x)}{g^2(x)}$.

2.4. Anwendung der Ableitung auf die quantitative Bestimmung lokaler Extrema. Eine Funktion wie $f(x) = x^3$ (mit Definitionsbereich \mathbb{R} verstanden) hat keine *globalen* Extrema, d.h. keinen minimalen und keinen maximalen Wert *bezogen auf den gesamten Definitionsbereich*. Denn es kommen beliebig große und beliebig kleine Werte heraus. Auch $g(x) = x^3 - x$ hat keine *globalen* Extrema, wohl aber *lokale*, d.h. *auf beliebig kleine Umgebungen bezogen*. (f dagegen besitzt auch keine lokalen Extrema.) Das wissen wir von der groben Skizze des Graphen von g . Nunmehr stellen wir das Problem, die betreffenden Stellen genau zu lokalisieren. Es genügt für alle Zwecke, die zugehörigen Abszissenwerte (x -Werte) bestimmen zu können, dann lassen sich die zugehörigen (minimalen oder maximalen) Funktionswerte, an denen man interessiert ist, ohne weiteres finden. Das lokale Verständnis erweist sich als recht fruchtbar, da es mittels der Ableitung bewältigt werden kann und zu für sich interessanten lokalen Extrema führt oder auch gegebenenfalls zu globalen. Daher definieren wir noch einmal ausführlicher:

DEFINITION 34. f hat an der Stelle x_0 ein lokales Minimum [bzw. lokales Maximum], wenn es ein (beliebig kleines!) Intervall $]x_0 - \varepsilon, x_0 + \varepsilon[$, $\varepsilon > 0$, um x_0 herum gibt, so dass für alle $x \in]x_0 - \varepsilon, x_0 + \varepsilon[$, d.h. $|x - x_0| < \varepsilon$, gilt:

$$f(x) - f(x_0) \geq 0 \text{ [bzw. } f(x) - f(x_0) \leq 0].$$

Es ist leicht auszurechnen, aber anschaulich noch einfacher zu sehen, dass im Falle der Differenzierbarkeit von f an der Stelle x_0 die Tangente an den Graphen von f im Punkte $(x_0, f(x_0))$ die Steigung Null besitzen muss, wenn f an dieser Stelle ein lokales Extremum hat. Also:

SATZ 19. Wenn $f'(x_0)$ existiert und f an der Stelle x_0 ein lokales Extremum hat, dann gilt $f'(x) = 0$. (Existenz von $f'(x_0)$ setzt insbesondere voraus, dass f in einer (beidseitigen!) Umgebung von x_0 definiert ist!)

Zur Anwendung dieser Aussage achte man unbedingt auf die Richtung des „wenn-so“-Pfeiles: Aus $f(x_0) \neq 0$ kann man (bei Existenz der Ableitung) schließen, dass f an der Stelle x_0 *keinen* Extremwert haben kann. (Nicht aber folgt etwa aus $f'(x_0) = 0$, dass dort ein Extremwert vorliegt. Man merke sich dazu das Gegenbeispiel $f(x) = x^3$: An der Stelle $x_0 = 0$ verschwindet die Ableitung, aber es handelt sich um keinen Extremwert, sondern um einen Sattel. Für den Fall, dass f nicht differenzierbar ist oder eine ansonsten differenzierbare Funktion auf ein Intervall eingeschränkt wird, macht der Satz keinerlei Aussage! Beispiele: $f : [0, 1] \rightarrow [0, 1]$, $f(x) = x^2$, hat ein absolutes Maximum in $x = 1$, und das findet man nicht über „Ableitung Nullsetzen“. Die Aufgabe, die Punkte eines Quadrats zu finden, die maximalen Abstand von einem Punkt im Innern des Quadrats haben, lässt sich nicht über Ableitung lösen, da die zugehörige Funktion gerade an den Stellen nicht differenzierbar ist, wo diese Maxima liegen.

$f'(x_0) = 0$ ist nur eine *notwendige*, aber keine *hinreichende* Bedingung für die Existenz eines lokalen Extremwertes - und das auch nur unter Voraussetzung der Existenz einer (beidseitigen) Ableitung. Das bedeutet: Die Nullstellen der ersten Ableitung bilden nur erst *Kandidaten für Extremstellen*. Man mache sich jedoch den Wert dieser Aussage klar: Nachdem zunächst einmal *jede* reelle Zahl in Frage kam, zumindest jedoch ein Kontinuum, so hat man in aller Regel nunmehr endlich viele Kandidaten (allenfalls abzählbar viele). Wie kommt man zur endgültigen Entscheidung? Hier sind zwei Möglichkeiten, zu denen man eher greifen sollte als zur zweiten Ableitung:

- Man hat bereits eine grobe Skizze und weiß, dass es mindestens k lokale Extremstellen geben muss, und findet nun genau k Nullstellen der ersten Ableitung: Dann müssen diese Nullstellen die Abszissenwerte dieser Extremstellen sein!
- Man betrachtet die Ableitung - und zwar nur ihr Vorzeichen - in einer Umgebung von einer fraglichen Stelle x_0 mit der Eigenschaft $f'(x_0) = 0$: Wenn f' an dieser Stelle einen *Vorzeichenwechsel* erleidet, dann handelt es sich um ein *Extremum*, sonst nicht! Man kann sogar sagen, ob es sich im ersteren Falle um ein Minimum oder Maximum handelt: Bei Wechsel des Vorzeichens von negativ auf positiv (von links nach rechts gesehen) handelt es sich um ein Minimum, wie man sich veranschauliche. Bei umgekehrter

Reihenfolge ist es ein Maximum. (Man beachte, dass bei einer in einem Intervall konstanten Funktion Minimum und Maximum dasselbe sind und die Ableitung konstant Null ist, diesen Fall wollten wir hier als trivialen nicht besonders betrachten. Daher interessieren hier nur echt positive oder negative Ableitungen abseits von x_0 .)

Beispiel:

$f(x) = x^3 - x$; die Ableitung wird Null für $3x^2 - 1 = 0$, also $x_1 = -1/\sqrt{3}$, $x_2 = 1/\sqrt{3}$. Bereits aus der groben Skizze erkennt man sofort, dass an der ersten Stelle ein lokales Maximum liegt, an der zweiten ein lokales Minimum. Alternativ könnten wir auch rein rechnerisch (über das Vorzeichen der Ableitung) bequem zu diesem Resultat gelangen: An der Stelle $x_1 = -1/\sqrt{3}$ beobachtet man den Wechsel $+|-$. Offensichtlich ist nämlich $f'(x) > 0$ für $x < x_1$, und offensichtlich gilt $f'(x) < 0$ für Werte x , die *ein wenig* größer sind als x_1 (nämlich noch kleiner als x_2). Klar liegt bei x_2 der Wechsel $-|+$ vor.

2.5. Ein fundamentales Resultat der Analysis: Mittelwertsatz bzw. Satz vom endlichen Zuwachs. Die angesprochenen Sätze handeln davon, wichtigste globale Eigenschaften einer Funktion (also solche, die sich auf ein ganzes Intervall beziehen), durch das lokale Ableitungsverhalten auszudrücken. Sie haben eine Fülle von interessanten Anwendungen.

SATZ 20 (Mittelwertsatz). *Sei f auf $[a, b]$ stetig, auf $]a, b[$ differenzierbar, $a < b$. Dann gibt es eine Zahl $\xi \in]a, b[$ mit der Eigenschaft*

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

Der Satz besagt also, dass die *mittlere* Steigung von f auf $[a, b]$, welche durch den Differenzenquotienten auf der rechten Seite ausgedrückt wird, an *mindestens einer* Stelle ξ im Innern des Intervalls als *lokale* Steigung (Ableitung an dieser Stelle) angenommen wird. Wir beweisen den Satz nicht, führen ihn aber auf ein noch grundlegenderes Resultat über stetige Funktionen zurück: Zunächst betrachtet man den Fall $f(a) = f(b)$. Wenn f konstant ist, so hat f überall im Innern den Ableitungswert Null. Damit ist die Aussage des Satzes in diesem Fall klar. Ist f nicht konstant, so nehme f etwa Werte $> f(a)$ im Innern an. Da f stetig ist, nimmt f auf $[a, b]$ an irgendeiner Stelle ein absolutes Maximum an (fundamentales Resultat zu stetigen Funktionen, das wir voraussetzen - dafür muss man einiges tun). (Im andern Fall, dass f Werte $< f(a)$ im Innern annimmt, erhält man derart ein absolutes Minimum.) Nennen wir diese Stelle ξ . Nach dem vorigen Satz gilt mit der Voraussetzung, dass f auf $]a, b[$ differenzierbar sei: $f'(\xi) = 0$, und das ist gerade der Wert des Differenzenquotienten mit der Zusatzvoraussetzung $f(a) = f(b) = 0$. (Dieser Spezialfall des Mittelwertsatzes heißt Satz von Rolle.) Nun verallgemeinert man für eine beliebige Funktion f , welche die Voraussetzungen des Satzes erfüllt, indem man definiert:

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

Dann erfüllt g die Voraussetzungen des Satzes von Rolle, insbesondere $g(a) = g(b) = 0$. (Stetigkeits- und Differenzierbarkeitseigenschaften von g sind dieselben wie von f !) Also mit der Folgerung des Satzes von Rolle: Es gibt $\xi \in]a, b[$, so dass $g'(\xi) = 0$. Aber mit den Ableitungsregeln:

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

was mit $g'(\xi) = 0$ sofort die Aussage des Mittelwertsatzes ergibt.

FOLGERUNG 3 (Satz vom endlichen Zuwachs). *Sei f auf $[a, b]$ stetig und auf $]a, b[$ differenzierbar ist, $a < b$. Dann gilt:*

$$\begin{aligned} \text{Wenn } f'(x) &\leq g'(x) \text{ für alle } x \in]a, b[, \text{ dann} \\ f(x) - f(a) &\leq g(x) - g(a) \text{ für alle } x \in [a, b]. \end{aligned}$$

Die Herleitung aus dem Mittelwertsatz ist leicht: Wenn es eine Zahl x_0 , $a < x_0 \leq b$, gäbe, welche dem Satz widerspräche, so hätten wir mit der Funktion $h(x) = f(x) - f(a) - (g(x) - g(a))$ die Voraussetzungen des Mittelwertsatzes erfüllt, also wegen $h(x_0) > 0$ und $h(a) = 0$ eine Zahl $\xi \in]a, x_0[$, so dass $h'(\xi) > 0$. Aber $h'(\xi) = f'(\xi) - g'(\xi)$, so dass $f'(\xi) > g'(\xi)$ im Widerspruch zur Voraussetzung des Satzes vom endlichen Zuwachs, da $a < \xi < b$.

Anwendungen der Sätze:

SATZ 21.

Wenn $f'(x) > 0$ für alle $x \in [a, b]$, dann ist f auf $[a, b]$ streng monoton steigend.

Wenn $f'(x) < 0$ für alle $x \in [a, b]$, dann ist f auf $[a, b]$ streng monoton fallend.

Wenn $f'(x) = 0$ für alle $x \in [a, b]$, dann ist f auf $[a, b]$ konstant. Insbesondere folgt aus $f' = g'$ auf $[a, b]$, dass $f = g + c$, mit einer Konstanten c , auf $[a, b]$ gilt.

Dies folgt unmittelbar aus dem Satz vom endlichen Zuwachs. Man beachte: Aus dem streng monotonen Steigen einer in einem Intervall differenzierbaren Funktion folgt nicht, dass $f' > 0$ überall im Intervall gelten muss: $f(x) = x^3$ ist ein Beispiel dafür; es ist $f'(0) = 0$, aber f ist global streng monoton wachsend.

Der Satz vom endlichen Zuwachs eignet sich auch dafür, konkrete wichtige Ungleichungen zu beweisen, dazu ein Beispiel:

$$\text{Behauptung: } \sin(x) \leq x - \frac{1}{6}x^3 + \frac{1}{120}x^5 \text{ für } x \geq 0.$$

Beweis: Da die Ungleichung für $x = 0$ gilt, genügt nach dem Satz vom endlichen Zuwachs die Ungleichung für die Ableitung auf beiden Seiten:

$$\cos(x) \leq 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 \text{ für } x \geq 0.$$

Diese gilt wieder für $x = 0$, also genügt mit demselben Argument die Ungleichung

$$-\sin(x) \leq -x + \frac{1}{6}x^3 \text{ für } x \geq 0,$$

dafür reicht nun offenbar

$$-\cos(x) \leq -1 + \frac{1}{2}x^2 \text{ für } x \geq 0,$$

und dafür genügt

$$\sin(x) \leq x \text{ für } x \geq 0.$$

Diese wiederum folgt direkt aus $\cos(x) \leq 1$, was klar ist. Also gelten (rückwärts bis oben hinauf) alle diese Ungleichungen. Verallgemeinert man auf höhere Potenzen, so erreicht man auf diese Weise sogar die Reihendarstellung für Sinus!

2.6. Die de L'Hospitalischen Regeln. Eine nützliche Anwendung der ersten Ableitung besteht darin, die Frage der Existenz von $\lim_{x \rightarrow a} \frac{f(x)}{g(x)}$ gerade in den kritischen Fällen $\lim_{x \rightarrow a} f(x) = 0$, $\lim_{x \rightarrow a} g(x) = 0$ (oder beide ∞) beantworten zu können. Dabei darf auch $a = \infty$ sein. Man hat den

SATZ 22 (de L'Hospitalische Regeln). *Es sei $a \in \mathbb{R} \cup \{\infty\}$. Weiter gelte $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$ oder aber $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = \infty$. Dann gilt:*

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

$$\text{und es gilt : } \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}.$$

Man mache sich klar: Der Satz hat Voraussetzungen. Zum Beispiel ist $\lim_{x \rightarrow 0, x > 0} \frac{1}{x} = \infty$, aber $\lim_{x \rightarrow 0, x > 0} \frac{0}{1} = 0$. In diesem Falle ist die Voraussetzung nicht erfüllt, dass Zähler und Nenner *beide* gegen Null gehen müssen. Aber in solchen Fällen ist die Antwort direkt klar. Dagegen können wir die Regeln heranziehen, um etwa zu zeigen: $\lim_{x \rightarrow \infty} x/e^x = 0$; denn $\lim_{x \rightarrow \infty} 1/e^x = 0$. Ebenso: $\lim_{x \rightarrow \infty} \ln(x)/x = 0$, da $\lim_{x \rightarrow \infty} (1/x)/1 = 0$. In manchen Fällen braucht man mehrfache Anwendung, z.B. für $x^2/e^x \rightarrow 0$ für $x \rightarrow \infty$. Entsprechend weiter muss man induktiv für x^n/e^x gehen. Man beachte, dass in all diesen Fällen die Voraussetzungen des Satzes erfüllt sind.

2.7. Zur Bedeutung der zweiten Ableitung. Man kann eine Funktion zwei man differenzieren, wenn denn die erste Ableitung existiert und diese wiederum differenzierbar ist. Damit $f''(x_0)$ existiert, muss selbstverständlich f' in einer ganzen Umgebung von x_0 existieren und eine Tangenzenzerlegung in x_0 besitzen. Rechnerisch ist das natürlich mit den üblich auftretenden Ausdrücken ganz leicht, z.B. ist für $f(x) = x^3 + \sin(2x)$: $f''(x) = 6x - 4\cos(2x)$. Auch diese zweite Ableitung hat dann klare geometrische Deutungen: Man kann sie zum Beispiel für Wendepunkte heranziehen. f hat definitionsgemäß in x_0 einen Wendepunkt, wenn f' in x_0 einen lokalen Extremwert annimmt. Kandidaten sind also die Stellen, an denen die zweite Ableitung verschwindet. Wie oben besprochen hat man dann noch zu prüfen, ob ein Extremwert der ersten Ableitung wirklich vorliegt. Indessen kann man auch die zweite Ableitung manchmal (nicht immer ist das rechnerisch günstig, schon gar nicht immer hat es Erfolg) heranziehen, um zu prüfen, ob eine Funktion einen Extremwert hat: Wenn $f'(x_0) = 0$ und $f''(x_0) > 0$, so hat f in x_0 ein lokales Minimum. Wenn $f'(x_0) = 0$ und $f''(x_0) < 0$, so hat f in x_0 ein lokales Maximum. Das sieht man schnell ein: Sei also $f'(x_0) = 0$ und z.B. $f''(x_0) > 0$. Aus $f''(x_0) > 0$ folgt, dass für hinreichend kleine Beträge von Δx gilt:

$$\frac{f'(x_0 + \Delta x) - f'(x_0)}{\Delta x} > 0.$$

Das bedeutet mit $f'(x_0) = 0$:

$$\begin{aligned} f'(x_0 + \Delta x) &> 0 \text{ für } \Delta x > 0, \\ f'(x_0 + \Delta x) &< 0 \text{ für } \Delta x < 0. \end{aligned}$$

Es folgt also, dass die erste Ableitung bei x_0 einen Vorzeichenwechsel $-/+$ macht, also die Funktion f selbst vom Fallen ins Steigen übergeht, d.h. in x_0 ein lokales Minimum vorliegen muss. Klar geht der Schluss für den Fall $f''(x_0) < 0$ analog.

Allgemeiner gibt f'' Auskunft über das Krümmungsverhalten von f : Ist $f''(x_0) > 0$ [bzw. < 0], so befindet man sich im Punkte $(x_0, f(x_0))$ in einer Linkskurve [bzw. Rechtskurve] des Graphen von f , so verstanden: Sie durchfahren in Richtung anwachsender x -Werte den Graphen von f , dann ist klar, was eine Linkskurve und eine Rechtskurve ist. Entsprechend liegt ein lokales Minimum in einer Linkskurve, ein lokales Maximum in einer Rechtskurve (damit wird der obenstehende Satz über Extremwerte von f anschaulich klar), und ein Wendepunkt bedeutet den Übergang einer Linkskurve in eine Rechtskurve bzw. umgekehrt. Dort muss also die zweite Ableitung verschwinden.

Zur geometrischen Bedeutung kommt eine weitere für die lokale Näherung einer Funktion durch eine einfacheren, diesmal nicht mehr lineare, sondern quadratische: Man geht zur Näherung von f um x_0 zweiter Ordnung über, indem man bildet: $f(x_0 + \Delta x) \approx f(x_0) + f'(x_0)\Delta x + \frac{1}{2}f''(x_0)\Delta x^2$, Voraussetzung: Existenz von $f''(x_0)$. Dann hat man einen Fehler, der noch durch Δx^2 geteilt gegen Null geht für $\Delta x \rightarrow 0$.

2.8. Ableitung von Kurven. Erinnerung: Eine Kurve im \mathbb{R}^3 ist eine Abbildung $\vec{x} : D \rightarrow \mathbb{R}^3$, mit einem Definitionsbereich $D \subseteq \mathbb{R}$, D ist gewöhnlich ein reelles Intervall, beschränkt oder unbeschränkt. Das einfachste Beispiel war $\vec{x}(t) = \vec{x}_0 + t\vec{a}$, $t \in \mathbb{R}$. Physikalisch wird man die unabhängige Variable gern t schreiben und als Zeit deuten. Dann beschreibt eine Kurve nicht nur eine Spur, eine Bahn, das eindimensionale Gebilde von Punkten - die Bahn ist einfach nur $Bild(\vec{x})$, sondern auch, wie eine solche Bahn in der Zeit durchlaufen wird - im Beispiel: mit konstanter Geschwindigkeit. Ein interessanteres Beispiel ist die Flugparabel: $\vec{x}(t) = \vec{x}_0 + t\vec{v}_0 + \frac{1}{2}t^2\vec{g}$, $t \in \mathbb{R}$. Das ist eine Orts-Zeit-Funktion, und man sollte erwarten, dass die Momentangeschwindigkeit zur Zeit t einfach

die Ableitung nach der Zeit an der Stelle t ist. Aber wie wäre die Ableitung einer vektorwertigen Abbildung wie \vec{x} zu bilden? Im Beispiel der Flugparabel wird man - richtig - vermuten, dass man einfach die konstanten Vektorfaktoren stehenzulassen und die von t abhängigen Zahl-Vorfaktoren abzuleiten hat, also $\vec{x}'(t) = \vec{v}_0 + t\vec{g}$. \vec{v}_0 kommt als Geschwindigkeitsvektor zur Zeit $t = 0$ heraus. Die zweite Ableitung sollte den Beschleunigungsvektor ergeben, und das tut sie auch: $\vec{x}''(t) = \vec{g}$. Aber einmal wäre dies zu rechtfertigen, zum andern wäre die Frage zu stellen, wie man allgemeinere Kurven abzuleiten hätte, etwa $\vec{x}(t) = (\cos(t), \sin(t), t)$ (Schraubenbewegung). Hervorragende Dienste leistet das sinngemäße Bilden einer Tangentenerlegung:

$$\text{Sei } \vec{x}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}, \quad t \in]a, b[, \quad a < b.$$

Dann setzen wir $\vec{x}(t_0 + \Delta t)$ an mit einer Stelle $t_0 \in]a, b[$ und kommen auf entsprechende Ausdrücke in den Komponentenfunktionen von \vec{x} , die wir mit x, y, z bezeichnet haben. Setzen wir nun einmal voraus, dass diese Komponentenfunktionen an der Stelle t_0 differenzierbar seien, dann haben wir mit Einsetzen von deren Tangentenerlegungen an der Stelle t_0 :

$$\begin{aligned} \vec{x}(t_0 + \Delta t) &= \begin{pmatrix} x(t_0 + \Delta t) \\ y(t_0 + \Delta t) \\ z(t_0 + \Delta t) \end{pmatrix} = \begin{pmatrix} x(t_0) + x'(t_0)\Delta t + R(\Delta t) \\ y(t_0) + y'(t_0)\Delta t + S(\Delta t) \\ z(t_0) + z'(t_0)\Delta t + T(\Delta t) \end{pmatrix} \\ &= \begin{pmatrix} x(t_0) \\ y(t_0) \\ z(t_0) \end{pmatrix} + \begin{pmatrix} x'(t_0) \\ y'(t_0) \\ z'(t_0) \end{pmatrix} \Delta t + \begin{pmatrix} R(\Delta t) \\ S(\Delta t) \\ T(\Delta t) \end{pmatrix}. \end{aligned}$$

Bezeichnen wir nun den Vektor bei Δt als \vec{a} und den dritten, den „Restvektor“ mit $\vec{u}(\Delta t)$, so haben wir eine Zerlegung völlig analog zur Tangentenerlegung bei einfachen Funktionen geschafft:

$$(2.9) \quad \vec{x}(t_0 + \Delta t) = \vec{x}(t_0) + \vec{a}\Delta t + \vec{u}(\Delta t).$$

Wir haben nur noch die Restbedingung für die Näherung 1. Ordnung für unsern allgemeineren Fall sinnvoll zu formulieren, sie lautet: $\vec{u}(\Delta t)/\Delta t \rightarrow \vec{0}$ für $\Delta t \rightarrow 0$, $\Delta t \neq 0$. Vektor geteilt durch Zahl $\neq 0$ ist ja sinnvoll, und es kommt ein Vektor heraus! Weiter wird klar, dass diese Bedingung nach unserer Voraussetzung über die Reste R, S, T erfüllt ist, welche den Tangentenerlegungen der Komponentenfunktionen entstammten: Alle Komponenten von $\vec{u}(\Delta t)$ gehen nach 0 für $\Delta t \rightarrow 0$, also geht $\vec{u}(\Delta t)$ nach $\vec{0}$. Drehen wir nun die Sache um, setzen also voraus, dass wir obenstehende Zerlegung von $\vec{x}(t_0 + \Delta t)$ haben mit einem festen Vektor \vec{a} und einem Restvektor $\vec{u}(\Delta t)$, welcher die gerade formulierte Bedingung erfüllt. Dann behaupten wir, dass die Komponentenfunktionen x, y, z alle eine Tangentenerlegung an der Stelle t_0 besitzen und der Vektor \vec{a} gerade deren Ableitungen an der Stelle t_0 als Komponenten hat. Dies einzusehen, müssen wir lediglich die frühere Einsicht beachten, dass eine Vektorgleichung in Komponentenform entsprechende Zahlgleichungen ergibt, also mit Bezeichnung der Komponenten a_i von \vec{a} und u_i von \vec{u} :

$$\vec{x}(t_0 + \Delta t) = \begin{pmatrix} x(t_0 + \Delta t) \\ y(t_0 + \Delta t) \\ z(t_0 + \Delta t) \end{pmatrix} = \begin{pmatrix} x(t_0) \\ y(t_0) \\ z(t_0) \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \Delta t + \begin{pmatrix} u_1(\Delta t) \\ u_2(\Delta t) \\ u_3(\Delta t) \end{pmatrix}.$$

Nun geht der Vektor $\vec{u}(\Delta t)/\Delta t$ für $\Delta t \rightarrow 0$ nach $\vec{0}$ genau dann, wenn jede seiner Komponenten nach 0 geht. Das ergibt $x(t_0 + \Delta t) = x(t_0) + a_1\Delta t + u_1(\Delta t)$, und $u_1(\Delta t)/\Delta t \rightarrow 0$ für $\Delta t \rightarrow 0$. Also haben wir mit Lesen der ersten Komponentengleichung eine Tangentenerlegung von x an der Stelle t_0 , so dass $a_1 = x'(t_0)$ sein muss. Entsprechendes führen wir für die andern Komponenten durch.

Damit haben wir folgende Definition und das wichtige Resultat dazu gewonnen:

DEFINITION 35. Sei die Kurve \vec{x} in einer beidseitigen Umgebung von t_0 definiert. Dann heißt \vec{x} ableitbar (differenzierbar) an der Stelle t_0 genau dann, wenn eine Tangentenzersetzung der Form

$$\vec{x}(t_0 + \Delta t) = \vec{x}(t_0) + \vec{a}\Delta t + \vec{u}(\Delta t)$$

existiert, mit einem festen Vektor \vec{a} und einem Restvektor $\vec{u}(\Delta t)$, welcher die Restbedingung erfüllt: $\vec{u}(\Delta t) \rightarrow \vec{0}$ für $\Delta t \rightarrow 0$. Dann ist \vec{a} eindeutig bestimmt und heißt Ableitung von \vec{x} an der Stelle t_0 , Bezeichnung: $\vec{x}'(t_0)$, bei Zeitdeutung in der Physik gern $\dot{\vec{x}}(t_0)$.

SATZ 23. Eine Kurve \vec{x} ist genau dann differenzierbar in t_0 , wenn jede ihrer Komponentenfunktionen es ist. Der Ableitungsvektor besteht dann genau aus den Ableitungen der Komponentenfunktionen. (Das gilt für jede Dimension.) Geometrisch ist der Ableitungsvektor tangential zur Bahn der Kurve im Punkt $\vec{x}(t_0)$. Er gibt allerdings nur dann einen Richtungsvektor für die Tangente, wenn er nicht Null wird.

Beispiele: Für $\vec{x}(t) = (\cos t, \sin t)$ hat man $\vec{x}'(t) = (-\sin t, \cos t)$, und dieser Vektor steht tangential zur Kreiskurve im Punkt $(\cos t, \sin t)$. Er steht in diesem Beispiel senkrecht auf $\vec{x}(t)$, und wir wissen, dass die Tangenten an den Kreis senkrecht zu den zugehörigen Radiusvektoren stehen. Betrachten wir noch die Beschleunigung: $\vec{x}''(t) = (-\cos t, -\sin t)$. Auch das ist höchst plausibel: Beschleunigung in Richtung des Kreisinnern ist vonnöten, soll die Bewegung auf einer Kreisbahn erzwungen werden. Es ist auch instruktiv, mit einer schnelleren Bewegung wie $\vec{y}(t) = (\cos 2t, \sin 2t)$ zu vergleichen: Dann ist $\vec{y}'(t) = (-2\sin 2t, 2\cos 2t)$, die Geschwindigkeitsvektoren haben also doppelten Betrag, die Beschleunigungsvektoren $\vec{y}''(t) = (-4\cos 2t, -4\sin 2t)$ sogar vierfachen. Zweites Beispiel: $\vec{x}(t) = (\cos t, \sin t, t)$ hat $\vec{x}'(t) = (-\sin t, \cos t, 1)$, die Geschwindigkeit in z -Richtung ist also konstant, die Beschleunigung in dieser Richtung Null. Kommen wir zurück zur Flugparabel: In Komponentenform haben wir dafür $\vec{x}(t) = (x_0, y_0, z_0) + t(v_1, v_2, v_3) + \frac{1}{2}t^2(g_1, g_2, g_3)$. Die erste Komponentenfunktion lautet $x(t) = x_0 + tv_1 + \frac{1}{2}t^2g_1$, deren Ableitung also $v_1 + tg_1$, analog für die anderen, so dass wir unsere anfängliche Überlegung voll bestätigt sehen, dass man beim Ableiten mit konstanten Vektoren wie mit Zahlkonstanten korrekt umgeht. Wir können auch ohne weiteres folgende substantielle Anwendung auf Flugparabeln machen: Zur gegebenen Flugparabel bestimme man den Scheitelpunkt. Das wirkt auf den ersten Blick nicht einfach, es wird jedoch einfach, sobald man den Scheitelpunkt charakterisiert als den Punkt der Parabel, in dem Geschwindigkeits- und Beschleunigungsvektor senkrecht aufeinander stehen. Also genügt es, die Gleichung $\vec{x}'(t)\vec{x}''(t) = 0$ (Skalarprodukt!) für die Unbekannte t anzusetzen und zu lösen. Mit der Lösung t_0 gilt dann, dass $\vec{x}(t_0)$ der gesuchte Scheitelpunkt ist. Ein weiteres Beispiel: Eine Kurve in \mathbb{C} der Form $z(t) = a(t) + jb(t)$ ist wie eine Kurve im \mathbb{R}^2 abzuleiten, daher $z'(t) = a'(t) + jb'(t)$, vorausgesetzt natürlich, dass $a(t)$ und $b(t)$ ableitbar sind. Man prüfe auf diese Weise nach, dass man tatsächlich $\frac{d}{dt}e^{jt} = je^{jt} = e^{j(t+\pi/2)}$ bekommt, wie man auch bei naiver Anwendung der Kettenregel erwarten sollte. (Man verwechsle die Ableitung von Abbildungen $\mathbb{R} \rightarrow \mathbb{C}$ allerdings nicht mit der Ableitung einer Funktion $\mathbb{C} \rightarrow \mathbb{C}$, das ist etwas ganz Anderes!)

Wir können natürlich die Ableitung von Kurvenparametrisierungen ausnutzen, um das Problem zu bewältigen, eine Tangente an eine Kurvenbahn durch einen bestimmten Punkt zu legen. Zum Beispiel wollen wir an die Ellipse $\vec{x}(t) = (a \cos t, b \sin t)$, $a, b \neq 0$, im Punkt $\vec{x}(t_0)$ eine Tangente legen und eine Parameterdarstellung für diese Tangente gewinnen. Dafür benötigen wir einen Aufpunktvektor, der mit $\vec{x}(t_0)$ sofort gefunden ist, und einen Richtungsvektor, der mit $\vec{x}'(t_0) = (-a \sin(t_0), b \cos(t_0))$ sofort bereitsteht, da dieser Vektor nicht $\vec{0}$ ist (warum?). Also lautet eine Parameterdarstellung für die gewünschte Tangente: $\vec{y}_g(\lambda) = (a \cos t_0, b \sin t_0) + \lambda(-a \sin t_0, b \cos t_0)$, $\lambda \in \mathbb{R}$. Es wäre viel aufwendiger gewesen, eine Gerade so zu bestimmen, dass sie durch den Punkt $\vec{x}(t_0)$ geht und die Ellipse nur berührt, d.h. nicht zwei mal schneidet!

2.9. Ein kleiner Ausblick auf die Ableitung von Funktionen $\mathbb{R}^n \rightarrow \mathbb{R}$: Gradienten und Niveauflächen. Man beachte: Der Wertebereich ist hier eindimensional, der Definitionsbereich mehrdimensional, also gerade umgekehrt wie bei den Kurven. Dies Problem ist theoretisch schwieriger zu bewältigen, man hat dafür allerdings auch reichere Anwendungen. Aber wieder kommt man mit sinngemäßer Übertragung der Tangentenzersetzung aus, während Differenzenquotienten in diesem Falle völlig unpassend sind. Die größere theoretische Schwierigkeit bedeutet jedoch keinerlei rechnerische: Um etwa $f(x, y, z) = x^2 + y^2 + z^2$ abzuleiten, braucht man

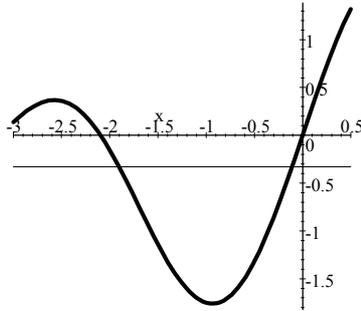
nur getrennt nach den verschiedenen unabhängigen Variablen abzuleiten, dabei die andern jeweils wie äußere Parameter zu behandeln. Diese Ableitungen nennt man dann *partielle* und schreibt dafür: $\frac{\partial f}{\partial x}(x_0, y_0, z_0) = 2x_0$, $\frac{\partial f}{\partial y}(x_0, y_0, z_0) = 2y_0$, $\frac{\partial f}{\partial z}(x_0, y_0, z_0) = 2z_0$ im Beispiel. Wenn nun diese partiellen Ableitungen existieren, so reicht das noch nicht für die Existenz der „totalen“ Ableitung von f . Aber wenn diese existiert, so lässt sie sich jedenfalls aus diesen partiellen Ableitungen zusammensetzen. Interessant ist nun der Vektor der partiellen Ableitungen. Man nennt ihn den Gradienten und schreibt dafür $\text{grad}f(x_0, y_0, z_0)$ oder auch $\text{grad}f(\vec{x}_0)$ in komponentenfreier Schreibweise. Im Beispiel also $\text{grad}f(x_0, y_0, z_0) = (2x_0, 2y_0, 2z_0)$. Er hat die interessante geometrische Bedeutung, senkrecht auf der Niveaulfläche des Skalarfeldes f im Punkt (x_0, y_0, z_0) zu stehen. Diese Niveaulfläche wird durch die Gleichung $f(x, y, z) = f(x_0, y_0, z_0)$ definiert, Bestimmungsgleichung mit x, y, z als Unbestimmten! Im Beispiel ist das die Kugelschale durch den Punkt (x_0, y_0, z_0) . Insbesondere kann man damit also das Problem lösen, eine Normale zu einer gekrümmten Fläche zu bekommen: Man definiert ein Skalarfeld, von dem die gegebene Fläche eine Niveaulfläche ist. Dann bildet man im gewünschten Punkt den Gradienten dieses Skalarfeldes. Ist er nicht Null, so hat man einen Normalenvektor zur Fläche in diesem Punkt. Im Beispiel war das besonders einfach und auch direkt zu sehen, aber bereits bei der Fläche (Bestimmungsgleichung!) $2x^2 + 3y^2 + 4z^2 = 4$ (das ist eine Ellipsoidoberfläche) wäre das nicht so direkt zu sehen, doch für einen Punkt dieser Fläche (x_0, y_0, z_0) bildet man nur den Gradienten $(4x_0, 6y_0, 8z_0)$ und ist fertig. Mittels des Gradienten kann man entsprechend auch etwa Fragen der Art beantworten, in welchem Winkel eine Gerade auf eine gekrümmte Fläche auftrifft usw. Eine weitere wichtige Anwendung des Gradienten fußt auf seiner Eigenschaft, in die Richtung des größten Anstieges der Feldwerte zu zeigen (wenn er nicht Null ist). Das nutzt man in numerischer Mathematik dahingehend aus, dass man experimentell ein Maximum einer mehrstelligen Funktion sucht, indem man an irgendeinem Punkt beginnt und dann jeweils ein geeignetes kleines Stück in Richtung des Gradienten wandert, immer iterativ so fort, mit immer kleineren Vielfachen des jeweiligen Gradientenvektors. („Gradienten-Suchmethode“). Sucht man ein Minimum, so geht man jeweils genau entgegengesetzt zum Gradienten fort. Das ist zwar keine ganz sichere Methode, liefert aber vielfach brauchbare Ergebnisse gerade bei solchen Funktionen, die eine unübersehbare Fülle von lokalen Extremwerten besitzt, die man unmöglich berechnen könnte, da man die zugehörigen nichtlinearen Gleichungssysteme nicht lösen kann. Solch ein Verfahren benutzt man übrigens beim Trainieren neuronaler Netzwerke.

Alles Gesagte gilt auch in höheren Dimensionen als 3, nur werden aus den Flächen eben Hyperflächen: Die Mengen der Punkte mit bestimmtem konstantem Feldwert eines Skalarfeldes $f : \mathbb{R}^n \rightarrow \mathbb{R}$, also die Niveaugebilde zu f , sind $n - 1$ - dimensionale Hyperflächen. Speziell für $n = 2$ hat man Niveaulkurven - man vergleiche die Höhenlinien auf Landkarten, Kurven gleicher Temperatur oder gleichen Luftdrucks auf einer solchen Karte (Isothermen bzw. Isobaren genannt).

3. Integrale

3.1. Die Idee des Integrals. Zurück zu den eindimensionalen gewöhnlichen reellen Funktionen. Es sollte einleuchten, dass es nicht nur wichtige *lokale* Merkmale gibt, wie sie durch Ableitungen, eventuell höhere, beschrieben werden, sondern auch wichtige *globale*. Ein Beispiel war Monotonie in einem ganzen Intervall. Zentrale Bedeutung hat immer wieder die Frage nach dem Mittelwert einer Funktion in einem Bereich. Kennt man etwa die mittlere Geschwindigkeit (vektoriell) einer Zeitspanne und den Anfangsort zu Beginn dieser Zeitspanne, so weiß man, wo man gelandet ist: Anfangs-Ortsvektor plus Zeitspanne mal mittlerer Geschwindigkeitsvektor. Ein Mittelwert ist offenbar abhängig von allen Werten einer bestimmten Menge. Zum Beispiel ist der Mittelwert des Einkommens in einem Jahr das arithmetische Mittel der einzelnen Einkommen in den einzelnen Monaten dieses Jahres. Ebenso ermittelt man Messwerte als arithmetische Mittel einer endlichen Reihe von Einzelmessungen. Aber was ist der Mittelwert *aller* Zahlen x^2 , für $x \in [0, 1]$? Das sind überabzählbar viele Werte, die hier zu mitteln wären. Der Rechenausdruck des arithmetischen Mittels hilft hier nichts; dennoch hat man den Eindruck, dass es nicht völlig abwegig wäre, etwa die Werte $0^2, 0.1^2, 0.2^2 \dots$ bis 1^2 zu bilden und davon das arithmetische Mittel auszurechnen: Das ergibt $(0 + 0.1^2 + \dots + 0.9^2 + 1^2) / 11 = 0.35$ und ist tatsächlich nicht so schlecht - $1/3$ wäre korrekt, wie wir sehen werden. Die Idee besteht nun darin, den genauen Mittelwert einer Funktion auf

einem Intervall geometrisch zu verstehen, nämlich einen **Zusammenhang mit den Flächeninhalten der Flächen zwischen dem Graphen der Funktion und der x -Achse** herzustellen. Intuitiv erkennt man nämlich bei einem Funktionsgraphen wie im folgenden Beispiel, wo der Mittelwert - für das gezeichnete Intervall - liegen sollte, angedeutet durch die horizontale Gerade:



Es sollten die Flächeninhalte, die zwischen dem Graphen und der Mittelwertlinie liegen, oberhalb und unterhalb dieser Linie einander die Waage halten. Mittelwert 0 wäre im Beispiel deutlich zu hoch! Das bedeutet aber: Der Flächeninhalt zwischen dem Graphen und der x -Achse, wobei die Flächen oberhalb der x -Achse positiv gezählt werden und die Fläche unterhalb der x -Achse negativ, muss gleich dem Flächeninhalt zwischen der x -Achse und der Mittelwertlinie sein, wobei auch diese Rechteckfläche positiv gezählt wird, wenn sie oberhalb der x -Achse liegt und negativ, wenn sie unterhalb liegt wie im Beispiel. Wir definieren und bezeichnen nunmehr die wesentlichen Bestandteile dieser Beobachtung:

DEFINITION 36. Für eine stückweise stetige Funktion f auf dem Intervall $[a, b]$, $a < b$, mit allenfalls endlich vielen Sprüngen definieren wir anschaulich:

$$\int_a^b f(x)dx = \text{der Flächeninhalt mit Orientierungsvorzeichen}$$

zwischen dem Graphen von f auf $[a, b]$ und der x -Achse
(„bestimmtes Integral von a nach b über die Funktion f “).

$$\bar{f}_{[a,b]} = \text{Mittelwert von } f \text{ auf } [a, b].$$

Nach unseren Beobachtungen sollte folgende Beziehung bestehen:

$$\bar{f}_{[a,b]} = \frac{1}{b-a} \int_a^b f(x)dx.$$

Damit sollte von vornherein klar sein: Mittelwerte und folglich auch Integrale sollten negativ sein können. Es ist kein Defekt des Integrals, sondern gerade seine Tugend, so eng mit dem Mittelwertbegriff zusammenzuhängen, dass es nicht so direkt geeignet ist, den absoluten Flächeninhalt zwischen dem Graphen und der x -Achse zu berechnen (dafür dürfte man in der Tat bei Vorzeichenwechseln immer nur von einer Nullstelle zur andern integrieren!). Für diesen Zusammenhang ist es gerade richtig, dass das Integral die Flächen oberhalb der x -Achse positiv und die andern negativ zählt. Zur Erläuterung des Wortes „Orientierungsvorzeichen“ denke man daran, dass bei Sicht längs der x -Achse von a , der unteren Integralgrenze, nach b , der oberen, die positiv gezählten Flächen links erscheinen. Schaut man dagegen von b nach a , so erscheinen die Flächen oberhalb der x -Achse rechts, die andern links. Zählen wir also dabei wieder positiv, was links erscheint, so erhalten wir $-\int_a^b f(x)dx$. Daher definieren wir sofort:

DEFINITION 37. Für eine Funktion mit denselben Voraussetzungen wie in der vorigen Definition und $a < b$ definieren wir:

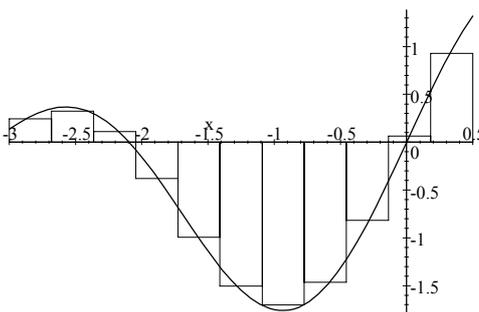
$$\int_b^a f(x)dx = - \int_a^b f(x)dx.$$

Bemerkung: Der Zusammenhang mit dem Mittelwert bleibt genau erhalten, es gilt auch

$$\bar{f}_{[a,b]} = \frac{1}{a-b} \int_b^a f(x)dx.$$

Der Grund für die Bemerkung: Auch der Bruch vor dem Integral hat sein Vorzeichen gewechselt.

Folgende Skizze zeigt nun, wie man naheliegender näherungsweise ein bestimmtes Integral berechnen kann:



Man teilt also das ganze Integrationsintervall durch n (nicht notwendig gleich breite) Streifen der Breiten Δx_i ein und wählt in jedem Teilintervall einen Punkt x_i , dann hat man:

$$\int_a^b f(x)dx \approx \sum_{i=n}^1 f(x_i)\Delta x_i.$$

Das stimmt auch für $b < a$, wenn man allgemein setzt: $a = a_0$, $b = a_n$, a_0, a_1, \dots, a_n monoton, $\Delta x_i = a_i - a_{i-1}$, $i \geq 1$. Offensichtlich wird die Näherung besser bei immer feinerer Einteilung. Auf diese Weise wird auch die merkwürdige Integralschreibweise motiviert - seit der Erfindung durch Leibniz (Ende 17. Jahrhundert) hat sie sich gehalten - das Integralzeichen ist ein stilisiertes Summenzeichen, und der Summenindex verschwindet, mit der Idee, über das Kontinuum zu summieren bei infinitesimal kleiner Streifenbreite dx (für Δx , das nicht infinitesimal klein ist). Mit der Summenapproximation des Integrals können wir nun auch das zuerst anvisierte Verfahren rechtfertigen, den Mittelwert $\bar{f}_{[a,b]}$ durch ein endliches arithmetisches Mittel zu nähern: Bei $\Delta x_i = \Delta x$ konstant haben wir $\Delta x = (b - a)/n$, also

$$\frac{1}{n} \sum_{i=n}^1 f(x_i) = \frac{\Delta x}{b-a} \sum_{i=n}^1 f(x_i) = \frac{1}{b-a} \sum_{i=n}^1 f(x_i)\Delta x \approx \frac{1}{b-a} \int_a^b f(x)dx = \bar{f}_{[a,b]}.$$

Wir kennen nun bereits Möglichkeiten, Integral und Mittelwert näherungsweise auszurechnen. Aber dieser Zusammenhang wird uns unmittelbar auch zu einer Möglichkeit exakter Berechnung führen, die allerdings nur für gewisse Funktionen praktisch gangbar sein wird, vgl. den nächsten Abschnitt. Daher behalten Näherungsmethoden eine überragende Bedeutung! Hier wollen wir zunächst noch die Idee verallgemeinern und zeigen, wie typisch Integrale aus inhaltlichen Aufgabenstellungen resultieren: Man hat eine Größe, die näherungsweise durch Summation über kleine Stücke eines Kontinuums (das kann eindimensional sein, wie hier zunächst ausschließlich betrachtet, aber auch zwei-, drei- oder noch höher dimensional) berechnet werden kann. Dann ist diese Größe durch ein Integral zu berechnen. Dazu Beispiele: Fährt man mit ungleichmäßiger Geschwindigkeit $v(t)$, so hat man für den zurückgelegten Weg kleine $v(t)\Delta t$ aufzusummieren, also $\int_{t_0}^{t_1} v(t)dt$ zu berechnen. Oder

gleichwertig: Man multipliziert die mittlere Geschwindigkeit $\frac{1}{t_1-t_0} \int_{t_0}^{t_1} v(t) dt$ mit der Zeitspanne $t_1 - t_0$. Hat man einen inhomogenen Körper mit einer kontinuierlichen Massenverteilung $\rho(\vec{x})$, so errechnet sich die Gesamtmasse durch Integration über das dreidimensionale Gebiet des Körpers der Funktion $\rho(\vec{x})$. Ebenso errechnet man eine Gesamtladung durch Integration der Ladungsdichte, eine Wahrscheinlichkeit durch Integration über eine Verteilungsdichte bzw. Wahrscheinlichkeitsdichte. Dazu das wichtigste Beispiel: Die Dichte der Standard-Normalverteilung lautet

$$\varphi_{0,1}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad x \in \mathbb{R}.$$

Die Wahrscheinlichkeit dafür, dass eine so verteilte Größe einen Wert $\leq a$ annimmt, ist dann

$$\int_{-\infty}^a \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \quad (\text{nur näherungsweise zu berechnen!})$$

Hier taucht $-\infty$ als Integrationsgrenze auf. Das bedeutet einfach: $\int_{-\infty}^a \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \lim_{b \rightarrow -\infty} \int_b^a \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$, und man wird mit der Existenz dieses Grenzwertes rechnen, weil der Integrand für $x \rightarrow -\infty$ sehr schnell nach Null geht. Daraus ergeben sich weitere Beispiele: Hat man eine zufällige Größe mit endlich vielen Werten x_i , die jeweils mit der Wahrscheinlichkeit p_i angenommen werden, so ist der Erwartungswert μ auszurechnen als $\sum x_i p_i$. Der Erwartungswert, auch einfach Mittelwert genannt, hat die Bedeutung: Führt man das Experiment sehr oft durch und bildet das arithmetische Mittel aller beobachteten Werte, so landet man nach Wahrscheinlichkeit nahe beim Erwartungswert. (Man kann sogar ausrechnen, mit welcher Wahrscheinlichkeit wie nahe). Der Erwartungswert einer Größe, die auf einem Kontinuum ihre Werte hat und deren Wahrscheinlichkeitsverteilung durch eine Dichtefunktion f gegeben ist, berechnet man folgerichtig als $\int x \cdot f(x) dx$ - zu integrieren über den Bereich aller Werte, in der Regel ein endliches oder unendliches Intervall. Man stelle sich unter $f(x) dx$ die infinitesimale Wahrscheinlichkeit vor, mit der genau der Wert x zu beobachten ist. So ist der Erwartungswert einer standardnormalverteilten Größe $\int_{-\infty}^{\infty} x \varphi_{0,1}(x) dx = 0$. Ebenso wichtig bei Wahrscheinlichkeitsverteilungen ist die Varianz σ^2 , die mittlere quadrierte Abweichung der Einzelwerte vom Mittelwert μ . Bei einer Dichte f lautet sie. $\sigma^2 = \int (x - \mu)^2 f(x) dx$. Für die Standard-Normalverteilung ergibt sich $\sigma^2 = \int_{-\infty}^{\infty} x^2 \varphi_{0,1}(x) dx = 1$. Sie heißt Standard-Normalverteilung, weil sie auf die einfachsten Parameter-Werte $\mu = 0, \sigma^2 = 1$ normiert ist. Machen wir dazu ein praktisches Beispiel: Messwertfehler sind im allgemeinen recht genau normalverteilt. Wiederholen wir eine Messung vielfach (n mal) unabhängig, so können wir näherungsweise die Varianz der Messwerte ermitteln, σ^2 . Dazu rechnet man das arithmetische Mittel aller Messwerte \bar{x} aus und dann $s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$ als Schätzwert für σ^2 . s^2 ist dann eine Schätzung für die Varianz der Messfehler. Sind diese nun normalverteilt mit Erwartungswert Null (d.h. wir haben keinen systematischen Fehler, stets zu hohe oder zu tiefe Werte zu messen - so etwas kommt natürlich unglücklicherweise auch vor), dann kann man die Wahrscheinlichkeit ausrechnen, mit welcher der absolute Messfehler unter einer Schwelle $\delta > 0$ bleibt: Das ist $\int_{-\delta/s}^{\delta/s} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$, mit $s = \sqrt{s^2}$. (s heißt auch Standardabweichung.) Zum Beispiel ergibt sich mit $\delta = 1/1000$ mm bei einer Längenmessung und beobachtetem $s = 1/1000$ mm eine Wahrscheinlichkeit von etwa 0.68 oder 68% dafür, dass eine Einzelmessung einen Fehler von höchstens $1/1000$ mm aufweist. Das kann man verbessern, wenn man statt einzelner Messungen mehrere - sagen wir n unabhängige Messwerte mittelt; dann reduziert sich die Varianz mit Faktor $1/n$. Bei 25 gemittelten Messungen erhalten wir im Beispiel $s' = 1/5000$ mm, also ist die Wahrscheinlichkeit dafür, mit der Messung durch Mittelwertbildung von 25 Einzelmessungen unter $1/1000$ mm Messfehler zu bleiben: $\int_{-5}^5 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$, und das ist so gut wie 1. Man kann also sicher sein. Es ist sogar noch die Wahrscheinlichkeit dafür, unter $1/2000$ mm Messfehler zu bleiben: $\int_{-2.5}^{2.5} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = .98758$, also fast 99% sicher.

3.2. Grundlage der exakten Berechnung von bestimmten Integralen. Wie angekündigt, führt der schöne Zusammenhang

$$\overline{f_{[a,b]}} = \frac{1}{b-a} \int_a^b f(x) dx$$

zum Ziel. Dazu nehmen wir an, wir hätten eine Funktion F mit der Eigenschaft $F' = f$. An jeder Stelle x gibt f also die Steigung von F . Wir betrachten nunmehr die mittlere Steigung von F auf $[a, b]$. Das ist einerseits der bekannte Differenzenquotient:

$$\text{mittlere Steigung von } F \text{ auf } [a, b] = \frac{F(b) - F(a)}{b - a}.$$

Andererseits haben wir

$$\text{mittlere Steigung von } F \text{ auf } [a, b] = \overline{f_{[a,b]}}.$$

Das ergibt zusammen:

$$\overline{f_{[a,b]}} = \frac{F(b) - F(a)}{b - a}.$$

Daraus folgt mit

$$\overline{f_{[a,b]}} = \frac{1}{b-a} \int_a^b f(x) dx,$$

dass gilt:

SATZ 24 (Hauptsatz der Differential- und Integralrechnung). *Wenn f auf $[a, b]$ stetig („stückweise“ genügt) ist und F eine Stammfunktion von f auf $[a, b]$ (d.h. $F' = f$ auf $[a, b]$), dann gilt:*

$$(3.1) \quad \int_a^b f(x) dx = F(b) - F(a), \text{ und entsprechend}$$

$$\overline{f_{[a,b]}} = \frac{F(b) - F(a)}{b - a}.$$

Weiter gilt auch noch, dass unter der genannten Voraussetzung für f stets eine Stammfunktion F zu f existiert. (Allerdings ist es vielfach unmöglich, einen üblichen Rechenausdruck für F zu berechnen, einfach, weil ein solcher nicht existiert!)

Es sei bemerkt, dass wir das nicht streng bewiesen haben, wohl aber eine sehr gute und präzisierbare intuitive Argumentation dafür anführten. Die systematische Anwendung geschieht im nächsten Abschnitt.

3.3. Praktische Berechnung von bestimmten und unbestimmten Integralen: Formeln und Beispiele. Grundlage ist der Hauptsatz. Seine Anwendung besteht natürlicherweise in zwei Schritten: Zuerst sucht man eine Stammfunktion F zu f , dann setzt man die Grenzen des Integrals ein und bildet die Differenz. Zu diesem Zweck ist es nützlich, folgende (übliche!) Notation einzuführen:

$$[F(x)]_a^b = F(b) - F(a),$$

und nun rechnet man z.B. für das Eingangsbeispiel zur Integration:

$$\int_0^1 x^2 dx = \left[\frac{x^3}{3} \right]_0^1 = \frac{1^3}{3} - \frac{0^3}{3} = \frac{1}{3}.$$

Das ist also auch der Mittelwert von x^2 auf dem Intervall, da seine Breite 1 ist. Hier sind zwei Formeln zum Umgang mit den Grenzen bei bestimmten Integralen. Die erste versteht sich unmittelbar aus der Flächendeutung, die zweite aus der Bedeutung der Orientierung:

$$(3.2) \quad \int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx$$

$$\int_a^b f(x)dx = -\int_b^a f(x)dx.$$

Ansonsten geht es im wesentlichen um das Auffinden von Stammfunktionen zu einer gegebenen Funktion. Dabei beachten wir, dass mit F auch stets $F + c$, c eine Konstante (sogenannte „Integrationskonstante“), eine Stammfunktion von f ist, dass man damit aber auch *alle* Stammfunktionen von f beschrieben hat, wie wir nach Satz wissen. Im folgenden Text wollen wir das ewige Wiederholen von „+c“ vermeiden und müssen dann nur daran denken, dass wir stets nur *eine* Stammfunktion beschreiben. Dazu ist folgende Notation in Gebrauch:

$$\int f(x)dx = F(x) (+c), \text{ mit } F' = f. \text{ (Unbestimmtes Integral)}$$

Der Vorgang verläuft wie bei den Ableitungsformeln: Formeln für Grundintegrale und für zusammengesetzte Funktionen, für letztere aber wesentlich unvollständiger bzw. gestörter.

3.3.1. *Grundintegrale.* Durch Umkehren der zugehörigen Ableitungsregeln erhalten wir unmittelbar - man denke daran, dass Addition einer Integrationskonstanten c ebenfalls zu einer Stammfunktion führt:

$$(3.3) \quad \int x^a dx = \frac{x^{a+1}}{a+1}, \text{ für } a \neq -1,$$

$$(3.4) \quad \int \frac{1}{x} dx = \ln|x|,$$

$$(3.5) \quad \int \frac{1}{1+x^2} dx = \arctan(x),$$

$$(3.6) \quad \int e^x dx = e^x,$$

$$(3.7) \quad \int \sin(x) dx = -\cos(x),$$

$$(3.8) \quad \int \cos(x) dx = \sin(x),$$

Wir fügen noch hinzu (zum Auffinden s. u., aber Nachprüfen durch Ableiten von $x \ln(x) - x$ ist sofort möglich):

$$(3.9) \quad \int \ln(x) dx = x \ln(x) - x.$$

3.3.2. *Integrationsregeln für Zusammensetzungen von Funktionen.* Linearität des Integrals:

$$(3.10) \quad \int (f(x) + g(x)) dx = \int f(x) dx + \int g(x) dx,$$

$$\int cf(x) dx = c \int f(x) dx.$$

Damit hat man z.B.

$$\int \left(3 + 2x + 5e^x - \frac{4}{x} \right) dx = 3x + 2\frac{x^2}{2} + 5e^x - 4 \ln|x|.$$

Man integriert also in Summen mit konstanten Faktoren einfach gliedweise, ähnlich wie beim Ableiten. (Nur sehr eingeschränkt taugliche) Produktregel: „Partielle Integration“:

$$(3.11) \quad \int F(x)g(x)dx = F(x)G(x) - \int f(x)G(x)dx.$$

Es bleibt also ein Integral übrig, das nur hoffentlich lösbar ist. Typisch ist das anzuwenden, wenn ein Faktor ein Polynom, der andere Exponential- oder Logarithmusfunktion (beide allenfalls leicht verändert) ist, Anwendungsbeispiel:

$$\int \ln(x)dx = \int (\ln(x) \cdot 1)dx = \ln(x) \cdot x - \int \frac{1}{x}x dx = x \ln(x) - x.$$

Es ist sehr einfach, die Regel aus der Produktregel des Differenzierens herzuleiten, sie ist einfach die Umkehrung davon:

$$\begin{aligned} (FG)' &= fG + Fg, \text{ also} \\ \int (f(x)G(x) + F(x)g(x))dx &= F(x)G(x), \text{ daher} \\ \int f(x)G(x)dx + \int F(x)g(x)dx &= F(x)G(x), \text{ nun erstes Integral hinüber.} \end{aligned}$$

$\frac{1}{\alpha}$ -Regel für lineare Transformationen von Funktionen, zu denen man schon Stammfunktionen kennt:

$$(3.12) \quad \int f(\alpha x + \beta)dx = \frac{1}{\alpha}F(\alpha x + \beta), \text{ wobei } \alpha \neq 0 \text{ und } F' = f \text{ sind.}$$

Anwendungsbeispiele:

$$\begin{aligned} \int (2x + 3)^{100} dx &= \frac{1}{2} \cdot \frac{(2x + 3)^{101}}{101}, \\ \int \frac{1}{2x + 3} dx &= \frac{1}{2} \ln |2x + 3|, \\ \int 2^x dx &= \int e^{x \ln(2)} dx = \frac{1}{\ln(2)} e^{x \ln(2)} = \frac{1}{\ln(2)} \cdot 2^x. \end{aligned}$$

Auch diese Regel bestätigt man unmittelbar durch Bildung von

$$\frac{d}{dx} \frac{1}{\alpha} F(\alpha x + \beta) = \frac{1}{\alpha} \alpha f(\alpha x + \beta) = f(\alpha x + \beta).$$

Umkehrung der Kettenregel (oder Substitutionsregel) - wir schreiben das dx hier links, das ist oft nützlich, insbesondere auch bei mehrfachen Integralen:

$$(3.13) \quad \int dx f'(x)g(f(x)) = \int du g(u) = G(f(x)), \text{ mit } G' = g.$$

Wieder erhält man durch Ableiten unmittelbar die Bestätigung:

$$\frac{d}{dx} G(f(x)) = g(f(x))f'(x).$$

Der eingefügte Zwischenschritt ist sehr nützlich: Für $f(x)$ wird die neue Integrationsvariable u eingeführt, und man hat die Ersetzungszeilen:

$$\begin{aligned} u &= f(x), \\ du &= f'(x)dx, \text{ gemäß } \frac{du}{dx} = f'(x). \end{aligned}$$

Nun kann man formal ersetzen:

$$\int \underbrace{dx f'(x)}_{du} \underbrace{g(f(x))}_{g(u)} = \int dug(u) = G(u) \underset{\text{Rückeinsetzen}}{=} G(f(x)).$$

Anwendungsbeispiel:

$$\int dx \frac{\ln x}{x} = \int du \cdot u = \frac{u^2}{2} = \frac{1}{2} \ln^2(x),$$

wie man folgendermaßen sieht:

$$\begin{aligned} u &= \ln(x), \quad g(u) = u, \\ du &= \frac{1}{x} dx. \end{aligned}$$

Weiteres Anwendungsbeispiel:

$$\begin{aligned} \int dx \cdot x \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} &= - \int dx(-x) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} = - \int du \frac{1}{\sqrt{2\pi}} e^u \\ &= - \frac{1}{\sqrt{2\pi}} e^u = - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \end{aligned}$$

mit der Substitution

$$\begin{aligned} u &= -\frac{1}{2}x^2, \quad g(u) = e^u, \\ du &= -x. \end{aligned}$$

Damit hat man

$$\int_0^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \lim_{b \rightarrow \infty} \int_0^b x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \lim_{b \rightarrow \infty} - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}b^2} = 0,$$

ebenso ergibt sich

$$\int_{-\infty}^0 x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \lim_{a \rightarrow -\infty} - \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}a^2} = 0,$$

daher ist der Erwartungswert einer standard-normalverteilten Größe

$$\int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \int_{-\infty}^0 x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx + \int_0^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 0.$$

Wir haben die Gelegenheit benutzt, nebenbei zu zeigen, wie man ein Integral mit unendlicher Grenze als Grenzwert von Integralen mit endlichen Grenzen berechnet. Aber das Integral, das man hauptsächlich bei der Standard-Normalverteilung benötigt, nämlich

$$\int dx \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2},$$

gerade das kann man nicht ausrechnen, sondern nur die Werte nähern wie in Beispielen weiter oben geschehen!

Index

- Abbildung
 - durch eine Matrix definiert, 33
- Ableitung einer Funktion
 - als Limes von Differenzenquotienten, 65
- Ableitungsregeln
 - Kettenregel, 71
 - Linearität, 71
 - Produkt- und Quotientenregel, 71
 - Regel für Umkehrfunktionen, 71
- Addition
 - komplexer Zahlen, 46
- allgemeine Lösung
 - eines linearen Gleichungssystems, 34
- Arcusfunktionen, 56
- Asymptote, 53
- Aufpunktvektor, 21

- Basis
 - eines Vektorraums, 31
- Bestimmungsgleichung
 - mit äußeren Parametern, 8
- Betrag
 - einer komplexen Zahl, 47
- Betrag (Absolutbetrag)
 - eines Vektors, 37
- Bild einer linearen Abbildung, 34
- Bilinearität, 38
- Binomialkoeffizienten, 3
- binomische Formeln
 - auch allgemeine, 3
- Bruchrechnung
 - Grundgesetze, 2

- Cauchy-Schwarzsche Ungleichung, 40
- Cosinus
 - Darstellung über das Skalarprodukt, 40

- Darstellungsformen geometrischer Gebilde, 13
- Determinante
 - in Dimensionen 2 und 3, 44
- Dimension eines Vektorraums, 30
- Dreiecksungleichung, 37
- Durchschnitt
 - zweier Mengen, 25

- Einsetzen, 7
- Erzeugendensystem
 - für einen Vektorraum, 31
- Exponentialfunktionen, 58
- exponentielles Wachstum, 60

- Fallunterscheidungen bei äußeren Parametern, 8
- Formel
 - allgemeingültige, 2, 7
 - definitorische, 8
 - unter spezieller Interpretation allgemeingültig, 6
- Funktion
 - Begriff einer reellen Funktion, 53
 - gerade, ungerade, 54
 - monoton steigend bzw. fallend, 54
- Funktionen
 - Exponential- und Logarithmusfunktionen, 58
 - Potenzfunktionen, 61
 - trigonometrische, 55

- Grenzwert
 - bei Zahlenfolgen und Funktionen, 58
- Grundfunktionen
 - im engeren und weiteren Sinne, 62
- Grundintegrale, 86

- Hauptsatz
 - der Differential- und Integralrechnung, 85
- Hinterinanderschaltung von Funktionen, 62
- homogenes lineares Gleichungssystem, 33

- Imaginärteil
 - einer komplexen Zahl, 46
- inhomogenes lineares Gleichungssystem, 34
- Integral
 - anschaulicher Begriff des bestimmten Integrals, 82
 - Näherung eines bestimmten Integrals, 83
 - praktische (exakte) Berechnung, so weit möglich, 86
- Integrationskonstante, 86
- Integrationsregeln
 - 1/alpha - Regel für lineare Transformationen im Urbildbereich, 87
 - Linearität, 87
 - partielle Integration, 87
 - Umkehrung der Kettenregel bzw. Substitutionsregel, 88
- inverse Funktion, 54

- Körperaxiome, 1
- kartesische Koordinaten
 - einer komplexen Zahl, 48

- Kern einer linearen Abbildung, 33
- Koeffizientenvergleich, 25
- komplexe Zahlen, 46
- Konjugation
 - komplexe, 47
- Konstante, 7
- Koordinatendarstellungen, 12
- Koordinatensysteme, 12
- Kreis, 14
- Kurvendiskussion, 63
 - Leitfragen, 63
- Lösungsmenge, 8
- Limes
 - bei Zahlenfolgen und Funktionen, 58
- lineare Abhängigkeit, 30
- lineare Algebra, 30
- lineare Transformation
 - von Funktionen, 55
- lineare Unabhängigkeit, 30
- Linearkombination von Vektoren, 21
- Logarithmusfunktionen, 59
- Lokale Extrema, 75
- Matrix, 32
 - Matrix mal Vektor, 32
- Menge aller (reellen) Zahlentripel, 13
- Mittelwertsatz, 76
- Monotonie
 - Begriff der Monotonie bei Funktionen, 54
 - und Ableitungseigenschaften, 77
- Multiplikation
 - komplexer Zahlen, 46
- Multiplikation eines Vektors mit einem Skalar, 17
- Näherung 0. Ordnung, 66
- Näherung 1. Ordnung, 68
- Normalenform
 - für Hyperebenen, 41
- Parameter
 - äußerer, 6
 - freier, 7
- Parameterdarstellung, 14
 - einer Ebene, 23
 - einer Geraden, 21
- Parameterdarstellung eines Kreises, 15
- Pol
 - einer Funktion an einer Stelle, 53
- Polarkoordinaten
 - einer komplexen Zahl, 48
- Potenzfunktionen, 61
- quadratische Ergänzung, 3
- quadratische Gleichungen
 - Lösungsformel, 3
- Realteil
 - einer komplexen Zahl, 46
- Rechenausdrücke
 - rekursiver Aufbau, 62
- rekursive Definition, 3
- reziproke Funktion, 54
- Richtungsvektor, 21
- Rollen von Buchstaben in Formeln und Gleichungen, 9
- Satz vom endlichen Zuwachs, 77
- Schnittmenge
 - zweier Mengen, 25
- senkrechte Projektion
 - eines Vektors auf einen anderen, 40
- Skalarprodukt, 38
- Spatprodukt, 43
- stumme Variable, 7
- Summenzeichen
 - großes, 4
- transzendente Zahl, 59
- triviale Linearkombination, 30
- Umkehrbarkeit
 - einer Funktion, 54
- Umkehrfunktion, 54
- Unterraum
 - (Vektor-), 30
 - affiner, 29
 - von einem System von Vektoren erzeugt, 31
- Urbildraum, 33
- Variable
 - freie, 2, 7
- Vektoraddition, 18
- Vektorprodukt
 - Berechnungsformel, 42
 - geometrische Charakterisierung, 42
- Vektorraum über einem Körper K , 19
- Vektorraumaxiome, 19
- Vektorraumoperationen, 17
- Verkettung von Funktionen, 62
- Werteraum, 33