Brückenkurs Mathematik

Guido Walz / Frank Zeilfelder / Thomas Rießinger

Brückenkurs Mathematik

für Studieneinsteiger aller Disziplinen

3. Auflage

Spektrum
AKADEMISCHER VERLAG

Autoren
Prof. Dr. Guido Walz
PD Dr. Frank Zeilfelder
Universität Mannheim, Institut für Mathematik, 68131 Mannheim

Prof. Dr. Thomas Rießinger
Fachhochschule Frankfurt a.M., Fachbereich 2, 60318 Frankfurt a.M.

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Springer ist ein Unternehmen von Springer Science+Business Media
springer.de

3. Auflage 2011

Spektrum Akademischer Verlag ist ein Imprint von Springer

11 12 13 14 15 5 4 3 2 1

Planung und Lektorat: Dr. Andreas Rüdinger, Anja Groth
Redaktion: Heidemarie Wolter
Satz: EDV-Beratung Frank Herweg, Leutershausen
Umschlaggestaltung: SpieszDesign, Neu-Ulm
Titelfotografie: © age; Mauritius

ISBN 978-3-8274-2763-2

Inhaltsverzeichnis

Die Mathematik als Fachgebiet ist so ernst, dass man keine Gelegenheit versäumen sollte, dieses Fachgebiet unterhaltsamer zu gestalten.
Blaise Pascal

Einleitung und Vorwort

Auch heute noch soll es ab und zu vorkommen, dass jemand in einer stillen Stunde ein Buch in die Hand nimmt, um darin zu schmökern; das wird allerdings in den seltensten Fällen ein Mathematik-Buch sein, aber das liegt nicht so sehr an der Mathematik selbst, sondern an der Art und Weise, wie sie in vielen Büchern dargestellt wird. Wenn Ihnen hierbei Begriffe wie ‚trocken', ‚spröde' oder ‚unverständlich' einfallen, sind wir schon auf einem guten Wege zu einer gemeinsamen Einschätzung.

Die Mathematik hat in der Gesellschaft leider einen hohen Igitt!-Faktor, und das liegt nicht selten an traumatischen Schulerlebnissen und schlechten Lehrbüchern für Schüler. Mit einem solchen schulischen Werdegang hat man aber meist Angst vor den mathematischen Teilen des Studiums, und ganz im Sinne einer selbst erfüllenden Prophezeiung bekommt man dann da auch wirklich Probleme.

Mit diesem Buch versuchen wir, Ihnen diese Einstiegsprobleme zu ersparen, indem wir Ihnen auf unterhaltsame Weise eine Brücke bauen, die Sie sanft über alle Untiefen hinweg ins Innere der Hochschulmathematik hineingeleitet. Diese Brücke beginnt auf der einen Seite beim einfachen Zahlenrechnen, wie es Ihnen vermutlich in der Mittelstufe schon begegnet ist, und führt Sie hinüber bis zu den Grundlagen von Linearer Algebra, Differenzialrechnung und Wahrscheinlichkeitsrechnung, die die Hauptinhalte Ihrer ersten Semester darstellen werden. Diesen Inhalten werden Sie dort immer gegenüberstehen, und bei deren Behandlung können Sie dann beruhigt sagen: „Kenn' ich schon!"

Das Buch wendet sich an Studierende aller Fachrichtungen, die im Laufe des Studiums oder auch Berufslebens mit Mathematik in Berührung kommen. Eines unserer Hauptanliegen ist es, zu beweisen, dass es möglich ist, ein Mathematik-Buch zu schreiben, das man von vorne bis hinten einfach lesen kann, ohne im Formalismus oder in humorloser Trockenheit verloren zu gehen, das einem nach dem Lesen aber dennoch das nötige Wissen und die fachliche Sicherheit für einen erfolgreichen Mathematikteil des Studiums vermittelt hat.

Eine große Zahl von Übungsaufgaben, deren Lösungen Sie am Ende des Buches finden, hilft bei der weiteren Vertiefung des Gelesenen. Ganz bewusst haben wir auf Beweise größtenteils verzichtet – damit werden Sie sich im Verlauf Ihres Studiums noch genügend oft herumschlagen müssen –, vielmehr wollen wir Ihnen ganz einfach Lust auf Mathematik machen, auch wenn Ihnen das momentan noch als Gegensatz erscheinen mag.

Ein großes Ärgernis ist in unseren Augen der ‚Krankenschwestern-Plural', den viele Lehrbuchautoren an den Tag legen („Wie geht's uns denn heute?", „Morgen wird uns also die Leber herausoperiert!", „Jetzt integrieren wir mal eine Funktion!"). Schrecklich! (Dies bezieht sich wohlgemerkt nur auf diese Formulierungsunart, *nicht* auf den Berufsstand der Krankenschwestern und -pfleger, eine Gruppe, deren Vertreter jeder für sich vermutlich mehr für die Menschheit leistet als alle Mathematikprofessoren der Welt zusammen.) Wir haben uns, um diese Unsitte zu vermeiden, dafür entschieden, Sie in jedem Kapitel ganz persönlich im Ich-Stil anzureden – jeder von uns hat diejenigen Kapitel verfasst, deren Inhalte er kompetent und (hoffentlich) instruktiv vermitteln kann, und dieses soll auch durch den Schreibstil ausgedrückt werden.

Und nun (und das meinen wir ernst!): Viel Spaß beim Studium der folgenden Seiten!

Sommer 2005

Guido Walz
Frank Zeilfelder
Thomas Rießinger

Vorwort zur zweiten Auflage

Die erfreulich gute Akzeptanz unseres Buches machte bereits jetzt eine zweite Auflage erforderlich. Wir haben die Gelegenheit genutzt, um einige kleinere Fehler im Text, auf die wir von aufmerksamen Leserinnen und Lesern hingewiesen wurden, zu beseitigen.

Außerdem haben wir ein neues Kapitel über die Grundlagen komplexer Zahlen aufgenommen. Wir hoffen, dass Ihnen dieses ebenso viel Spaß macht wie der restliche Text.

Frühjahr 2007

Guido Walz
Frank Zeilfelder
Thomas Rießinger

Vorwort zur dritten Auflage

Trotz aller Bemühungen um Fehlerfreiheit – ein Zustand, den wohl kein Buch der Welt jemals erreichen wird – fanden sich in der zweiten Auflage immer noch ein paar Druckfehler, die wir nun korrigiert haben.

Diese dritte Auflage enthält als Neuerung eine Formelsammlung zum Herausnehmen, in der Sie die wichtigsten Formeln des Buchs in komprimierter Form finden und die Sie für jede Gelegenheit bei sich tragen können.

Frühjahr 2011

Guido Walz
Frank Zeilfelder
Thomas Rießinger

1 Elementare Rechenmethoden

Bevor man in die höheren Sphären der Mathematik aufsteigt, sollte man zunächst sicher sein, dass der Umgang mit den elementaren Rechenverfahren kein Problem darstellt. In diesem ersten Kapitel will ich daher zunächst einmal die Grundrechenarten, insbesondere das Bruchrechnen, nochmals in Erinnerung rufen sowie den Umgang mit Termen, Klammern und sonstigen mathematischen Symbolen und Notationen üben.

1.1 Grundrechenarten

Ob Sie es glauben oder nicht, ich beginne mit dem ganz einfachen Zahlenrechnen – wobei zunächst klar gemacht werden muss, von was für „Zahlen" hier überhaupt gesprochen wird.

Die einfachsten Zahlen sind sicherlich die Zahlen $1, 2, 3, \ldots$, denen man bereits als kleines Kind begegnet, wenn man seine Bauklötzchen oder die Kekse, die vor sich auf dem Tisch liegen, zählt. (Ich persönlich bin hier nie sehr weit gekommen, da ich die Kekse lieber gegessen als gezählt habe, aber das gehört nicht hierher.) Man nennt diese Zahlen **natürliche Zahlen** und bezeichnet die Menge aller dieser Zahlen mit dem Symbol $\mathbb{N}$. Eine **Menge** ist in der Mathematik immer eine Zusammenfassung von Objekten, man sagt auch **Elementen**, und die Elemente, die eine Menge ausmachen, schreibt man in geschweiften Klammern $\{$ und $\}$, man sagt auch **Mengenklammern**, auf.

Mit dieser Notationsregelung ist also

$$\mathbb{N} \;=\; \{1, 2, 3, \ldots\}\;. \tag{1.1}$$

Anhand dieser einfachsten Zahlen können wir die ersten Rechengesetze schon einführen, deren Übertragung auf allgemeinere Zahlbereiche später kein Problem ist.

Es ist übrigens innerhalb der Mathematiker-Gemeinde umstritten, ob zur Menge $\mathbb{N}$ noch die Null gehört oder nicht. Ich will mich an dieser Stelle nicht über diese Diskussionen äußern, sondern bitte Sie einfach, die in Zeile (1.1) gegebene Definition zu übernehmen. Für die Menge der natürlichen Zahl unter Einschluss der Null werde ich falls nötig das Symbol $\mathbb{N}_0$ verwenden, also $\mathbb{N}_0 \;=\; \{0, 1, 2, 3, \ldots\}$.

Mit der Definition von Summe oder auch Produkt zweier natürlicher Zahlen will ich Ihnen jetzt wirklich nicht kommen, sonst würden Sie dieses Buch gelangweilt zur

G. Walz et al., *Brückenkurs Mathematik*

Seite legen, und das will ich nicht. Stattdessen schauen wir uns einfache Eigenschaften dieser Rechenoperationen an.

Wie Sie möglicherweise wissen, ist es völlig unerheblich, ob Sie (an einem feucht-fröhlichen Abend) zuerst zwei und bald darauf noch mal vier Gläser Rotwein trinken oder ob Sie gleich am Anfang vier und später noch mal zwei Gläser niedermachen: das Ergebnis, sprich die Kopfschmerzen am Folgetag, ist identisch. Das liegt schlicht und einfach daran, dass $2+4 = 4+2$ ist, und das gilt natürlich nicht nur für diese speziellen Zahlen, sondern für alle; so etwas nennt man in der Mathematik eine Gesetzmäßigkeit oder kurz ein Gesetz, und dieses hier heißt Kommutativgesetz. Dieses gilt auch für die Multiplikation, und das formuliere ich jetzt mal allgemein, wobei ich Ihnen auch gleich noch eine verwandte Gesetzmäßigkeit, das Assoziativgesetz, unterjuble:

Kommutativgesetz und Assoziativgesetz

Eine Rechenoperation $*$ auf einer Menge M heißt **kommutativ**, wenn für alle a und b aus M gilt:

$$a * b \; = \; b * a \; .$$

Man sagt in diesem Falle auch, dass die Operation $*$ das **Kommutativgesetz** erfüllt.

Eine Rechenoperation $*$ auf einer Menge M heißt **assoziativ**, wenn für alle a, b und c aus M gilt:

$$a * (b * c) \; = \; (a * b) * c \; .$$

Man sagt in diesem Falle auch, dass die Operation $*$ das **Assoziativgesetz** erfüllt.

Haben Sie eine Ahnung, warum ich diese Definitionen gerade hier hingeschrieben habe? Nun, darum:

Addition und Multiplikation sind kommutativ und assoziativ

Die Addition $+$ wie auch die Multiplikation $\cdot$ sind auf der Menge $\mathbb{N}$ sowohl kommutativ als auch assoziativ.

Womit wir die ersten beiden Rechengesetze für diese Operationen bereits unter Dach und Fach hätten. Beispielsweise ist $2+(3+4) = (2+3)+4$ $(= 9)$, und auch dem oben erwähnten Kleinkind ist klar, dass es am Ende des Tages neun Kekse gegessen hat, unabhängig davon, ob es morgens zwei und am Nachmittag dicht nacheinander drei und vier Kekse verputzt, oder ob es gleich morgens zwei und noch mal drei Kekse isst und sich nachmittags mit vieren begnügt; womit wir auch ein Beispiel für das Assoziativgesetz gesehen hätten.

Übungsaufgabe 1.1:
Zeigen Sie, dass für drei natürliche Zahlen a, b, c stets gilt:

$$a + b + c = c + b + a \ . \qquad \blacksquare$$

Auch die nächste Eigenschaft, die ich vorstellen will, nämlich die Distributivität, kann man für allgemeine Rechenoperation formulieren, aber da hier immer zwei verschiedene Operationen kombiniert werden müssen, will ich das doch konkret hinschreiben, um Sie nicht gleich am Anfang zu verwirren.

Distributivgesetz

Mit der Multiplikation $\cdot$ und der Addition $+$ erfüllt die Menge $\mathbb{N}$ das **Distributivgesetz**, d. h., es gilt für drei beliebige natürliche Zahlen a, b, c:

$$a \cdot (b + c) \;=\; a \cdot b + a \cdot c \ . \tag{1.2}$$

Die Klammern in diesem Ausdruck dienen übrigens nicht der Verschönerung, sondern bewirken, dass die Rechnung $b + c$ zuerst ausgeführt werden muss, danach erst wird mit a multipliziert. Dies ist die elementarste Regel der *Klammerrechnung,* die wir weiter unten noch genauer anschauen werden.

Als Beispiel für (1.2) berechne ich $3 \cdot (2 + 5) = 3 \cdot 7 = 21$, denn das erhält man auch, wenn man gemäß der rechten Seite von (1.2) berechnet: $3 \cdot 2 + 3 \cdot 5 = 6 + 15 = 21$.

Sie haben sich vielleicht schon gewundert, dass ich die Subtraktion, die ja gemeinhin als die zweite Grundrechenart gilt, bisher noch nicht erwähnt habe. Das liegt daran, dass die Subtraktion leider innerhalb der Menge der natürlichen Zahlen allein nicht möglich ist: Zwar ist klar, dass $5 - 2 = 3$ ist, aber was soll $2 - 5$ sein? Offenbar muss das eine Zahl sein, zu der man 3 addieren muss, um 0 zu erhalten. Diese Zahl bezeichnet man mit -3 und nennt sie gelegentlich auch die Gegenzahl von 3. Jede natürliche Zahl hat eine solche Gegenzahl, die Gesamtheit aller dieser Gegenzahlen der natürlichen Zahlen nennt man die **negativen Zahlen**, und diese bilden zusammen mit den natürlichen Zahlen und der Null die Menge der **ganzen Zahlen**, bezeichnet mit $\mathbb{Z}$:

$$\mathbb{Z} \;=\; \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\} \ .$$

Diese neu eingeführten negativen Zahlen sind übrigens keine künstliche Erfindung weltfremder Mathematiker, sondern durchaus praxisrelevant; denken Sie beispielsweise an die Außentemperaturen an einem kalten Wintertag oder Ihren Kontostand am Ende des Monats. Innerhalb des Zahlbereichs $\mathbb{Z}$ kann man nun fröhlich addieren, subtrahieren und multiplizieren, ohne aus der Menge herauszufallen.

Allerdings hat die Subtraktion, derentwegen wir ja letztendlich die ganzen Zahlen eingeführt hatten, einen dicken Nachteil: Sie erfüllt die einfachsten Rechengesetze nicht, denn sie ist weder kommutativ noch assoziativ. Das Fehlen der Kommutativität habe ich bereits beim Eingangsbeispiel gezeigt ($5 - 2 \neq 2 - 5$), dass die Subtraktion auch nicht assoziativ ist, zeigt das folgende Beispiel:

$$7 - (3 - 1) = 7 - 2 = 5 ,$$

aber

$$(7 - 3) - 1 = 4 - 1 = 3 ,$$

und 3 ist nicht gleich 5. Das ist natürlich ziemlich willkürlich ausgewählt, es gibt unzählige andere Zahlenbeispiele, die es hier auch getan hätten. Das ist aber gerade das Wesen eines wichtigen Beweisprinzips in der Mathematik, welches wir hier sozusagen nebenbei kennen gelernt haben: Um eine Annahme (in diesem Fall die Annahme, dass die Subtraktion assoziativ wäre) zu *widerlegen,* genügt *ein einziges* Gegenbeispiel.

Um mit ganzen, insbesondere negativen Zahlen rechnen zu können, muss noch geklärt werden, was ein Ausdruck wie $-(-a)$ bedeutet. Nun ja, das soll ja wohl die Gegenzahl von $-a$ sein, also diejenige, die bei Addition zu $-a$ gerade null ergibt. Da hätten wir aber bereits einen Kandidaten, nämlich a. Wir haben also gezeigt:

$$-(-a) = a . \tag{1.3}$$

Dies ist sozusagen der Prototyp aller Rechenregeln, die unter uns allen solche Assoziationen wie „Minus mal Minus ergibt Plus“ erzeugen, denn aus (1.3) folgt, dass für beliebige ganze Zahlen a und b gilt:

Produkte negativer Zahlen

$$(-a)b = a(-b) = -ab$$

und

$$(-a)(-b) = ab .$$

Alle weiteren Regeln, die wir u. a. im Abschnitt über Klammerrechnung kennen lernen werden, folgen aus diesen recht einfachen Regeln.

Übungsaufgabe 1.2:
Berechnen Sie folgende Zahlen:
a) $(3 - 6)(-2 - 3)$;
b) $-(8 - 3)(-2 + 5)$;
c) $(9 - 3 - 8)(3 - 1 - 7)(-2 + 1)$. ■

1.2 Bruchrechnung und rationale Zahlen

Es gibt natürlich noch eine vierte Grundrechenart, auf die Sie sicherlich schon sehnsüchtig gewartet haben: die Division. Zu deren Behandlung müssen wir aber die ganzen Zahlen verlassen und uns in die Welt der Brüche begeben. Nur Mut, ich bin bei Ihnen!

Beginnen wir wieder mit einem Beispiel, und da ich ähnlich wie Obelix immer ans Essen denke, auch hier wieder etwas aus diesem Bereich: Stellen Sie sich vor, Sie sitzen an einem schönen Sommerabend zu dritt beim Abendessen und vor Ihnen auf dem Grill liegen sechs saftige Steaks. Da alle gleich viel Hunger haben, werden diese sechs Steaks gerecht unter den drei Leuten aufgeteilt, man kann auch sagen „geteilt" oder „dividiert", und jeder erhält $6 : 3 = 2$ Steaks. So weit, so gut. Am nächsten Tag ist schlechtes Wetter, Grillen fällt flach und es gibt die beliebte Tiefkühlpizza. Leider haben Sie vor lauter Mathematik-Lernen vergessen, einzukaufen und es sind nur noch zwei Pizzen (diesen Plural von Pizza gibt es wirklich, ich habe im Duden nachgesehen) für drei Leute da. Das Problem ist klar: Bei gerechter Aufteilung bekommt jetzt niemand eine ganze Pizza, und das liegt eben daran, dass $2 : 3$ keine ganze Zahl ist.

Was tun? Bei der Pizza hilft ein scharfes Messer, bei der Division hilft das, was Mathematiker immer tun, wenn ein Problem zunächst keine Lösung hat: Sie definieren einfach etwas Neues. In diesem Fall erweitert man den Bereich der ganzen Zahlen, in dem wir uns die ganze Zeit bewegen konnten, zur Menge der **rationalen Zahlen** oder auch **Brüche**, symbolisiert durch das Zeichen $\mathbb{Q}$: Es ist

$$\mathbb{Q} = \left\{ \frac{p}{q} \text{ mit } p, q \in \mathbb{Z} \text{ und } q \neq 0 \right\} .$$

Das bedarf wohl einiger erklärender Worte, da ich hier erstmals in diesem Buch eine Menge nicht einfach aufgezählt, sondern definiert habe. Diese Definition bedeutet ausgeschrieben, dass in der Menge $\mathbb{Q}$ alle Ausdrücke (Brüche) der Form $\frac{p}{q}$ liegen, wobei p und q beliebige ganze Zahlen sein dürfen, q allerdings nicht null sein darf. Man nennt p den **Zähler** und q den **Nenner** des Bruchs.

Brüche, deren Zähler gleich eins ist, nennt man auch **Stammbrüche**. Umgekehrt sind natürlich Brüche, deren Nenner gleich eins ist, stets gleich dem Zähler, einer ganzen Zahl also. Daher sind die ganzen Zahlen als Teilmenge in der Menge der rationalen Zahlen enthalten.

Ein Bruch ist sozusagen eine Notation für eine „noch nicht durchgeführte" Division, denn $\frac{p}{q}$ bedeutet im Prinzip dasselbe wie $p : q$; jeder Bruch lässt sich durch Ausführung dieser Division in eine Dezimalzahl umformen, die entweder nur endlich viele Nachkommastellen hat oder deren Nachkommastellen sich periodisch wiederholen.

Beispiele sind

$$\frac{1}{4} = 0{,}25$$
$$\frac{2}{13} = 0{,}153846153846153846153846153846......$$
$$\frac{35}{7} = 5{,}0.$$

Unser obiges Pizza-Problem ist mit dieser neuen Notation ganz einfach zu lösen, denn jeder erhält jetzt $\frac{2}{3}$ Pizzen.

Die schöne Erfahrung, dass man jetzt beliebige (ganze) Zahlen dividieren kann, wird einem allerdings ein wenig verleidet durch die unschöne Tatsache, dass man sich jetzt mit der allseits beliebten **Bruchrechnung** herumschlagen muss. Dafür gibt es ein paar Regeln, und die werde ich Ihnen jetzt, ob Sie wollen oder nicht, näher bringen. Die erste betrifft das so genannte **Erweitern** eines Bruchs:

Erweitern von Brüchen

Multipliziert man Zähler und Nenner eines Bruchs mit der gleichen, von Null verschiedenen Zahl, so ändert sich sein Wert nicht; in Formeln:

$$\frac{p}{q} = \frac{p \cdot a}{q \cdot a} \text{ für alle } a \neq 0.$$

Man nennt dies genauer das Erweitern des Bruchs mit dem Faktor a. Praktisches Beispiel? Bitte sehr: Ob Sie eine halbe Pizza oder zwei Viertel Pizzen zu sich nehmen, ist sowohl ernährungstechnisch als auch mathematisch dasselbe, denn

$$\frac{1}{2} = \frac{1 \cdot 2}{2 \cdot 2} = \frac{2}{4} .$$

Die Umkehrung des Erweiterns ist das **Kürzen** eines Bruchs:

Kürzen von Brüchen

Enthalten Zähler und Nenner eines Bruchs den gleichen von Null verschiedenen Faktor, so kann man beide durch diesen Faktor dividieren, ohne den Wert des Bruchs zu verändern, in Formeln:

$$\frac{p \cdot a}{q \cdot a} = \frac{p}{q} \text{ für alle } a \neq 0.$$

Dies nützt man aus, um unnötig große Zahlen in Zähler und Nenner eines Bruchs zu vermeiden. Der gemeinsame Faktor a steht natürlich in den seltensten Fällen explizit da, vielmehr muss man ihn zunächst als gemeinsamen Faktor von Zähler und Nenner identifizieren. Beispielsweise sieht man relativ schnell, dass in dem Bruch $\frac{8}{12}$ der gemeinsame Faktor 4 enthalten ist und man berechnet

$$\frac{8}{12} = \frac{2 \cdot 4}{3 \cdot 4} = \frac{2}{3} .$$

Schwieriger wird's schon bei Dingen wie $\frac{104}{152}$. Manche Autoren raten nun, eine Zerlegung in Primfaktoren vorzunehmen, um den gemeinsamen Faktor (hier ist es 8) zu erkennen und mit einem Schlag zu erledigen. Ich persönlich halte es nicht für ehrenrührig, zunächst einmal kleine leicht erkennbare Faktoren auszukürzen und sein Glück danach evtl. noch mal zu versuchen; beachten Sie hierbei, dass bei obiger Definition kein Mensch verlangt hat, dass a der *größtmögliche* Faktor sein muss, der in beiden Termen steckt.

Im Beispiel erkennt man sofort, dass Zähler und Nenner beide gerade sind, also den Faktor 2 enthalten, und ermittelt

$$\frac{104}{152} = \frac{52}{76} .$$

Mit etwas Glück sieht man nun, dass der verbliebene Bruch sogar durch 4 kürzbar ist (wenn man dieses Glück nicht hat, muss man eben zweimal durch 2 kürzen) und findet als Endergebnis

$$\frac{104}{152} = \frac{52}{76} = \frac{13}{19} .$$

Da 13 und 19 Primzahlen sind, ist kein weiteres Kürzen möglich.

Übungsaufgabe 1.3:
Kürzen Sie die folgenden Brüche so weit wie möglich:

$$\frac{231}{22}, \quad \frac{52}{28}, \quad \frac{16}{64} .$$

■

Bisher haben wir mit Brüchen ja nicht wirklich gerechnet, sondern sie nur umgeformt. Die erste Rechenart, die ich nun vorstellen werde, hat es gleich in sich: die **Addition** (und synchron **Subtraktion**) **von Brüchen**; ich bin sicher, dass Ihnen bei der Nennung dieses Themas das eine oder andere traumatische Schulerlebnis in Erinnerung kommt. Das liegt daran, dass man Brüche leider *nicht* zähler- und nennerweise addieren darf, sondern zunächst dafür sorgen muss, dass sie denselben Nenner haben. Es gilt nämlich:

Summe zweier Brüche mit gleichem Nenner

Die Summe zweier Brüche der Form $\frac{p_1}{q}$ und $\frac{p_2}{q}$ ist

$$\frac{p_1}{q} + \frac{p_2}{q} = \frac{p_1 + p_2}{q} , \tag{1.4}$$

was sich mit der Lebenserfahrung deckt, dass ein Viertel einer Pizza plus zwei Viertel derselben Pizza eben insgesamt drei Viertel dieser Pizza darstellen.

Das geht aber wirklich *nur dann,* wenn die Ausgangsbrüche *denselben Nenner* haben. Da man selten so viel Glück haben wird, muss man im allgemeinen Fall durch Erweitern die Nenner zunächst angleichen. Eine Methode, die ohne viel Nachdenken immer klappt, ist das Erweitern mit dem Nenner des jeweils anderen Bruchs. Haben Sie das so verstanden? Ich glaube nicht, und das ist nicht *Ihr* Fehler, sondern meiner, manchmal ist eine formelmäßige Darstellung wirklich angebrachter. Also:

Will man die Summe der beiden Brüche $\frac{p_1}{q_1}$ und $\frac{p_2}{q_2}$ ermitteln, so kann man zunächst den ersten Bruch mit q_2, den zweiten mit q_1 erweitern und erhält – unter Benutzung von (1.4) – die Rechenvorschrift

$$\frac{p_1}{q_1} + \frac{p_2}{q_2} = \frac{p_1 q_2}{q_1 q_2} + \frac{p_2 q_1}{q_1 q_2} = \frac{p_1 q_2 + p_2 q_1}{q_1 q_2}$$

und ebenso natürlich die Vorschrift für die Subtraktion

$$\frac{p_1}{q_1} - \frac{p_2}{q_2} = \frac{p_1 q_2}{q_1 q_2} - \frac{p_2 q_1}{q_1 q_2} = \frac{p_1 q_2 - p_2 q_1}{q_1 q_2} \,.$$

Man berechnet hiermit beispielsweise

$$\frac{2}{3} + \frac{3}{7} = \frac{14}{21} + \frac{9}{21} = \frac{23}{21} \,.$$

Da ich in meinem Leben noch nie drei Siebtel einer Pizza gesehen habe, fällt mir hier leider kein Beispiel ein, aber ich bin sicher, dass Sie das durch Nachrechnen auch so verifizieren können.

Diese Methode, zwei Brüche zu addieren bzw. zu subtrahieren, klappt prinzipiell immer, ist aber manchmal nicht sinnvoll: Wenn Sie beispielsweise die Summe

$$\frac{1}{5\,000} + \frac{1}{10\,000}$$

berechnen wollen, müssten Sie beide Brüche auf den Nenner 50 000 000 bringen, was aber nicht sinnvoll ist, da man mit bloßem Auge erkennen kann, dass $\frac{1}{5\,000} = \frac{2}{10\,000}$ ist, die Summe also

$$\frac{1}{5\,000} + \frac{1}{10\,000} = \frac{3}{10\,000}$$

ergibt.

Allgemein kann man die Addition zweier Brüche effizienter gestalten, indem man das **kleinste gemeinsame Vielfache (kgV)** der beiden Nenner q_1 und q_2 als Ziel der Erweiterung nimmt. Das kgV ist, wie der Name schon andeutet, die kleinste aller Zahlen, die Vielfaches sowohl von q_1 als auch von q_2 ist. In obigem Beispiel ist dieses kgV gleich 10 000, denn es ist ein Vielfaches sowohl von 10 000 selbst (als 1-faches) als auch von 5 000. Ganz so einfach ist es natürlich nicht immer; im ungünstigsten Fall ist das kgV gleich dem Produkt der beiden beteiligten Zahlen; dies passiert beispielsweise, wenn beide Primzahlen sind.

Die folgende Formelzeile sollte Sie nicht allzu sehr erschrecken; ich möchte nur einmal die Summe zweier beliebiger Brüche als geschlossene Formel hinschreiben; manchmal brauche ich so etwas ganz einfach:

Summe zweier Brüche

$$\frac{p_1}{q_1} + \frac{p_2}{q_2} = \frac{p_1 \cdot \frac{kgV}{q_1} + p_2 \cdot \frac{kgV}{q_2}}{kgV}$$

Bei der Suche nach dem kgV – es gibt dafür auch raffinierte Algorithmen, aber die will ich Ihnen und mir hier nicht antun – kann man beispielsweise mit dem Produkt als ersten Kandidaten starten und dann schauen, ob es noch kleinere gemeinsame Vielfache gibt. So findet man beispielsweise heraus, dass das kgV von 6 und 10 gleich 30 ist: der erste Kandidat, 60, ist noch zu groß, die Hälfte tut's auch, und noch kleinere gemeinsame Vielfache als 30 gibt es offenbar auch nicht. Daher ist beispielsweise

$$\frac{1}{6} + \frac{3}{10} = \frac{5}{30} + \frac{9}{30} = \frac{14}{30} = \frac{7}{15},$$

wobei ich das Endergebnis noch mal gekürzt habe.

Keine Angst, das Schlimmste haben wir überstanden, denn die **Multiplikation von Brüchen** geht nun wirklich so vor sich, wie sich das jeder Schüler in seinen mathematischen Wunschträumen vorstellt: zähler- und nennerweise.

Produkt zweier Brüche

Das **Produkt zweier Brüche** wird wie folgt berechnet:

$$\frac{p_1}{q_1} \cdot \frac{p_2}{q_2} = \frac{p_1 \cdot p_2}{q_1 \cdot q_2}.$$

Ob Sie's glauben oder nicht, aber ich komme Ihnen schon wieder mit einem Pizza-Beispiel: Erinnern Sie sich noch an den verregneten Abend, wo Sie mit $\frac{2}{3}$ Pizza auskommen mussten? Nun, nehmen wir weiterhin an, dass an dieser Stelle Ihr großer Bruder auftaucht und $\frac{3}{4}$ Ihres Stückes für sich verlangt (es ist ein *sehr großer* Bruder!). Wie viel müssen Sie abtreten? Nach den gerade gelernten Regeln

$$\frac{3}{4} \cdot \frac{2}{3} = \frac{3 \cdot 2}{4 \cdot 3} = \frac{2}{4} = \frac{1}{2},$$

also eine halbe Pizza! Was Ihnen dann noch bleibt, rechne ich jetzt nicht aus, es ist zu deprimierend.

Übungsaufgabe 1.4:
Berechnen Sie die folgenden Produkte:

$$\frac{9}{4} \cdot \frac{6}{3}, \quad \frac{13}{7} \cdot \frac{21}{26}, \quad \frac{52}{76} \cdot \frac{19}{13}.$$

■

Da ja jede ganze Zahl als Bruch mit dem Nenner 1 interpretiert werden kann, ist jetzt auch klar, wie das Produkt eines Bruchs mit einer Zahl zu berechnen ist: Man multipliziert den Zähler mit der ganzen Zahl und lässt den Nenner unverändert:

Produkt eines Bruchs und einer ganzen Zahl

Für $a \in \mathbb{Z}$ und $\frac{p}{q} \in \mathbb{Q}$ ist

$$a \cdot \frac{p}{q} = \frac{a \cdot p}{q} .$$

Pizza? Na gut: Wenn Sie vier Mal ein Viertel Pizza gegessen haben, so macht das insgesamt

$$4 \cdot \frac{1}{4} = \frac{4 \cdot 1}{4} = 1 ,$$

also eine Pizza. So einfach und lebensnah kann Mathematik sein!

Die letzte Grundrechenart ist die Division, die aber auch nicht wesentlich schwieriger ist als die Multiplikation. Um die Divisionsvorschrift für Brüche formulieren zu können, brauche ich den Begriff des Kehrbruchs (manche sagen auch Kehrwerts) eines Bruchs: Der **Kehrbruch** des Bruchs $\frac{p}{q}$ ist der Bruch $\frac{q}{p}$, wobei natürlich p nicht null sein darf. Insbesondere ist der Kehrbruch einer ganzen Zahl a der Stammbruch $\frac{1}{a}$.

Man erhält also den Kehrbruch eines Bruchs, indem man Zähler und Nenner vertauscht. Jetzt können wir uns an die Division heranwagen:

Dividieren von Brüchen

Man dividiert durch einen Bruch, indem man mit seinem Kehrbruch multipliziert, in Formeln:

$$\frac{p_1}{q_1} : \frac{p_2}{q_2} = \frac{p_1}{q_1} \cdot \frac{q_2}{p_2} = \frac{p_1 \cdot q_2}{q_1 \cdot p_2} .$$

Das Ergebnis der Division nennt man den **Quotienten** der beiden Ausgangsbrüche.

Ohne lange Vorrede zwei Zahlenbeispiele; es ist

$$\frac{13}{15} : \frac{2}{7} = \frac{13 \cdot 7}{15 \cdot 2} = \frac{91}{30}$$

und

$$\frac{1}{5} : \frac{3}{5} = \frac{1 \cdot 5}{5 \cdot 3} = \frac{1}{3} .$$

Auch hier kann man wieder ein lebenspraktisches Beispiel formulieren. Da zu viel Pizza nicht gut für Ihren Körper ist, jetzt einmal etwas Gesundes: Kartoffelchips! Nehmen Sie an, Sie wollen einen Film anschauen, der 150 Minuten, also $\frac{5}{2}$ Stunden dauert. (Im täglichen Leben würde man hier wohl eher „zweieinhalb Stunden“ sagen, aber beim Rechnen sollte man „gemischte Zahlen“ wie $2\frac{1}{2}$ vermeiden, da dies allzu leicht mit dem Produkt $2 \cdot \frac{1}{2} (= 1)$ verwechselt werden kann.) Leider haben Sie nur noch eine dreiviertel Tüte Chips zur Verfügung und möchten diese möglichst gleichmäßig über die Laufzeit des Films verteilen. Die notwendige Rechnung lautet

$$\frac{3}{4} : \frac{5}{2} = \frac{3 \cdot 2}{4 \cdot 5} = \frac{3}{10} ,$$

wobei ich das Ergebnis gleich gekürzt habe. Sie sollten also pro Stunde drei Zehntel Tüten zu sich nehmen. Viel Spaß beim Abmessen!

Übungsaufgabe 1.5:
Berechnen Sie die folgenden Quotienten:

$$\frac{3}{4} : \frac{7}{8} , \quad \frac{52}{76} : \frac{13}{19} , \quad \frac{143}{11} : \frac{130}{22} .$$ ■

Bleibt nur noch zu betonen, dass alle eingangs für die natürlichen und die ganzen Zahlen vorgestellten Rechengesetze auch für die rationalen Zahlen ihre Gültigkeit behalten.

Zum Ende dieses Abschnitts will ich Sie nun noch in die Geheimnisse der Dreisatzrechnung einweihen – falls dies überhaupt notwendig ist, denn die Dreisatzrechnung ist natürlich Gegenstand der Schulmathematik. Aber ebenso natürlich haben die meisten Leute eine gewisse Scheu vor dem eigentlich ganz harmlosen Dreisatz, und falls Sie auch zu dieser Gruppe gehören, will ich Ihnen diese nun nehmen.

Am besten beginne ich mit einem Beispiel: Ein Mann kauft 5 Brote ein und muss dafür 6 Euro bezahlen. Am nächsten Tag (es handelt sich um eine große Familie) kauft er 7 Brote ein. Wie viel kosten diese 7 Brote?

„Eigentlich“ müsste man nun zunächst berechnen, was ein Brot kostet, nämlich

$$\frac{6\,\text{Euro}}{5\,\text{Brote}} = 1{,}20\,\text{Euro/Brot} ,$$

und dann den Preis für 7 Brote durch Multiplikation bestimmen:

$$7\,\text{Brote} \cdot 1{,}20\,\text{Euro/Brot} = 8{,}40\,\text{Euro} .$$

Der Dreisatz umgeht diese Rückrechnung auf eine Einheit (hier: Preis für ein Brot), indem er ausnutzt, dass das Verhältnis von 6 Euro zu 5 Broten dasselbe ist wie das vom unbekannten Preis p zu 7 Broten. In Formeln lautet dies:

$$\frac{6\,\text{Euro}}{5\,\text{Brote}} = \frac{p\ \text{Euro}}{7\,\text{Brote}} .$$

Löst man dies nun nach p auf, ergibt sich

$$p\,\text{Euro} = \frac{6\,\text{Euro}}{5\,\text{Brote}} \cdot 7\,\text{Brote} = \frac{42}{5}\,\text{Euro} = 8{,}40\,\text{Euro} .$$

Das ist das Wesen des Dreisatzes, also der Aufgabe, aus drei vorgegebenen Größen eine vierte, unbekannte, direkt zu berechnen:

Dreisatz

Beim Dreisatz wird die unbekannte Größe ins Verhältnis zu einer Gesamtheit gesetzt, und dieses Verhältnis muss gleich einem bekannten Verhältnis sein. Durch Auflösen dieser Gleichung nach der Unbekannten kann man diese dann direkt berechnen.

Nein, das müssen Sie nicht gleich verstehen. Ich mache lieber noch ein, zwei Beispiele: Eine Truppe von 15 Arbeitern benötigt zum Ausheben einer Grube 9 Stunden. Am nächsten Tag erscheinen krankheitsbedingt nur 6 Arbeiter zur gleichen Tätigkeit auf einer identischen Baustelle. Frage: Wie lange brauchen diese?

Der Dreisatz

$$\frac{x\,\text{Stunden}}{15\,\text{Arbeiter}} = \frac{9\,\text{Stunden}}{6\,\text{Arbeiter}}$$

ermöglicht es nun, das Umrechnen auf „Mannstunden“ zu umgehen und die Lösung

$$x\,\text{Stunden} = \frac{9\,\text{Stunden}}{6\,\text{Arbeiter}} \cdot 15\,\text{Arbeiter} = 22{,}5\,\text{Stunden}$$

direkt zu ermitteln.

An diesem Beispiel kann man auch sehr schön die Grenzen der mathematischen Modellbildung illustrieren: Derselbe Dreisatz ergibt nämlich, dass die gewünschte Arbeit von einer Million Bauarbeitern in deutlich weniger als einer Sekunde erledigt werden kann, was sicherlich aus verschiedenen Gründen nicht ganz der Realität entspricht.

Sie wissen ja schon, dass ich eine Vorliebe für Pizza habe, warum also nicht noch ein Beispiel aus diesem Genre? Nun denn: Zur Herstellung von drei Pizzen benötigt man 750 Gramm Teig. Wieviel Teig wird zur Vorbereitung einer Party mit 36 Pizzen benötigt? Zur Lösung benutzt man den Dreisatz:

$$\frac{x\,\text{Gramm}}{36\ \text{Pizzen}} = \frac{750\,\text{Gramm}}{3\ \text{Pizzen}} \;.$$

Auflösen nach x ergibt

$$x\,\text{Gramm} = \frac{750\,\text{Gramm}}{3\ \text{Pizzen}} \cdot 36\ \text{Pizzen} = 9\,000\,\text{Gramm}$$

Sie müssen also für Ihre Pizza-Party 9 Kilogramm Teig vorbereiten. Ihr Problem!

1.3 Klammerrechnung

Ich habe im bisherigen Verlauf bereits an einigen Stellen nicht auf den Einsatz von Klammern verzichten können, beispielsweise bei der Formulierung des Distributivgesetzes oder bei den Eigenschaften der Subtraktion. In beiden Fällen dienten die Klammern weniger der Verschönerung der Ausdrücke, sondern legten die Reihenfolge fest, in der die Rechnungen auszuführen waren. Generell gilt:

> **Klammerregel**
>
> Rechenoperationen, die durch Klammern eingeschlossen werden, sind stets zuerst auszuführen.

Das haben wir beim Distributivgesetz ja schon intuitiv richtig gemacht, denn $4\cdot(2+5)$ bedeutet eben, dass die Addition in der Klammer *vor* der Multiplikation ausgeführt werden muss; das Ergebnis ist 28. Der ohne Klammern notierte Ausdruck $4\cdot 2+5$ würde dagegen bedeuten, dass man $4 \cdot 2$ berechnen muss, und erst wenn man das geschafft hat, 5 addieren darf; das Ergebnis ist jetzt also 13. Der Grund ist, dass irgendwann vor langer Zeit einmal festgelegt wurde, dass Multiplikation und Division stets *vor* Addition und Subtraktion durchzuführen sind. („Punktrechnung vor Strichrechnung“, erinnern Sie sich?)

Hat man die Rechnung in der Klammer durchgeführt, so schreibt man die Klammer nicht mehr, sie hat sich sozusagen in Luft aufgelöst, und daher nennt man diesen Vorgang auch „Auflösen“ von Klammern; z. B. ist

$$-(3-5)\cdot(-2+4) \;=\; -(-2)\cdot 2 \;=\; 2\cdot 2 \;=\; 4\;.$$

Klammern tauchen häufig im Zusammenhang mit der Subtraktion auf, da diese nicht assoziativ ist; beispielsweise ist $11-5-3$ eben *nicht* dasselbe wie $11-(5-3)$. (Falls Sie zu müde sind, das selbst nachzurechnen: Der erste Ausdruck ergibt 3, der zweite 9).

Sie denken, dass Rechnungen wie $11-(5-3)$ im außermathematischen Leben nicht vorkommen können? Nun, stellen Sie sich vor, Sie schauen ein Fußballspiel Ihrer Lieblingsmannschaft an. Es handelt sich um ein ziemlich rüdes Match und plötzlich liegen fünf Spieler Ihrer Mannschaft verletzt am Spielfeldrand. Nach kurzer Behandlung können drei davon wieder mitspielen, zwei müssen ins Krankenhaus. Wie viele Spieler Ihres Teams stehen dann noch auf dem Platz? (Da es sich um ein *sehr* rüdes Spiel handelt, sind die Ersatzspieler bereits alle geflohen.) Nun, die Rechnung, die Sie hierfür anstellen müssen, ist genau die am Anfang des Absatzes genannte, und Sie sehen, dass noch neun Spieler dabei sind. Wenn Sie die Klammer weglassen würden, wären es nur noch drei und der Schiedsrichter würde das Spiel abbrechen.

Zurück zur Mathematik; Klammern können auch im Rudel auftreten, dann gilt:

Geschachtelte Klammern

Sind Klammern geschachtelt, so ist zuerst die innerste aufzulösen.

Zwei kleine Beispiele hierzu. Es ist

$$-(18-(2+3\cdot(4+1))) = -(18-(2+3\cdot 5)) = -(18-(2+15)) = -(18-17) = -1$$

und

$$\begin{aligned} 1-(1-(1-(1-(1-1)))) &= 1-(1-(1-(1-0))) = 1-(1-(1-1)) \\ &= 1-(1-0) = 1-1 = 0\,. \end{aligned}$$

Das war ein recht kurzer und wohl auch nicht allzu schwieriger Abschnitt. Vielleicht halten Sie ihn ja auch für überflüssig (was ein gutes Zeichen ist, denn dann war Ihnen das alles ja schon vertraut), aber wenn wir weiter unten mit Termen und Variablen rechnen, werden Sie froh sein, das Klammerrechnen hier schon einmal angeschaut zu haben.

Übungsaufgabe 1.6:
Berechnen Sie folgende Ausdrücke:
a) $(((2-3)\cdot 4)-2)\cdot(-2)$
b) $2-((1-4)\cdot(3-2)+4)$ ■

1.4 Potenzen und Wurzeln

Auch ohne dass ich das hier schon behandelt hätte, wissen Sie sicherlich noch, wie man den Inhalt eines Quadrats mit der Seitenlänge 2 herausfindet: Man berechnet einfach $2\cdot 2(=4)$. Wenn man nun noch eine dritte Dimension hinzunimmt, also einen Würfel der Kantenlänge 2 betrachtet, so berechnet sich dessen Inhalt zu $2\cdot 2\cdot 2(=8)$. Damit ist für uns einfache Menschenwesen das Ende der Fahnenstange schon ereicht, mehr als drei Dimensionen hat unser Lebensraum nicht, aber nehmen wir mal an, Captain Kirk würde auf einen vierdimensionalen Raum stoßen, wo man dann eben vierdimensionale Würfel der Kantenlänge 2 berechnen kann; deren Inhalt wäre dann $2\cdot 2\cdot 2\cdot 2(=16)$. Und wenn dann selbst Captain Kirk nicht mehr weiter weiß, kommt Luke Skywalker und findet einen zehndimensionalen Würfel! Nein, keine Sorge, ich schreibe jetzt nicht ein Produkt aus zehn Zweien hin, denn genau hierfür hat man in der Mathematik eine Abkürzung gefunden, die *Potenzschreibweise;* unser zehndimensionaler Würfel hätte dann den Inhalt $2^{10}(=1024)$, und allgemein definiert man das so:

Potenzierung mit natürlichen Zahlen

Ist a eine beliebige Zahl und n eine natürliche Zahl, so ist a^n definiert als das n-fache Produkt von a mit sich selbst, also

$$a^n = \underbrace{a \cdot a \cdots a}_{n-mal} .$$

Man nennt a die **Basis** und n den **Exponenten** von a^n und bezeichnet a^n selbst als **n-te Potenz** von a.

Weiterhin definiert man $a^0 = 1$, wobei dann allerdings a nicht null sein darf.

Beispielsweise ist

$$3^4 = 3 \cdot 3 \cdot 3 \cdot 3 = 81$$

und

$$\left(\frac{7}{2}\right)^3 = \frac{7}{2} \cdot \frac{7}{2} \cdot \frac{7}{2} = \frac{343}{8} .$$

Mit weiteren Zahlenbeispielen will ich uns hier nicht aufhalten, ich denke, das ist nicht nötig, und schließlich gibt es noch viel zu tun. Beispielsweise erste Rechengesetze formulieren:

Potenzgesetze

Für alle von null verschiedenen Zahlen a und b und alle natürlichen Zahlen m und n gelten die folgenden Gleichungen:

$$a^m \cdot a^n = a^{m+n} \qquad \text{(P1)}$$

$$a^m \cdot b^m = (a \cdot b)^m \qquad \text{(P2)}$$

$$(a^m)^n = a^{m \cdot n} \qquad \text{(P3)}$$

Einfache Zahlenbeispiele hierzu: Es ist $2^3 \cdot 2^2 = 2^5$, denn $2^3 = 8$, $2^2 = 4$, also $2^3 \cdot 2^2 = 32$, was mit 2^5 übereinstimmt. Das illustriert (P1).

Zur Erläuterung von (P2) berechne ich $3^2 \cdot 5^2$, was $9 \cdot 25 = 225$ ergibt. Nach Regel (P2) kann man ebenso gut $(3 \cdot 5)^2$, also 15^2 berechnen, und tatsächlich ist $15^2 = 225$.

Schließlich schauen wir uns an, was $(3^2)^4$ ergibt: Es ist $9^4 = 6561$, und das erhält man auch als Ergebnis von $3^{2 \cdot 4} = 3^8$, was nach Regel (P3) auch zu erwarten war.

Nachdem Sie nun das Rechnen mit Potenzen beherrschen, solange die Exponenten natürliche Zahlen sind, wagen wir uns jetzt weiter vor, indem wir als Exponenten alle ganzen Zahlen zulassen, also auch negative. Nur, was soll bespielsweise a^{-1} sein? Damit auch in diesem Fall Regel (P1) gilt, muss das eine Zahl sein, die die Eigenschaft

$$a^{-1} \cdot a = a^{-1+1} = a^0 = 1$$

hat, die also mit a multipliziert auf die 1 führt. Da gibt es aber nur einen echten Kandidaten, nämlich die Zahl $\frac{1}{a}$. Wir haben also gesehen:

$$a^{-1} = \frac{1}{a}$$

und allgemein:

Potenzen mit negativen Exponenten

Für alle ganzen Zahlen n gilt:

$$a^{-n} = \frac{1}{a^n} . \tag{P4}$$

Die oben formulierten Potenzregeln (P1) bis (P3) gelten auch für nicht positive Exponenten m und n.

Beispiel 1.1:
Ein paar Beispiele hierzu; nach der ersten Regel kann man in dem Term $5^4 \cdot 5^{-5} \cdot 5^3$ einfach die Exponenten addieren und erhält

$$5^4 \cdot 5^{-5} \cdot 5^3 = 5^2 = 25 .$$

Auch ein Ausdruck der Form $(x^{-2} \cdot y^{-3})^{-3} \cdot x$ muss uns nicht schrecken, denn mithilfe der Potenzregeln findet man heraus, dass

$$(x^{-2} \cdot y^{-3})^{-3} \cdot x = (x^{(-2)\cdot(-3)} \cdot y^{(-3)\cdot(-3)}) \cdot x = x^6 \cdot y^9 \cdot x = x^7 \cdot y^9$$

ist. Schließlich kann man den unangenehm aussehenden Bruch

$$\frac{(17^2)^{-6}}{17^9 \cdot 17^{-3}}$$

umformen zu

$$\frac{(17^2)^{-6}}{17^9 \cdot 17^{-3}} = \frac{17^{-12}}{17^6} = \frac{1}{17^{12} \cdot 17^6} = \frac{1}{17^{18}} .$$

Auch kein reines Vergnügen, aber immerhin kompakt zusammengefasst. ■

Übungsaufgabe 1.7:
Berechnen Sie folgende Terme:

a) $\left(\frac{1}{2}\right)^3 \cdot \left(\frac{2}{3}\right)^5 \cdot \left(\frac{3}{4}\right)^2$

b) $(a^2 b^3 c^{-1})^2 \cdot (c^2)^2$ ■

So weit, so gut. Aber nachdem das mit den ganzzahligen Exponenten so gut läuft, werden wir jetzt tollkühn und versuchen, rationale Zahlen in den Exponenten zu nehmen. Aber was um alles in der Welt soll beispielsweise

$$17^{\frac{173}{229}}$$

sein? Na ja, ganz so kompliziert muss man ja nicht anfangen, beginnen wir lieber mit den einfachsten rationalen Zahlen, den Stammbrüchen, und hier auch wieder mit dem einfachsten (wenn ich den trivialen Stammbruch $\frac{1}{1}$ weglasse), nämlich $\frac{1}{2}$.

Bevor ich weitermache: Mir ist gerade die bei Mathematik-Dozenten und -Autoren beliebte Vokabel „trivial" aus der Feder bzw. der Tastatur geflossen; die sollte ich vielleicht erklären. Ein *trivialer Fall* ist in der Mathematik einer, der so einfach bzw. offensichtlich richtig ist, dass man auf ihn nicht weiter eingehen und schon gar nichts beweisen muss. (Böse Zungen behaupten allerdings, dass ein Autor auch gerne mal „trivial" schreibt, wenn er selbst keine Ahnung hat, wie die Sache anzugehen wäre, und lieber den Leser mit dem Problem alleine lässt.)

Zurück zum Thema. Was ist nun also $a^{\frac{1}{2}}$, beispielsweise konkret $3^{\frac{1}{2}}$. Nun, nach Regel (P3) muss das eine Zahl sein, für die gilt:

$$(3^{\frac{1}{2}})^2 \;=\; 3^{\frac{1}{2}\cdot 2} \;=\; 3\,,$$

die also, wenn man sie quadriert, 3 ergibt. So eine Zahl gibt es aber in der gesamten Menge der rationalen Zahlen nicht (was man streng genommen beweisen müsste, aber das hat ein gewisser Euklid von Alexandria netterweise schon vor ein paar tausend Jahren für uns erledigt), und wenn es etwas nicht gibt, dann definiert man es eben, das kennen Sie ja schon. Und da das mit der speziellen Zahl 3 offenbar nicht das Geringste zu tun hat, macht man das gleich allgemein, und zwar wie folgt:

Potenzen mit dem Exponenten $\frac{1}{2}$ – Quadratwurzeln

Für jede positive Zahl a ist $a^{\frac{1}{2}}$ definiert als diejenige positive Zahl, die mit sich selbst multipliziert a ergibt. Man bezeichent $a^{\frac{1}{2}}$ als **Quadratwurzel** oder einfach als **Wurzel** aus a und schreibt dafür auch das Symbol $\sqrt{a}$, also

$$a^{\frac{1}{2}} \;=\; \sqrt{a}\;.$$

Bei einigen ganzen Zahlen, die man auch **Quadratzahlen** nennt, ist die Wurzel ebenfalls eine ganze Zahl, beispielsweise ist $\sqrt{16} = 4$, denn $4^2 = 16$, und ebenso ist $\sqrt{36} = 6$. In den allermeisten Fällen ist die Wurzel jedoch keine ganze Zahl, noch nicht einmal eine rationale. Was es denn sonst sein soll? Bitte noch ein paar Zeilen Geduld, dann verrate ich es Ihnen. Vorher muss ich Sie nämlich leider daran erinnern, dass wir unserem Ziel, beliebige rationale Zahlen als Exponenten zuzulassen, bisher nur minimal näher gekommen sind, denn wir haben die Sache bisher nur für eine einzige rationale Zahl, die $\frac{1}{2}$, geklärt.

Aber keine Sorge, damit ist der Bann gebrochen und der Rest ist leicht. Zunächst definiert man Potenzen mit einem beliebigen Stammbruch im Exponenten auf die bei $\frac{1}{2}$ bewährte Art und Weise:

Potenzen mit einem Stammbruch im Exponenten – n-te Wurzeln

Für jede positive Zahl a und jede natürliche Zahl n ist $a^{\frac{1}{n}}$ definiert als diejenige positive Zahl, die n-mal mit sich selbst multipliziert a ergibt. Man bezeichnet $a^{\frac{1}{n}}$ als **n-te Wurzel** aus a und schreibt dafür auch das Symbol $\sqrt[n]{a}$, also

$$a^{\frac{1}{n}} = \sqrt[n]{a}$$

Auch hier gibt es einige ausgesuchte ganze Zahlen, deren n-te Wurzeln wieder ganzzahlig sind, beispielsweise ist $\sqrt[3]{64} = 4$, denn $4^3 = 64$, oder $\sqrt[4]{16} = 2$, denn $2^4 = 16$. Für $n = 3$ nennt man solche Zahlen **Kubikzahlen**, für größere Werte von n haben sie keine eigene Bezeichnung mehr.

Langweilig? Bitte sehr:

Übungsaufgabe 1.8:
Berechnen Sie die folgenden Wurzelausdrücke:
a) $\sqrt[4]{81}$
b) $\sqrt[3]{2^2 \cdot 2}$
c) $\sqrt[5]{(a^5)^2}$ ■

Wenn Sie jetzt denken: „Nun mach' mal voran, wir haben noch unendlich viele rationale Exponenten vor uns!“, dann kann ich Ihnen versichern, wir haben es so gut wie geschafft. Denn wenn Sie eine Zahl a mit einer beliebigen rationalen Zahl $\frac{p}{q}$ als Exponenten vorfinden, so können Sie mithilfe der Regel (P3) folgende Umformung vornehmen:

$$a^{\frac{p}{q}} = a^{\frac{1}{q} \cdot p} = (a^{\frac{1}{q}})^p \ . \tag{1.5}$$

Damit ist aber alles klar, denn das Potenzieren mit dem Stammbruch $\frac{1}{q}$ haben wir gerade gelernt und das Potenzieren mit der *ganzen Zahl* p können wir auch. In (nicht sehr schönen, aber vielleicht notwendigen) Worten heißt das wie folgt:

Potenzieren mit einer beliebigen rationalen Zahl

Für eine beliebige positive Zahl a bezeichnet der Ausdruck $a^{\frac{p}{q}}$ die p-te Potenz der q-ten Wurzel aus a, in Formeln:

$$a^{\frac{p}{q}} = (a^{\frac{1}{q}})^p = (\sqrt[q]{a})^p \ .$$

Und es kommt noch besser: Wegen des Kommutativgesetzes der Multiplikation (erinnern Sie sich?) kann man anstelle der Umformung in (1.5) auch die folgende machen:

$$a^{\frac{p}{q}} = a^{p \cdot \frac{1}{q}} = (a^p)^{\frac{1}{q}} . \tag{1.6}$$

Das bedeutet, dass $a^{\frac{p}{q}}$ auch gleich der q-ten Wurzel aus der p-ten Potenz von a ist. Das ist in der Zusammenfassung einen Textkasten wert:

Identitäten für $\mathbf{a^{\frac{p}{q}}}$

Für alle positiven Zahlen a und alle rationalen Exponenten $\frac{p}{q}$ gilt:

$$a^{\frac{p}{q}} = (a^{\frac{1}{q}})^p = (\sqrt[q]{a})^p = (a^p)^{\frac{1}{q}} = \sqrt[q]{a^p} .$$

Beispielsweise erhalten Sie nach jeder der hier angegebenen Identitäten, dass $8^{\frac{2}{3}} = 4$ ist, denn

$$8^{\frac{2}{3}} = (8^{\frac{1}{3}})^2 = 2^2 = 4 ,$$

aber auch

$$8^{\frac{2}{3}} = (8^2)^{\frac{1}{3}} = 64^{\frac{1}{3}} = 4 ,$$

weil $4^3 = 64$ ist.

So kann man also Potenzen mit rationalem Exponenten berechnen. Die Potenzrechenregeln gelten unverändert auch für rationale Exponenten, beispielsweise ist

$$3^{\frac{1}{2}} \cdot 3^{\frac{3}{2}} = 3^{\frac{1}{2}+\frac{3}{2}} = 3^2 = 9 .$$

Das sollte man sicherlich noch ein wenig üben, daher hier noch ein paar Vorschläge zum Selbststudium:

Übungsaufgabe 1.9:
Berechnen Sie folgende Zahlen:
a) $4^{\frac{1}{3}} \cdot 2^{\frac{1}{3}} \cdot 16^{\frac{1}{4}}$
b) $120^{\frac{1}{2}} \cdot 900^{\frac{1}{4}}$
c) $\sqrt{0{,}16}$ ■

Ich hatte oben schon erwähnt, dass Wurzeln in den seltensten Fällen rationale Zahlen sind. Was aber sind sie dann? Nun ja, zunächst einmal verpasst man ihnen und ihren Leidensgenossen einen Namen: Zahlen, die nicht rational sind, nennt man **irrationale Zahlen**.

Mit Namensgebung allein ist es aber nicht getan, man will ja auch wissen, wie groß die jeweilige irrationale Zahl ist. Und das kann man, indem man ihre Dezimaldarstellung näherungsweise bestimmt. Ich zeige Ihnen das einmal am Beispiel der irrationalen Zahl $\sqrt{2}$.

Ich pirsche mich dabei schrittweise an die Dezimaldarstellung dieser Zahl heran, indem ich eine Nachkommastelle nach der anderen bestimme. Um eine erste Orientierung über die Zahl $\sqrt{2}$ zu bekommen, stelle ich fest, dass $1^2 = 1$ kleiner als 2 und $2^2 = 4$ größer als 2 ist. Damit ist klar, dass $\sqrt{2}$ zwischen 1 und 2 liegt, die Dezimaldarstellung also mit 1,irgendwas beginnt. Um „irgendwas" näher zu untersuchen, fokussiere ich nun auf die erste Nachkommastelle und berechne

$$1{,}4^2 = 1{,}96 \quad \text{und} \quad 1{,}5^2 = 2{,}25\ .$$

Das eine Quadrat ist also kleiner, das andere größer als 2, somit liegt $\sqrt{2}$ zwischen $1,4$ und $1,5$ und ist also von der Form $1,4+$irgendwas. Sie wissen schon, wie ich das neue „irgendwas" untersuche: Ich berechne

$$1{,}41^2 = 1{,}988\,1 \quad \text{und} \quad 1{,}42^2 = 2{,}016\,4\ .$$

Damit wissen wir, dass $\sqrt{2}$ zwischen 1,41 und 1,42 liegt und somit von der Form $1,41+$irgendwas ist.

Sie ahnen es schon: Es gibt unendlich viele „irgendwasse", denn die Dezimaldarstellung der irrationalen Zahl $\sqrt{2}$ bricht niemals ab. Allerdings kann man sie beliebig genau berechnen und man findet heraus:

$$\sqrt{2} \;=\; 1{,}414\,213\,562\,373\,095\,048\,801\,688\,724\,209.....$$

Sie müssen zugeben, dass damit eine „ungefähre" Einschätzung der Größe dieser Zahl gelungen ist.

So oder ähnlich kann man mit jeder irrationalen Zahl umgehen, jede davon lässt sich beliebig genau größenmäßig ermitteln und somit in die vorhandene Skala der rationalen Zahlen einsortieren. Man kann sogar zeigen, dass es zwischen zwei beliebigen rationalen Zahlen immer unendlich viele irrationale gibt; diese Zahlen liegen, wie man auch sagt, beliebig dicht beieinander. Wenn man sie als nebeneinander liegende Punkte markiert, dann ergeben diese Punkte eine geschlossene Linie, die man auch **Zahlengerade** nennt. Die Menge dieser Zahlen hat eine spezielle Bezeichnung:

Reelle Zahlen

Die Zusammenfassung aller rationalen mit allen irrationalen Zahlen bezeichnet man als die Menge der **reellen Zahlen**, bezeichnet mit dem Symbol $\mathbb{R}$.

Die Menge $\mathbb{R}$ der reellen Zahlen umfasst nun tatsächlich alle Zahlen, die Sie sich so vorstellen können, natürlich zunächst einmal die ganzen und die rationalen Zahlen, aber auch berühmte irrationale Zahlen wie die Kreiszahl π oder die eulersche Zahl e, die Ihnen später noch begegnen werden, und natürlich auch unendlich viele unspektakuläre irrationale Zahlen wie etwa $\sqrt{17}$, die meist unbeachtet ihr Dasein fristen. Wichtig ist, dass man mit allen reellen genau so rechnen kann wie ich es im bisherigen

Verlauf für natürliche, ganze und rationale Zahlen herausgearbeitet habe. Das möchte ich abschließend noch als zitierfähige Regel zusammenfassen:

Rechenregeln für reelle Zahlen

Alle Rechenregeln, die wir bisher für natürliche, ganze und rationale Zahlen herausgefunden haben, gelten unverändert auch für die reellen Zahlen.

1.5 Spezielle Ausdrücke und Notationen

In der Mathematik wimmelt es von speziellen Notationen und Formelzeichen. Bis zu einem gewissen Maß (manche Autoren übertreiben es ein wenig, wie ich finde) ist das auch richtig so, denn es unterstützt die in der Mathematik notwendige Präzision und Knappheit. Beispielsweise ist der Ausdruck

$$\sqrt{16} = 4 \tag{1.7}$$

viel kürzer als

„Diejenige positive Zahl, die – mit sich selbst multipliziert – 16 ergibt, lautet 4“.

Das Problem damit ist nur, dass viele Studierende gerade am Anfang durch dieses ungewohnte Formelchinesisch abgeschreckt werden und inhaltlich eigentlich ganz leicht zu begreifende Zusammenhänge nicht verstehen, weil sie schlicht und ergreifend die Formelzeichen nicht „übersetzen“ können; beispielsweise nützt Ihnen die Knappheit und Präzision des Ausdrucks in (1.7) gar nichts, wenn Sie nicht wissen, was das Zeichen „$\sqrt{}$“ bedeutet. Es ist wohl schon so manchem Dozenten passiert, dass er eine Tafel voll elegantester Aussagen und Zusammenhänge angeschrieben hat, ohne zu merken, dass die Hälfte seiner Zuhörer bereits am Anfang beim ersten Summensymbol ausgestiegen ist.

Um das in Ihrem Fall zu verhindern, will ich im Folgenden ein paar in der Mathematik gebräuchliche Notationen und Konventionen erklären.

Obwohl ich im bisherigen Verlauf dieses ersten Kapitels versucht habe, möglichst konkret mit Zahlen zu arbeiten, so musste ich doch an der einen oder anderen Stelle „mit Buchstaben rechnen“, wie mein damals kleiner Sohn einmal gesagt hat. Nun, diese „Buchstabenrechnung“ ist Grundlage der gesamten Mathematik, ohne sie wären Formeln, Lösungsverfahren, Funktionen, Gleichungssysteme und andere schöne Dinge nicht formulierbar. Buchstaben, mit denen man rechnet, nennt man **Variablen**, je nach Situation auch **Unbekannte** oder **Parameter**. Sie dienen als Platzhalter in einem mathematischen Ausdruck und werden später durch Elemente einer vorgegebenen Grundmenge, meist Zahlen, ersetzt. Einen mathematischen Ausdruck, der Variablen enthält, nennt man einen **Term**. Das hat übrigens nichts mit Thermik oder Thermalbad zu tun, weshalb man es auch ohne „h“ schreibt.

Diese Begrifflichkeiten sollten Sie nicht erschrecken, so etwas kennen Sie schon lange; beispielsweise ist

$$2 \cdot (a + b)$$

ein einfacher Term. Nennen Sie die beiden Seitenlängen eines Rechtecks a und b, so gibt dieser Term den Umfang des Rechtecks an. Der (hier noch sehr kleine) Vorteil ist, dass man damit den Umfang jedes beliebigen Rechtecks auf der Welt (übrigens auch einer rechteckigen Pizza) berechnen kann, ohne jedes Mal die vier Seitenlängen addieren zu müssen.

Wir werden später, im Abschnitt über Funktionen, den Umgang mit Termen, die Wurzeln, Potenzen oder ähnlich scheußliche Dinge enthalten, ausführlich üben, so dass ich uns hier nicht damit aufhalten will. Vielmehr will ich Ihnen anhand der Klammerrechnung, die ich in Abschnitt 1.3 für Zahlen eingeführt hatte, nun den Umgang mit Termen näher bringen.

Im Folgenden seien a, b, c und x, y, z Variablen, die eine beliebige reelle Zahl darstellen; später wird man hierfür einfach sagen: Seien a, b, c, d und x, y, z beliebige reelle Zahlen.

Mit dem einfachen Auflösen von Klammern ist es nun nicht mehr so leicht, denn in einem Term wie $3a \cdot (b+2c)$ kann ich ja nicht $b+2c$ verkürzen, das muss schon so stehen bleiben. Allerdings kann man derartige Terme durch Ausmultiplizieren vereinfachen, denn das **Distributivgesetz** gilt auch für Terme. Somit ist

$$3a \cdot (b + 2c) \;=\; 3ab + 6ac \;.$$

Natürlich können auch zwei oder mehrere geklammerte Terme multipliziert werden; für zwei sieht das so aus:

Ausmultiplizieren von Klammern

Für alle reellen Zahlen a, b, x und y ist:

$$(a + b) \cdot (x + y) \;=\; ax + bx + ay + by \;. \tag{1.8}$$

Zur Begründung muss man nur zweimal nacheinander das Distributivgesetz anwenden:

$$\begin{aligned}(a + b) \cdot (x + y) \;&=\; (a + b) \cdot x + (a + b) \cdot y \\ &=\; ax + bx + ay + by \;.\end{aligned}$$

Natürlich muss man auch bei der Anwendung von (1.8) die Vorzeichenregeln beachten, beispielsweise ist

$$(a - b) \cdot (x + y) \;=\; ax - bx + ay - by \;.$$

Beispiel 1.2:
Zwei leichte Beispiele dazu: Es ist

$$(3x - 5z) \cdot (-b + 3c) \;=\; -3bx + 5bz + 9cx - 15cz$$

und

$$-(xy+c)\cdot(ab-2xz) \;=\; -(abxy+abc-2x^2yz-2cxz) \;=\; -abxy-abc+2x^2yz+2cxz\,.$$

■

Hier habe ich von der Konvention Gebrauch gemacht, dass man bei einem Produkt von Variablen die einzelnen Faktoren in der alphabetischen Reihenfolge notiert. Das ist natürlich (wegen der Kommutativität der Multiplikation) mathematisch dasselbe wie jede andere Reihenfolge auch, aber hilft einem bei langen Ausdrücken den Überblick zu bewahren; oder hätten Sie auf Anhieb gesehen, dass

$$abcxyz - zaxcby + yxbazc - byaxcz \;=\; 0$$

ist? Wohl kaum, ich jedenfalls nicht. Wenn Sie aber die vier Produkte jeweils in die alphabetisch richtige Anordnung bringen, sehen Sie sofort, dass da vier Mal $abcxyz$ steht.

Spezialfälle der Regel (1.8) sind die **binomischen Formeln**, eine echte Krone der Schulmathematik:

> **Binomische Formeln**
>
> Für alle reellen Zahlen a und b gilt:
>
> $$\begin{aligned}(a+b)^2 &= a^2 + 2ab + b^2\\ (a-b)^2 &= a^2 - 2ab + b^2\\ (a+b)\cdot(a-b) &= a^2 - b^2\end{aligned}$$

Diese sind übrigens nicht, wie ich einmal in einem Lehrbuch gelesen habe, nach dem italienischen Mathematiker Allesandro Binomi benannt (ganz einfach weil es den nicht gibt), vielmehr kommt die Bezeichnung aus dem Griechischen und sagt, dass hier „zwei Dinge“ (nämlich a und b) behandelt werden. Die binomischen Formeln, die in der oben angegebenen Reihenfolge oft auch als erste, zweite und dritte binomische Formel angesprochen werden, beweist man einfach durch Anwendung von (1.8) und anschließendes Zusammenfassen. So erhält man beispielsweise

$$(a+b)\cdot(a-b) \;=\; a^2 - ab + ab - b^2 \;=\; a^2 - b^2\,,$$

was die dritte Formel beweist.

Das Zusammenfassen gleichnamiger Terme habe ich Ihnen hier quasi nebenbei untergejubelt, es sollte natürlich immer im Anschluss an das Ausmultiplizieren durchgeführt werden, um allzu lange Ausdrücke zu vermeiden. Halten wir das einmal fest:

> **Zusammenfassen gleichnamiger Terme**
>
> Nach dem Ausmultiplizieren von Klammern sollte man stets gleichnamige Terme (also solche, die dieselben Variablen enthalten) zusammenfassen.

Nachdem Ihnen inzwischen wahrscheinlich schon allein beim Gedanken an Pizza übel wird und die Kartoffelchips leider alle sind, komme ich Ihnen jetzt allen Ernstes einmal mit Äpfeln und Birnen, dem Klassiker des elementaren Rechnens schlechthin: Wenn Sie als Ergebnis einer Multiplikation, die ich mir so genau auch nicht vorstellen kann, erhalten:

$$5 \text{ Äpfel} + 6 \text{ Birnen} - 3 \text{ Äpfel} - 2 \text{ Birnen}\,,$$

so werden Sie dies sicherlich nicht so stehen lassen, sondern zusammenfassen zu

$$2 \text{ Äpfel} + 4 \text{ Birnen}\ .$$

Genau das ist mit dem oben erwähnten „Zusammenfassen gleichnamiger Terme“ gemeint, wobei diese Terme in der Mathematik nicht „Äpfel und Birnen“ heißen, sondern ab, x^2y o. ä. Ein in diesem Sinne eher mathematisch angehauchtes Beispiel, bei dem auch gleich das Ausmultiplizieren nach (1.8) noch mal geübt wird, ist das folgende:

$$\begin{aligned}(2x+3)\cdot(4xy+y) &= 8x^2y+12xy+2xy+3y\\ &= 8x^2y+14xy+3y\ .\end{aligned}$$

Natürlich können in solchen Ausdrücken auch mehr als zwei Klammern oder auch mehr als zwei Ausdrücke in einer Klammer auftreten. Das konsequente mehrfache Anwenden der gerade gelernten Regeln führt auch hier zum Ziel, beispielsweise ist

$$\begin{aligned}(a+1)\cdot(4a+3)\cdot(1-a) &= (a+1)\cdot(4a+3-4a^2-3a)\\ &= (a+1)\cdot(a+3-4a^2)\\ &= a^2+a+3a+3-4a^3-4a^2\\ &= -4a^3-3a^2+4a+3\end{aligned}$$

oder

$$\begin{aligned}(1+x+y-xy)\cdot(x-y) &= x-y+x^2-xy+xy-y^2-x^2y+xy^2\\ &= x-y+x^2-y^2-x^2y+xy^2\ .\end{aligned}$$

Weitere Vereinfachungen kann ich nicht erkennen, daher lasse ich das so stehen.

Die Kombination dieser beiden Verallgemeinerungen, also sowohl mehr als zwei Klammern als auch mehr als zwei Terme in den Klammern, ist mir zu kompliziert; zum Glück ist mir gerade aufgefallen, dass ich schon lange keine Übungsaufgabe mehr gestellt habe, das käme doch jetzt gerade recht, finden Sie nicht auch?

Übungsaufgabe 1.10:
Vereinfachen Sie folgende Terme so weit wie möglich:
a) $(a^2 + b - c)(a + bc)(a^2 b)$
b) $(1 + x + x^2 + x^3 + x^4)(1 - x)$ ■

Ebenso wie beim einfachen Zahlenrechnen können natürlich auch beim Umgang mit Termen **geschachtelte Klammern** auftreten; daran ist jetzt überhaupt nichts Geheimnisvolles mehr, auch hier gilt die Regel, dass in diesem Fall immer die innerste Klammer zuerst aufgelöst werden muss.

Eine formelmäßige Rechenvorschrift hierfür aufzustellen, ist wenig sinnvoll, ich denke, ein oder zwei Beispiele verdeutlichen die Situation besser.

Beispiel 1.3:
Als Erstes betrachte ich den Term

$$T = 3b - (2b \cdot (3 - (2a + 3b) \cdot (3b \cdot (1 - a)))) .$$

Getreu dem Motto „von innen nach außen“ wird zunächst die innerste Klammer aufgelöst, das liefert

$$T = 3b - (2b \cdot (3 - (2a + 3b) \cdot (3b - 3ab))) .$$

Nun werden die beiden inneren Klammern ausmultipliziert, wonach T die folgende Gestalt hat:

$$T = 3b - (2b \cdot (3 - (6ab + 9b^2 - 6a^2 b - 9ab^2))) .$$

Das Auflösen der jetzt innersten Klammer bedeutet nun einfach, das Minuszeichen davor hineinzuziehen; das ergibt

$$T = 3b - (2b \cdot (3 - 6ab - 9b^2 + 6a^2 b + 9ab^2)) .$$

Nach der Multiplikation der nunmehr innersten Klammer haben wir

$$T = 3b - (6b - 12ab^2 - 18b^3 + 12a^2 b^2 + 18ab^3) ,$$

und schließlich als Endergebnis

$$T = -3b + 12ab^2 + 18b^3 - 12a^2 b^2 - 18ab^3 .$$ ■

Im Vorwort wurde gesagt, dass dieses Buch Lust auf das Lesen von Mathematik-Texten machen soll, und ich bin nicht so *ganz* sicher, ob mir das mit diesem Beispiel gelungen ist; ich glaube eigentlich eher nicht. Daher verzichte ich jetzt auf das ursprünglich hier geplante zweite Beispiel und gehe zum nächsten Thema über.

In gewissem Sinn behandle ich jetzt die Umkehrung des Ausmultiplizierens von Klammern: das Ausklammern eines gemeinsamen Faktors der einzelnen Summanden einer Summe. Noch Fragen, Watson? Ich denke schon, denn das war nicht sehr klar. Ausnahmsweise ist hier vielleicht einmal eine formelmäßige Darstellung klarer als eine verbale:

Ausklammern eines Faktors

Die Umformung einer Summe der Form $ax + ay$ in die Form $a \cdot (x + y)$ nennt man **Ausklammern** des Faktors a aus der Summe.

Bei diesem Thema ist das Arbeiten mit Termen ausnahmsweise einmal leichter als das mit Zahlen, denn bei Termen kann man die gemeinsamen Faktoren der einzelnen Summanden oft mit bloßem Auge erkennen. Ein Beispiel: Die Summanden in der Summe

$$abx^2y + axy^2$$

enthalten offensichtlich beide den Faktor axy (und keinen größeren), somit kann man ausklammern wie folgt:

$$abx^2y + axy^2 \;=\; axy \cdot (bx + y)\ .$$

Der Knackpunkt beim effizienten Ausklammern ist offenbar das Auffinden des **größten gemeinsamen Teilers** der Summanden, den man meist mit ggT bezeichnet. Diese Bezeichnung ist eigentlich selbst erklärend: Der ggT ist der größte Term, der alle beteiligten Summanden teilt, also der größtmögliche Faktor, der herausgezogen werden kann. Beim Auffinden dieses Faktors geht man, speziell bei Summen mit mehr als zwei Summanden, pragmatischerweise wie folgt vor: Man nimmt einen beliebigen – meist den ersten – Summanden als Kandidaten für den *ggT* und stutzt ihn durch Überprüfung der anderen Summanden nach und nach zurück, bis am Ende der größte *gemeinsame* Faktor übrig bleibt.

Beispiel? Na klar, Sie haben Recht:

Beispiel 1.4:
Betrachen wir den Term

$$6abx^2 - 6a^2x + 3ax + 12ab\ .$$

Da ich um diese Zeit nicht mehr allzu kreativ bin, nehme ich als ersten Kandidaten für den größten gemeinsamen Faktor einfach den ersten Summanden, $6abx^2$. Schon beim Testen des zweiten Summanden, $-6a^2x$, muss der Kandidat Federn lassen, denn dieser zweite Summand enthält nicht b als Faktor, und x tritt auch nur in der ersten Potenz auf, also verbleibt $6ax$ als Kandidat. In den dritten Summanden, $3ax$, passt das fast hinein, lediglich der Vorfaktor wird zur 3, der Kandidat also zu $3ax$. Im letzten Summanden ist leider kein x enthalten, sodass als ggT der Faktor $3a$ identifiziert ist. Somit gilt:

$$6abx^2 - 6a^2x + 3ax + 12ab \;=\; 3a \cdot (2bx^2 - 2ax + x + 4b)\ . \tag{1.9}$$

■

Übrigens wäre es wohl schneller gegangen, wenn ich die Testerei mit dem dritten Summanden $3ax$ begonnen hätte, aber man trifft im Leben eben nicht nur glückliche Entscheidungen.

Übungsaufgabe 1.11:
Klammern Sie den größtmöglichen Faktor aus:
a) $34xyz^2 - 17x^3yz + 51y^4$
b) $2a^2bc - 4ab^2c + 8abc^2$ ■

Eine wichtige Anwendung des Ausklammerns gemeinsamer Faktoren ist das **Kürzen von Brüchen**. Hierzu muss man nämlich – das ist bei Termen nicht anders als es oben bei Zahlen war – im Zähler und im Nenner einen gemeinsamen Faktor identifizieren, und falls Zähler und/oder Nenner aus einer Summe von Termen bestehen, bedeutet das eben Ausklammern. Ein erstes einfaches Beispiel ist der Bruch

$$\frac{4ax - 2ay}{ay} \ .$$

Hier sieht man mit bloßem Auge, dass im Zähler der Faktor $2a$ ausklammerbar ist, im Nenner ist nichts zu tun. Damit können wir umformen:

$$\frac{4ax - 2ay}{ay} = \frac{2a \cdot (2x - y)}{ay}$$

und (durch a) kürzen:

$$\frac{2a \cdot (2x - y)}{ay} = \frac{2 \cdot (2x - y)}{y} \ .$$

Mehr geht hier nicht, so verlockend der Faktor 2 im Zähler oder gar das y in der Klammer auch aussehen mag. Apropos „verlockend“: Ich habe lange überlegt, aber mir ist zu diesem Thema einfach keine Pizza-Anwendung eingefallen; wenn Sie eine gute Idee haben, schreiben Sie mir bitte, Sie werden dann in der nächsten Auflage des Werkes auch lobend erwähnt, versprochen!

Irgendwie scheint in mir der trockene Mathematiker doch tiefer zu stecken als ich dachte, denn im Gegensatz zu Pizza-Beispielen fallen mir gerade massenhaft formelmäßige Beispiele ein; eines davon tun wir uns noch an, in Ordnung? Danke:

Beispiel 1.5:
Dann schauen wir also den folgenden Bruch einmal an:

$$\frac{6abx^2 - 6a^2x + 3ax + 12ab}{6axy^2 - 12ax^2 + 6a^2x} \ .$$

Sie haben's so gewollt! Aber ganz so schlimm ist es nicht, denn wir haben oben schon Vorarbeit geleistet und in Formel (1.9) den Zähler bereits komplett erledigt. Betrachtet man den Nenner, also den Term $6axy^2 - 12ax^2 + 6a^2x$, so sticht ins Auge, dass alle drei Summanden den Faktor $6ax$ enthalten, und eine genauere Inspektion ergibt, dass das auch schon der größte gemeinsame Faktor ist. Damit ist

$$6axy^2 - 12ax^2 + 6a^2x = 6ax \cdot (y^2 - 2x + a) \ ,$$

und der gesamte Bruch wird zu

$$\frac{6abx^2 - 6a^2x + 3ax + 12ab}{6axy^2 - 12ax^2 + 6a^2x} = \frac{3a \cdot (2bx^2 - 2ax + x + 4b)}{6ax \cdot (y^2 - 2x + a)},$$

wobei ich das Ergebnis in (1.9) verwendet habe. Das Kürzen (durch $3a$) ist jetzt ein reines Vergnügen und liefert das Ergebnis

$$\frac{6abx^2 - 6a^2x + 3ax + 12ab}{6axy^2 - 12ax^2 + 6a^2x} = \frac{3a \cdot (2bx^2 - 2ax + x + 4b)}{6ax \cdot (y^2 - 2x + a)} = \frac{2bx^2 - 2ax + x + 4b}{2x \cdot (y^2 - 2x + a)}.$$

■

Mehr kann man hier nicht tun, aber da Untätigkeit in der Mathematik wie auch im Leben ganz schlecht ist, habe ich hier was für Sie:

Übungsaufgabe 1.12:
Kürzen Sie so weit wie möglich:

a) $\dfrac{4a^3(bc)^2}{(2abc - 4ab)a^2}$

b) $\dfrac{(3x^2yz)^3}{9(xy^2z)^3}$ ■

Damit sind wir mit dem Thema Klammerrechnung so gut wie durch, aber eine spezielle Sache aus der Dirty-Tricks-Abteilung will ich Ihnen noch zeigen: Wenn Sie die dritte binomische Formel rückwärts lesen, dann steht da

$$a^2 - b^2 = (a + b) \cdot (a - b),$$

und das wird nun sehr häufig benutzt, um Ausdrücke der Form $a^2 - b^2$ „auseinander zu ziehen". Beispielsweise ist für alle a der Welt

$$a^4 - a^2 = (a^2 + a) \cdot (a^2 - a) \tag{1.10}$$

und vor allem die Beziehung

$$1 - x^2 = (1 + x) \cdot (1 - x)$$

wird immer wieder gerne genommen. (Wenn Sie bei der Umformung in (1.10) ins Grübeln gekommen sind, schauen Sie noch mal ganz kurz in den Abschnitt über Potenzrechnung hinein, ich warte hier so lange.) Ich bin fast sicher, dass Ihnen so etwas in einer Ihrer ersten Vorlesungen oder auch Übungs- oder Klausuraufgaben begegnen wird und dann werden Sie mir dankbar sein!

Bei den bisherigen Rechnungen hatten wir es eigentlich immer nur mit einer Hand voll Zahlen oder Variablen zu tun, die man alle explizit hinschreiben konnte. Das wird nicht immer so sein; will man beispielsweise die Summe von, sagen wir, 16 Zahlen $a_1, a_2, a_3, \ldots, a_{16}$ berechnen, so wird man die schon nicht mehr alle hinschreiben wollen; man kann dann natürlich die Pünktchenschreibweise

$$a_1 + a_2 + \cdots\cdots + a_{16} \tag{1.11}$$

benutzen. Das ist aber nicht nur sehr platzaufwendig, sondern manchmal auch nicht exakt genug, denn der Ausdruck in (1.11) kann ja auch die Summe $a_1+a_2+a_4+a_8+a_{16}$ bedeuten, das Bildungsgesetz der zu summierenden Indizes ist nicht eindeutig. Daher hat man das Summenzeichen $\sum$ erfunden. Dies ist ein Sigma, also ein griechisches „S", und das steht für „Summe". Was summiert werden soll, schreibt man hinter das Summenzeichen, in unserem Beispielfall wäre das $\sum a_i$, der Index i, über den summiert werden soll, wurde hier also als Variable gesetzt. Der Bereich, den dieser Index durchlaufen soll, wird ober- und unterhalb des Summenzeichens geschrieben, unten der erste Wert, oben der letzte Wert; die Summe der o. g. 16 Zahlen lautet somit in exakter Schreibweise:

$$\sum_{i=1}^{16} a_i \, ,$$

und wenn man das einmal aussprechen muss, so sagt man: „Summe über a_i für i von 1 bis 16." Beachten Sie, dass hier keine Unsicherheit über den Laufbereich des Indexes mehr besteht, denn nach Übereinkunft durchläuft i hier *alle* ganzen Zahlen zwischen der unteren und oberen Grenze. Will man nicht alle Zahlen in der Summe haben, so muss man ein wenig tricksen; beispielsweise ist

$$a_1 + a_2 + a_4 + a_8 + a_{16} \;=\; \sum_{i=0}^{4} a_{2^i} \; .$$

(Prüfen Sie das lieber mal nach, bei mir weiß man nie!) Natürlich kann die obere und/oder die untere Grenze selbst eine variable Größe sein und der Index muss nicht i heißen (gebräuchlich, aber nicht zwingend vorgeschrieben sind hier die Buchstaben i, j, k, m, n, manchmal auch die griechischen Buchstaben μ (mü) und ν (nü)).

Einige Beispiele dazu: Die Summe der ersten 100 geraden Zahlen ist einfach

$$\sum_{i=1}^{100} 2i \, ,$$

die Summe der Stammbrüche zwischen $\frac{1}{5}$ und $\frac{1}{n}$, wobei n eine beliebige natürliche Zahl größer oder gleich 5 bezeichnet, lautet

$$\sum_{j=5}^{n} \frac{1}{j} \; .$$

Schließlich ist

$$\sum_{k=-3}^{2} 2^{k+1} \;=\; 2^{-2}+2^{-1}+2^{0}+2^{1}+2^{2}+2^{3} \;=\; \frac{1}{4}+\frac{1}{2}+1+2+4+8\,,$$

wofür auch immer das gut sein soll. Bevor ich Sie mit ein paar Übungsaufgaben alleine lasse, noch eine Bemerkung: Gerade bei variablen Summengrenzen kann es passieren, dass der Startwert des Indexes (der unter der Summe steht) schon größer ist als der Endwert oben. Da hier offenbar nichts zu summieren ist, vereinbart man, dass der Summenwert in diesem Fall null ist:

Leere Summen

$$\sum_{i=n}^{m} a_i \;=\; 0\,,\ \text{falls } n > m\ .$$

Jetzt aber Sie:

Übungsaufgabe 1.13:
Stellen Sie mithilfe des Summenzeichens dar:
a) die Summe aller durch 3 teilbaren Zahlen unter 100;
b) die Summe der ersten 11 ungeraden Zahlen. ■

Summen können auch geschachtelt auftreten, beispielsweise bedeutet

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{i,j}$$

die Summe aller Zahlen $a_{i,j}$ (welche natürlich gegeben sein müssen), wobei der erste Index zwischen 1 und m, der zweite zwischen 1 und n läuft. Diese Summe hat also insgesamt $m \cdot n$ Summanden.

Wenn es ganz dicke kommt, dann hängt die Summationsgrenze der inneren Summe noch vom Laufindex der äußeren ab; betrachten Sie doch mal dieses Sahnestückchen:

$$\sum_{i=1}^{3}\sum_{j=-i}^{i-2} a_{i,j}\ .$$

Ausgeschrieben lautet es:

$$\begin{aligned}\sum_{i=1}^{3}\sum_{j=-i}^{i-2} a_{i,j} \;&=\; \sum_{j=-1}^{-1} a_{1,j} \;+\; \sum_{j=-2}^{0} a_{2,j} \;+\; \sum_{j=-3}^{1} a_{3,j}\\ &=\; a_{1,-1} \;+\; (a_{2,-2}+a_{2,-1}+a_{2,0}) \;+\; (a_{3,-3}+a_{3,-2}+a_{3,-1}+a_{3,0}+a_{3,1})\,,\end{aligned}$$

wobei ich zur Verdeutlichung eigentlich unnötige Klammern gesetzt habe.

Übungsaufgabe 1.14:
Berechnen Sie die Summe

$$\sum_{i=-1}^{1} \sum_{j=-2i}^{i} i^2 \cdot j \ . \qquad \blacksquare$$

Ebenso wie bei Summen möchte man auch bei mehrfachen Produkten die Pünktchenschreibweise wie etwa

$$a_1 \cdot a_2 \cdot a_3 \cdot \cdots \cdot a_n \tag{1.12}$$

vermeiden, und ganz genauso wie bei Summen verfährt man auch hier: Die Abkürzung für „Produkt" ist „P" und der griechische Buchstabe P ist das $\prod$ (Pi). Damit wird (1.12) zu

$$\prod_{i=1}^{n} a_i \ .$$

Alle Bemerkungen, die ich oben bei den Summen gemacht habe, übertragen sich wörtlich auf die Produkte, weshalb ich sie mir und Ihnen hier eigentlich ersparen will.

In der Wahrscheinlichkeitsrechnung und Statistik geht man häufig mit dem Produkt mehrerer aufeinander folgender natürlicher Zahlen um, meist startend mit 1. Dafür haben wir zwar bereits eine Notation, das Produkt der ersten n natürlichen Zahlen ist ja gerade $\prod\limits_{i=1}^{n} i$, aber auf die Dauer ist das ein wenig unhandlich. Man hat dafür daher die abkürzende Bezeichnung $n!$ (sprich: „n Fakultät") eingeführt:

Fakultät

Für jede natürliche Zahl n ist

$$n! \ = \prod_{i=1}^{n} i \ .$$

Zusätzlich setzt man $0! = 1$.

Das Fatale ist, dass das hier verwendete Ausrufzeichen „!" im Gegensatz zum Summen- und Produktsymbol auch in der Alltagssprache bzw. -schrift verwendet wird und man daher bei mathematischen Texten manchmal sehr genau hinschauen muss, was gemeint ist. Stellt man beispielsweise in einer Klausur die Aufgabe: „Berechnen Sie 8!", so kann man als Antwort durchaus erhalten: „Erstens brauche ich 8 nicht zu berechnen, weil's schon da steht, und zweitens empfinde ich das Ausrufzeichen am Ende der Aufgabe als ziemlich unhöflich." Nun ja.

Mit der Einführung des Fakultätszeichens ist dieser Abschnitt auch fast schon wieder zu Ende, denn was man damit alles berechnen und darstellen kann, das lernen Sie im Laufe Ihres Studiums. Zeigen möchte ich Ihnen nur noch, wie man damit ein beliebiges (also nicht bei 1 beginnendes) Produkt natürlicher Zahlen ganz einfach darstellen kann: Sind k und n natürliche Zahlen und ist $k \leq n$, dann ist

$$\prod_{i=k+1}^{n} i = \frac{n!}{k!} . \tag{1.13}$$

So langsam müssen Sie sich daran gewöhnen, dass man in der Mathematik prinzipiell alles beweisen muss. Allerdings ist der Beweis dieser Aussage nicht allzu schwer, man muss nur die Definition der Fakultät benutzen, wobei ich ausnahmsweise, da hier kein Zweifel über den Laufbereich der Variablen vorliegen kann, noch mal auf die Pünktchenschreibweise zurückgreife: Es ist

$$\frac{n!}{k!} = \frac{1 \cdot 2 \cdot 3 \cdots (k-1) \cdot k \cdot (k+1) \cdots (n-1) \cdot n}{1 \cdot 2 \cdot 3 \cdots (k-1) \cdot k} ,$$

wobei ich zur Verdeutlichung der Situation im Zähler einige der inneren Faktoren mit hineingeschrieben habe. Da k nicht größer als n ist, kann ich nun den kompletten Nenner herauskürzen und erhalte

$$\frac{n!}{k!} = (k+1) \cdot (k+2) \cdots (n-1) \cdot n ,$$

was mit der linken Seite von (1.13) übereinstimmt.

Übungsaufgabe 1.15:
a) Berechnen Sie 8!.
b) Stellen Sie das Produkt $9 \cdot 10 \cdot 11$ mithilfe von Fakultäten dar. ■

Damit sind wir – erschöpft, müde und hungrig (hat da jemand „Pizza!“ gesagt?) – am Ende dieses einführenden Kapitels angelangt. Sicherlich war Ihnen vieles schon vertraut, oder Sie konnten es zumindest mithilfe dieser Zeilen wieder aus den Tiefen Ihres Gedächtnisses hervorholen, und vielleicht habe Sie sich sogar ab und an gefragt: „Muss das sein?“ Ja, ich denke schon, denn für die nachfolgenden Kapitel und erst recht für das gesamte Studium brauchen Sie eine Basis, auf der Sie aufbauen können, und die wurde nun gelegt; und wenn Sie sich vielleicht die ganze Zeit schon fragen: „Die ganze Zahlenrechnerei und Notationslyrik ist ja gut und schön, aber wann kommt denn endlich *richtige* Mathematik?“, dann kann ich antworten: „Jetzt gleich, lesen Sie weiter!“

2 Grundlegendes über Funktionen

Für die meisten Menschen beginnt die „richtige“ Mathematik beim ersten Auftreten von Funktionen: Sobald man ein Funktionsschaubild hinzeichnen und Verläufe von Funktionen bestimmen kann, geht ihnen das mathematische Herz auf. Ich muss gestehen, dass ich auch nicht ganz frei bin von dieser Denkweise, wenngleich ich betonen will, dass all das, was Gegenstand des ersten Kapitels war, durchaus ernst zu nehmende Mathematik ist. Aber dennoch will ich gern einstimmen in den Chor derer, die sagen: „Jetzt geht's richtig los!“

Das Problem ist allerdings: Kaum ist die Freude über das Auftreten der Funktionen verflogen, beginnt der Bammel vor denselben: Was ist das eigentlich, eine Funktion, was für wichtige Eigenschaften kann sie haben, welche speziellen Funktionen muss man kennen usw.?

Diese Befürchtungen sind aber wirklich nicht begründet, Funktionen etwas „ganz Normales“. Um Ihnen ein Gefühl dafür zu geben, beginne ich mit einem ganz einfachen Beispiel; stellen Sie sich vor, Sie fahren mit Ihrem Auto einige Stunden lang mit der konstanten Geschwindigkeit von 120 km/h. (Ich weiß, dass das heutzutage schwer vorstellbar ist, aber versetzen Sie sich einfach mal gedanklich ins Innere Australiens, da geht so was.) Zur Berechnung des nach einer Zeit von t Stunden zurückgelegten Weges von s Kilometern kann man dann die Formel

$$s \; = \; 120 \cdot t$$

aufstellen, wobei ich nach bester Mathematiker-Manier großzügig auf die Einheiten verzichtet habe. Bei Anwendungen, diese Warnung möchte ich hier gleich anbringen, sollte man die physikalischen Einheiten allerdings sicherheitshalber stets mitführen, entweder explizit auf beiden Seiten oder in eckigen Klammern für die ganze Gleichung.

Mit obiger Formel kann man ausrechnen, dass Sie nach einer Stunde 120 Kilometer zurückgelegt haben, nach zwei Stunden 240 Kilometer und beispielsweise nach fünf Stunden 600 Kilometer. Es ist aber ebenso gut möglich, auszurechnen, wie weit Sie nach einer halben Stunde gekommen sind (nämlich $120 \cdot \frac{1}{2} = 60$, also 60 km), wie weit nach einer Minute (2 km), und wenn Sie Lust haben auch die zurückgelegte Strecke nach einer Sekunde oder auch nach $\frac{273}{334}$ Stunden. Sie sehen also, dass man für t *jeden beliebigen* positiven Wert einsetzen kann, und abhängig von diesem Wert liefert die Formel die zugehörige Strecke s. Um diese Abhängigkeit stärker zu betonen, schreibt man meist

$$s(t) \; = \; 120 \cdot t \; , \tag{2.1}$$

und schon haben wir – heimlich, still und leise – das erste Beispiel einer Funktion gesehen. Da t hier von überhaupt nichts abhängt, nennt man es auch die **unabhängige Variable**, entsprechend s die (von t) abhängige Variable; für jeden beliebigen Eingabewert von t, mit dem Sie die Funktion füttern, liefert diese Ihnen einen eindeutig bestimmten Ausgangswert s. Genau das ist das Wesen einer Funktion, und das werde ich im nächsten Abschnitt erst einmal präzise hinschreiben.

2.1 Definitionsbereich, Wertevorrat und Bildmenge

Da es im Leben nicht nur Autos und gefahrene Zeiten gibt (was man bei manchen Leuten allerdings glauben könnte), nennt man die unabhängige Variable in der Mathematik meist x (und nicht t) und die Funktion bezeichnet man mit f (und nicht s).

Funktion, Definitionsbereich und Wertevorrat

Es seien D und W Mengen, die nicht leer sein sollten; üblicherweise, aber nicht zwingend notwendig, werden das Mengen von Zahlen sein. Eine **Funktion** f ist eine Vorschrift (Abbildungsvorschrift, Formel, ...), die jedem Element x der Menge D auf eindeutig bestimmte Art und Weise ein Element $y = f(x)$ von W zuordnet.

Man nennt D den **Definitionsbereich** von f (weil f für alle Werte x aus D definiert sein soll) und W den **Wertevorrat** von f, weil y als **Funktionswert** von x bezeichnet wird und W eben alle möglichen Funktionswerte von f in sich versammelt.

Oft schreibt man kurz $f : D \to W$ und hat so auf einen Schlag Definitionsbereich und Wertevorrat von f definiert. Ich werde das im Folgenden gelegentlich, aber nicht immer so tun.

Ich vermute, Sie brauchen jetzt erst mal einen starken Kaffee, um das Gelesene zu verdauen. Vielleicht kann ich Ihnen aber auch mit ein paar Beispielen helfen:

Beispiel 2.1:
Ich wähle erst mal der Bequemlichkeit halber sowohl als Definitionsmenge als auch als Wertevorrat die Menge der reellen Zahlen, also $D = W = \mathbb{R}$. Die erste Abbildungsvorschrift, die ich testen will, ist

$$f : \mathbb{R} \to \mathbb{R}\ , \quad f(x) \ = \ 2x\ .$$

Das ist nun eine Funktion reinsten Wassers, denn sie weist jeder reellen Zahl x einen eindeutig bestimmten Wert zu, nämlich das Doppelte von x, und dieser Wert liegt auch im zulässigen Wertebereich, nämlich der Menge der reellen Zahlen. ■

Ursprünglich hatte ich an dieser Stelle vorgesehen, in einer Übungsaufgabe einige Funktionswerte von f berechnen zu lassen, aber das scheint mir jetzt doch zu kindisch.

Beispiel 2.2:
Stattdessen ein zweites Beispiel, die Abbildungsvorschrift sei dieses Mal

$$f : \mathbb{R} \to \mathbb{R} \, , \quad f(x) \; = \; x^2 \; .$$

Auch dieses ist eine Funktion, denn jedem reellen Eingabewert x wird eine eindeutige Zahl, nämlich gerade sein Quadrat, zugewiesen. Dennoch weist dieses Beispiel eine Neuigkeit auf: Getreu der Regel „Minus mal Minus ergibt Plus“ treten hier als Funktionswerte nämlich niemals negative Zahlen auf, obwohl sie prinzipiell durchaus zugelassen wären. ■

Es ist also wichtig zu betonen, dass der Wertevorrat W die Menge aller als Funktionswert zugelassenen Zahlen enthält, diese müssen aber nicht unbedingt auch wirklich von der Funktion „getroffen“ werden. Das sollte festgehalten werden:

Der Wertevorrat muss nicht ausgeschöpft werden

Der Wertevorrat W einer Funktion enthält die Menge aller *potenziellen* Funktionswerte; er kann durchaus Zahlen enthalten, die niemals als Funktionswert auftreten.

Die Menge aller Elemente von W, die tatsächlich als Funktionswert auftreten, bezeichne ich als **Bildmenge** von D und symbolisiere sie mit $f(D)$.

Die Bildmenge der Funktion $f(x) = x^2$ ist also gerade die Menge der nicht negativen reellen Zahlen.

Ich gebe zu, dass die feinsinnige Unterscheidung zwischen Wertevorrat und Bildmenge nicht jedermanns Lieblingsthema ist, aber wichtig ist sie eben doch; schlimmer noch: Die Bezeichnungsweise und Notation ist in der Literatur nicht einheitlich, und es kann Ihnen durchaus passieren, dass später einmal ein Dozent oder Fachbuchautor das, was ich hier als Wertevorrat bezeichne, lieber Wertebereich nennen wird, die Bildmenge wiederum heißt dann vielleicht Wertemenge usw. Ich kann’s nicht ändern, Sie müssen einfach immer sorgfältig schauen, was der jeweilige Autor am Anfang des Textes definiert.

Ein weiteres, ziemlich eintöniges, aber korrektes Beispiel einer Funktion, die ihren Wertevorrat nicht ausschöpft, ist das folgende: Es sei $D = \mathbb{R}$, $W = \{0, 1, 2\}$ und $f(x) = 1$.

Was ist hier los? Nun, dies ist sicherlich eine Funktion, denn es wird jeder reellen Zahl x in eindeutiger Weise ein Wert zugeordnet, nämlich die Zahl 1. Der Wertevorrat dieser Funktion ist schon nicht allzu groß, er umfasst nur drei Zahlen, und dennoch schöpft sie ihn nicht aus, denn die Zahlen 0 und 2 treten niemals als Funktionswert in Erscheinung. Hier ist also die Bildmenge gerade die ein-elementige Menge $\{1\}$.

Sehr oft ist der Definitionsbereich einer Funktion ein Intervall, und diesen Begriff sowie die damit zusammenhängende Bezeichnungsweise will ich an dieser Stelle gleich einführen:

Intervall

Sind a und b reelle Zahlen, wobei a kleiner oder gleich b sein soll, also $a \leq b$, so nennt man die Menge aller reellen Zahlen, die zwischen a und b liegen, das **Intervall** mit den Grenzen a und b.

Gehören a und b selbst auch zu diesem Intervall, so nennt man dies ein **abgeschlossenes Intervall** und bezeichnet es mit eckigen Klammern:

$$[a, b] \ = \ \{x \in \mathbb{R} \ \text{ mit } a \leq x \leq b\} \ .$$

Gehören a und b nicht zu diesem Intervall, so nennt man dies ein **offenes Intervall** und bezeichnet es mit runden Klammern:

$$(a, b) \ = \ \{x \in \mathbb{R} \ \text{ mit } a < x < b\} \ .$$

Auch die Mischformen

$$(a, b] \ = \ \{x \in \mathbb{R} \ \text{ mit } a < x \leq b\}$$

und

$$[a, b) \ = \ \{x \in \mathbb{R} \ \text{ mit } a \leq x < b\}$$

existieren, man nennt sie **halboffene Intervalle**.

Mithilfe dieser Notation kann man also beispielsweise eine Funktion, deren Definitionsbereich die Menge aller rellen Zahlen zwischen 2 und 3 (unter Einschluss dieser beiden) sein soll, kurz in der Form

$$f : [2, 3] \to \mathbb{R}$$

notieren.

Wie wäre es denn an dieser Stelle mit einem kleinen *mental snack* zwischendurch?

Übungsaufgabe 2.1:
Bestimmen Sie die genaue Bildmenge der Funktion

$$f : [-1, 2] \to \mathbb{R} \ , \quad f(x) \ = \ \frac{1}{x-3} \ . \qquad \blacksquare$$

Wenn Sie nun so langsam denken, dass so ziemlich alles unter der Sonne, was mit „$f(x) =$“ losgeht, eine Funktion ist, so muss ich Ihnen das Folgende zeigen:

Es sei $D = \mathbb{R}$, $W = \{-1, 1\}$ und für alle $x \in D$

$$f(x) \ = \ 1 \text{ oder } -1, \text{ je nachdem, wie ich mich fühle.}$$

Das ist nun sicherlich keine eindeutige Vorschrift zur Ermittlung des Wertes $f(x)$, denn woher sollen Sie wissen, wie ich mich fühle, ich weiß es ja selbst oft nicht so genau; somit ist f *keine* Funktion.

Aber auch die folgenden beiden, wieder eher mathematisch angehauchten Beispiele stellen – aus unterschiedlichen Gründen – keine Funktionen dar:

Setzen wir beispielsweise

$$g : \mathbb{N} \to \mathbb{N} \ , \quad g(x) = \sqrt{x} \, ,$$

so stellt sich relativ schnell heraus, dass beispielsweise der Wert $g(2)$, also $\sqrt{2}$, nicht im angegebenen Wertevorrat von g liegt, denn $\sqrt{2}$ ist keine natürliche Zahl (nicht einmal eine rationale, aber darauf kommt's jetzt auch nicht mehr an).

Aber auch wenn ich den Definitionsbereich und den Wertevorrat von g erweitere zu $D = W = \mathbb{R}$, erweise ich der Möchtegern-Funktion einen Bärendienst, denn jetzt ist sie nicht mal mehr auf der gesamten Menge D definiert, da reelle Wurzeln aus negativen Zahlen nicht erklärt sind.

Sie sehen also, man muss die Funktionsvorschrift stets im Zusammenhang mit dem Definitionsbereich und dem Wertevorrat prüfen, um sicher sagen zu können, ob es sich wirklich um eine Funktion handelt. Und – um doch einmal in den bei vielen Autoren beliebten Krankenschwestern-Plural zu verfallen – das wollen wir doch gleich noch ein wenig üben:

Übungsaufgabe 2.2:
Welche der folgenden Vorschriften stellt eine Funktion dar?

a) $f : \mathbb{N} \to \mathbb{N} \ , \quad f(x) = x^2$

b) $g : [-1, 1] \to \mathbb{R} \ , \quad g(x) = \dfrac{1}{x}$

c) $h : \mathbb{Q} \to \mathbb{Q} \ , \quad h(x) = 2x^4 - \dfrac{5}{3}x^3 + \dfrac{1}{2}x$ ■

Liegen sowohl der Definitionsbereich als auch der Wertevorrat einer Funktion in den rellen Zahlen – man sagt dann auch kurz, es handle sich um eine **reelle Funktion** –, so kann man ihren Verlauf auch bildlich darstellen. Wenngleich Sie das ganz sicher von verschiedensten Gelegenheiten her bereits kennen, will ich es hier der Vollständigkeit halber kurz aufschreiben:

Man beginnt damit, ein **Koordinatensystem** zu zeichnen, also zwei senkrecht zueinander stehende Achsen, die sich üblicherweise im jeweiligen Nullpunkt schneiden. Auf diesen nimmt man durch Markierung von mindestens einer Zahl eine Einteilung vor. Die horizontale, nach rechts gerichtete Achse bezeichnet man meist als x-Achse, da man darauf die x-Werte markiert, die vertikale, nach oben gerichtete Achse heißt meist y-Achse, da man zumindest in früherer Zeit die Funktionswerte, welche man auf dieser Achse markiert, nicht als $f(x)$, sondern einfach als y bezeichnete.

Nach dem großen Philosophen und Mathematiker René Descartes, dessen Name in damals üblicher Weise latinisiert wurde zu Renatus Cartesius, nennt man ein derartiges Koordinatensystem auch **kartesisches Koordinatensystem**.

Um sich nun im wahrsten Sinne des Wortes ein Bild vom Verlauf einer Funktion zu machen, interpretiert man für jedes x aus dem Definitionsbereich der Funktion das

Zahlenpaar $(x, f(x))$ als Koordinaten eines Punktes in diesem gerade gezeichneten System und zeichnet diesen Punkt ein. So ergeben sich beispielsweise für die Funktion $f(x) = 3x$ die Paare $(0, 0)$, $(1, 3)$, $(-2, -6)$ und $(100, 300)$.

Macht man das nicht nur für vier, sondern für ganz viele ganz dicht beieinander liegende Pärchen, so ergeben die solchermaßen gezeichneten Punkte bei nicht allzu verrückten Funktionen eine durchgehende Linie, die man als **Schaubild** oder **Graph** der Funktion bezeichnet. Ein ziemlich willkürlich gewähltes Beispiel sehen Sie in Abbildung 2.1.

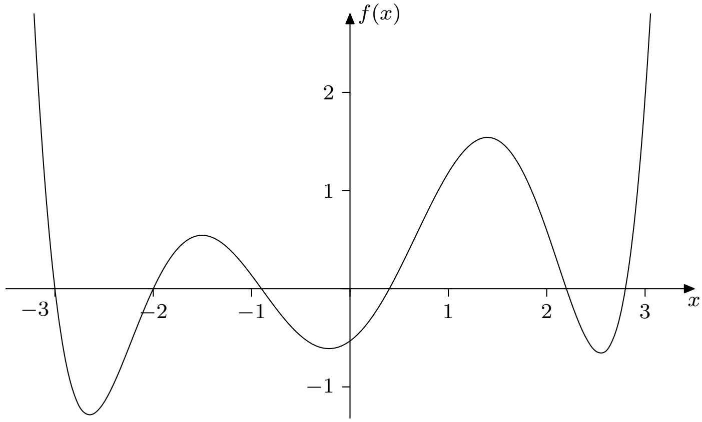

Abb. 2.1 Graph einer Funktion

Übrigens schreibt man auch nach der vor ein paar Jahren eingeführten neuen Rechtschreibung das Wort „Graph“ mit „ph“, ganz im Gegensatz zu so schönen damit verwandten Wörtern wie „grafisch“ oder „Biografie“. Meiner persönlichen Vermutung nach liegt das einfach daran, dass in der „f“-Schreibweise der Ausdruck „Graf von f“ doch gar zu sehr an „Graf von Monte Christo“ erinnern würde.

Aber nun zurück zur Mathematik.

2.2 Verkettung von Funktionen; Monotonie und Umkehrbarkeit

Nehmen wir einmal an, Sie sind glücklicher Inhaber einer kleinen Firma, die für eine Reihe stark nachgefragter Produkte das Monopol besitzt, und sie wollen die Preise für Ihre Produkte neu kalkulieren. Da Sie ordentlich verdienen wollen, gehen Sie dabei wie folgt vor: Zuerst wird der alte Preis um 100 erhöht, und danach wird das Ganze noch mal verdreifacht. (Sollte Ihnen das unverschämt erscheinen, so bin ich ganz Ihrer

Meinung, aber beispielsweise machen das die Ölkonzerne bei der Neubestimmung ihrer Benzinpreise genauso.)

Da Sie heimlich ein Mathematikbuch gelesen haben, beschreiben Sie diese beiden Operationen jeweils durch eine Funktion, die auf der Menge der positiven Zahlen definiert sein soll: Ist x der alte Preis, so lautet die zuerst anzuwendende Funktion

$$g(x) \;=\; x + 100\,,$$

und die zweite ist

$$f(y) \;=\; 3 \cdot y\;.$$

So berechnen Sie beispielsweise für $x = 100$, dass $g(x) = 200$ und $f(200) = 600$ ist; ein Produkt, das vorher 100 Euro kostete, schlägt jetzt mit 600 Euro zu Buche.

So können Sie das nach und nach mit allen Produktpreisen machen. Beispielsweise kommt ein Produkt, das eingangs $x = 10$ Euro kostete, nach Ihrer kleinen Neukalkulation auf $g(x) = 110$ und schließlich $f(110) = 330$ Euro.

Allerdings fällt Ihnen wohl ziemlich bald auf, dass Sie die zwischenzeitliche Bestimmung von $g(x)$ eigentlich weglassen können und lieber gleich den Ausgabewert von g wieder in f hineinstecken können. Was Sie dabei machen, nennt man in der Mathematik die **Verkettung** der Funktionen g und f. Die durch die Verkettung entstandene neue Funktion nenne ich für den Moment h; ich berechne sie explizit, indem ich $g(x)$ in f hineinstecke und ausrechne:

$$h(x) \;=\; f(g(x)) \;=\; f(x+100) \;=\; 3\cdot(x+100) \;=\; 3x + 300\;.$$

Es ist also

$$h(x) \;=\; 3x + 300 \quad ,$$

und auch diese Funktion ist auf der Menge aller positiven Zahlen definiert. In dieser Darstellung können Sie nun etwas schneller als oben nachrechnen, dass $h(100) = 600$ und $h(10) = 330$ ist.

So geht das prinzipiell für alle Funktionen der Welt, wobei nur gewährleistet sein muss, dass die Funktion f mit dem Ausgabewert von g etwas anzufangen weiß, dass also die Bildmenge von g im Definitionsbereich von f liegt.

Das sollte ich jetzt wirklich mal präzise aufschreiben:

Verkettung von Funktionen

Es seien $f : F \to W$ und $g : D \to E$ zwei Funktionen mit der Eigenschaft, dass der Definitionsbereich F von f im Bildbereich $g(D)$ von g liegt.

Dann heißt die Funktion $f \circ g : D \to W$, definiert durch

$$f \circ g : D \to W\;, \quad f \circ g(x) = f(g(x))$$

für alle $x \in D$, die **Verkettung** von f und g; gelegentlich sagt man auch **Hintereinanderausführung** oder **Komposition**.

Zu beachten ist die Reihenfolge: Es wird *zuerst* g und *danach* f angewendet.

Ich habe keine Ahnung, wie man das Zeichen $\circ$ aussprechen soll, ich habe es in meinem Leben bisher immer nur hingeschrieben. Manche Leute sagen dazu „Kringel“, andere „verknüpft mit“, und wieder andere sagen einfach „nach“, „$f \circ g$“ wird also in diesem Fall als „f nach g“ ausgesprochen. Das ist gar nicht so schlecht, wie es auf den ersten Blick scheinen mag, denn es verdeutlicht noch mal die Reihenfolge, in der die beiden Funktionen nacheinander auszuführen sind: Zuerst g und danach f. Das macht man gerade am Anfang oft falsch, da man eben von links nach rechts zu lesen gewohnt ist (möglicherweise machen Japaner hier weniger Fehler, ich weiß es nicht). Die Schreibweise $(f \circ g)(x)$ soll suggerieren, dass zunächst g auf x angewendet wird, weil es direkt daneben steht, und erst danach f an die Reihe kommt.

Bevor ich mich nun weiter in philosophische Betrachtungen über Vor- und Nachteile dieser Schreibweise verliere, lieber Beispiele; um diese etwas kompakter hinzuschreiben, möchte ich gerne für die Menge der nicht negativen reellen Zahlen das Symbol $\mathbb{R}^+$ einführen, also

$$\mathbb{R}^+ \;=\; \{x \in \mathbb{R}; x \geq 0\}\,.$$

Dann kann's losgehen:

Beispiel 2.3:
Betrachten wir die Funktionen

$$g : \mathbb{R} \to \mathbb{R}\;,\quad g(x) \;=\; (x+9)^2$$

und

$$f : \mathbb{R}^+ \to \mathbb{R}\;,\quad f(y) \;=\; \sqrt{y} - 8\;.$$

Da die Bildmenge von g gerade die nicht negativen reellen Zahlen sind („Minus mal Minus ...“!), kann ich unbesorgt f mit g verketten und erhalte:

$$(f \circ g) : \mathbb{R} \to \mathbb{R}\;,\quad (f \circ g)(x) \;=\; f((x+9)^2) \;=\; \sqrt{(x+9)^2} - 8 \;=\; x+9-8 \;=\; x+1\,,$$

also $(f \circ g)(x) = x + 1$. Die Verkettung bewirkt also gerade mal die Erhöhung von x um 1. Sie müssen zugeben, dass wir hier gegenüber der separaten Auswertung von f und g eine Menge Arbeit gespart haben! ■

Beispiel 2.4:
Jetzt seien

$$g : \mathbb{R}^+ \to \mathbb{R}\;,\quad g(x) \;=\; \sqrt{\frac{x+2}{x+1}}$$

und

$$f : \mathbb{R} \to \mathbb{R}\;,\quad f(y) \;=\; \frac{1}{y^2}.$$

Die Funktion f verdaut als Eingabewert jede reelle Zahl außer der Null, und da diese als Wert von g nicht vorkommt, brauchen wir uns um das Zusammenpassen der Bereiche keine Sorgen zu machen; das Berechnen bzw. Vereinfachen der Verkettung mache ich jetzt mal in Einzelschritten: Setzt man $g(x)$ in f ein, so wird $g(x)$ quadriert und in den Nenner gepackt; das Quadrieren hebt aber gerade die Wurzelbildung auf, sodass der *Nenner* von f jetzt $\frac{x+2}{x+1}$ heißt. Und da man durch einen Bruch dividiert, indem man mit seinem Kehrbruch multipliziert, erhalten wir als Ergebnis:

$$(f \circ g) : \mathbb{R}^+ \to \mathbb{R} \ , \quad (f \circ g)(x) \ = \ \frac{x+1}{x+2} \ . \qquad \blacksquare$$

Vielleicht haben Sie sich die ganze Zeit schon gefragt, was passiert, wenn man in obigen Beispielen bei der Verkettung die Reihenfolge der verketteten Funktionen vertauscht? Zu spät, ich war schneller und frage *Sie:*

Übungsaufgabe 2.3:
Prüfen Sie bei den in den obigen Beispielen behandelten Funktionenpaaren, ob auch die Verkettung $g \circ f$ möglich ist; falls ja, geben Sie die verkettete Funktion möglichst einfach an. ■

Eine gewisse Sonderrolle bei der Verkettung von Funktionen spielt die so genannte Umkehrfunktion einer Funktion. Das ist diejenige Funktion, die eine gegebene Funktion durch Verkettung wieder rückgängig macht; ich werde das gleich noch genauer definieren, erst mal ein paar Beispiele:

Verkettet man die Funktion

$$g : \mathbb{R} \to \mathbb{R} \ , \quad g(x) = x - 28$$

mit der ebenfalls auf ganz $\mathbb{R}$ definierten Funktion $f(y) = y + 28$, so ergibt sich

$$(f \circ g) : \mathbb{R} \to \mathbb{R} \ , \quad (f \circ g)(x) \ = \ (x - 28) + 28 \ = \ x$$

und ebenso

$$(g \circ f) : \mathbb{R} \to \mathbb{R} \ , \quad (g \circ f)(x) \ = \ (x + 28) - 28 \ = \ x \ .$$

Die Verkettung der beiden Funktionen lässt also unabhängig von der Reihenfolge jedes x unverändert, so als ob gar keine Funktion angewandt worden wäre. Anders formuliert: Die jeweils zuerst angewandte Funktion wird durch die Anwendung der zweiten umgekehrt. Das ist natürlich völlig unabhängig von der speziellen Zahl 28 und entspricht im Übrigen völlig der täglichen Lebenserfahrung: Wenn Sie beispielsweise im Roulette einen bestimmten Betrag gewinnen und diesen im nächsten Spiel gleich wieder komplett verlieren, so hat sich insgesamt an Ihrer finanziellen Situation nichts geändert, getreu dem alten Sprichwort: „Wie gewonnen, so zerronnen."

Auch im nächsten Beispiel klappt das mit dem Umkehren noch ganz gut: Um die Funktion

$$g : \mathbb{R} \to \mathbb{R} \ , \quad g(x) = -2x$$

wieder umzukehren, muss man sie nur mit der Funktion $f(y) = -\frac{1}{2}y$ verketten, denn dann ist für alle $x \in \mathbb{R}$:

$$(f \circ g) : \mathbb{R} \to \mathbb{R}\ , \quad (f \circ g)(x) \ = \ -\frac{1}{2} \cdot (-2x) \ = \ x$$

und ebenso

$$(g \circ f) : \mathbb{R} \to \mathbb{R}\ , \quad (g \circ f)(x) \ = \ -2 \cdot (-\frac{1}{2}x) = \ x \ .$$

Alles schön und gut, werden Sie sagen, wo ist das Problem? Die Probleme kommen noch, keine Sorge. Zuvor wird es aber höchste Zeit, den Begriff Umkehrfunktion exakt zu definieren. Dabei sollen natürlich die gerade betrachteten Beispiele als Leitfaden dienen.

Umkehrfunktion

Es sei f eine Funktion mit dem Definitionsbereich D und der Bildmenge $f(D)$. Eine Funktion, nennen wir sie f^{-1}, deren Definitionsbereich $f(D)$ ist und die die Eigenschaft

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

für alle $x \in D$ hat, nennt man **Umkehrfunktion** von f. Existiert zu einer Funktion f eine Umkehrfunktion, so nennt man f selbst **umkehrbar**.

Nicht jede Funktion ist umkehrbar.

Vielleicht liegt Ihnen ja schon eine ganze Weile die Frage auf der Zunge, warum um alles in der Welt man eigentlich an der Umkehrfunktion interessiert ist, man könnte ja auch gleich die Ausführung der Funktion bleiben lassen und sparte sich dadurch auch die Durchführung der Umkehrfunktion. Das ist für sich betrachtet schon richtig, aber meist ist man eben nicht an der Unwirksammachung einer Funktion interessiert, sondern man möchte anhand eines vorgelegten Funktionswertes $f(x)$ Rückschlüsse auf den Wert x ziehen, von dem der Funktionswert herstammt. Und dazu dient die Umkehrfunktion.

Noch irgendwelche Fragen, mein lieber Watson? Nun, ich glaube schon, und das liegt wahrscheinlich an meiner verquasten Ausdrucksweise. Ich mache lieber mal ein Beispiel. Denken Sie einmal an das ganz am Anfang dieses Kapitels formulierte Beispiel über das Autofahren zurück. Dort hatte ich – in Formel (2.1) – die Funktion

$$s(t) \ = \ 120 \cdot t$$

aufgestellt, die für ein mit Geschwindigkeit 120 fahrendes Auto zu jedem Zeitpunkt t die zurückgelegte Strecke $s(t)$ berechnet.

Nun möchten Sie aber vielleicht auch einmal in umgekehrter Fragestellung wissen, wie lange Sie zum Zurücklegen einer bestimmten Strecke brauchen, wenn Sie konstant 120 fahren. Mit anderen Worten: Sie fragen danach, zu welchem Wert der Zeit t eine von Ihnen ausgewählte Strecke s gehört. Das ist aber nichts anderes als die Frage nach der Umkehrfunktion der Funktion $s(t)$.

Mit ein wenig Überlegung finden Sie heraus, dass diese Umkehrfunktion s^{-1} gegeben ist durch

$$s^{-1}(y) \;=\; \frac{y}{120}\;,$$

denn

$$s^{-1}(s(t)) \;=\; \frac{s(t)}{120} \;=\; \frac{120t}{120} \;=\; t$$

für alle t. Wenn Sie also beispielsweise wissen wollen, wie lange Sie brauchen, um 300 Kilometer weit zu kommen, rechnen Sie einfach

$$s^{-1}(300) \;=\; \frac{300}{120} \;=\; \frac{5}{2}\;.$$

Sie brauchen also $\frac{5}{2}$, sprich zweieinhalb Stunden. So einfach kann das Leben mit ein wenig Mathematik sein!

Die letzte, etwas verschämt am Ende des obigen Kastens stehende Bemerkung ist es, die uns auf die oben bereits erwähnten Probleme hinführt. Als Beispiel zeige ich Ihnen die Funktion

$$f : \mathbb{R} \to \mathbb{R}\;,\quad f(x) \;= x^2\;,$$

die – bzw. deren Funktionsgraphen – man auch als **Normalparabel** bezeichnet. Sie ist ein Spezialfall der so genannten Polynome, die ich weiter unten mit Ihnen anschauen will; fürs Erste haben wir mit dieser einen Funktion genug zu tun. Bereits oben haben wir überlegt, dass die Bildmenge dieser Funktion gerade die Menge der nicht negativen reellen Zahlen ist.

Falls f auf dem gesamten Definitionsbereich $\mathbb{R}$ umkehrbar sein soll, dann muss man also eine Funktion finden, die auf ganz $\mathbb{R}^+$ definiert ist und überall dort die Wirkung von f gerade umkehrt.

Sie werden denken: „Nun sag' schon endlich, dass es die Wurzelfunktion ist, und lass' mich einen Kaffee trinken gehen.“ Tut mir Leid, Kaffee ist erst mal gestrichen, und zwar aus folgendem Grund: Zunächst einmal sieht die Sache ja ganz gut aus, denn die Wurzelfunktion, also die Funktion $w(x) = \sqrt{x}$, die jeder Zahl ihre (Quadrat-) Wurzel zuordnet, ist ja tatsächlich für jede nicht negative ganze Zahl, mit anderen Worten also auf $\mathbb{R}^+$, definiert.

Auch wenn man beispielsweise $w(4) = 2$ berechnet ist man noch guter Dinge, denn da $f(2) = 2^2 = 4$ ist, kehrt die Wurzel tatsächlich diesen Vorgang um, da

$$w(f(2)) \;=\; w(4) \;=\; 2\;,$$

und das kann man nicht nur für 2, sondern für jede positive Zahl machen.

Aber jetzt kommt's: Versuchen Sie mal, die Funktion $f(x) = x^2$ an der Stelle $x = -2$ umzukehren! Sie erhalten als Zwischenresultat $f(-2) = 4$, und wenn Sie darauf nun die Wurzel anwenden, erhalten Sie wie gerade gesehen als Wert 2, und *nicht* -2. Liegen also sowohl $+2$ als auch -2 im Definitionsbereich, so ist die Funktion nicht umkehrbar, und genauso wenig ist sie es an irgendeinem anderen negativen Wert, wenn der zugehörige positive Wert im Definitionsbereich liegt. Insgesamt hat sich also herausgestellt:

> Die auf ganz $\mathbb{R}$ definierte Funktion $f(x) = x^2$ ist nicht umkehrbar.

Woran liegt das? Nun, dazu muss man sich noch mal überlegen, was „Umkehrung“ einer Funktion f eigentlich bedeutet: Es geht dabei darum, jeden Funktionswert $f(x)$ von f auf seinen Ausgangswert – man sagt auch sein **Urbild** – x „zurückzuwerfen“, so, als ob gar keine Abbildung stattgefunden hätte. Dazu ist es aber doch offensichtlich nötig, dass man bei jedem Funktionswert eindeutig sagen kann, woher er kam, d. h., von welchem x-Wert er herstammt. Das geht aber bei der Normalparabel $f(x) = x^2$ nicht, denn schon bei meinem Beispiel 4 als Funktionswert kann niemand mehr sagen, woher dieser Wert kam, ob von -2 oder von $+2$.

Warum geht denn aber nun bei dieser doch eigentlich auch recht harmlosen Funktion nicht mehr, was bei den beiden oben genannten Beispielen noch recht gut ging? Nun, wenn man sich die Schaubilder der drei Funktionen in Abbildung 2.2 so ansieht, dann fällt auf, dass die ersten beiden, also $f_1(x) = x-28$ und $f_2(x) = -2x$, wunderbar gradlinig von unten nach oben bzw. umgekehrt verlaufen, ohne unterwegs noch einen Schlenker zu machen und noch mal ein Stück nach oben bzw. unten zurückzugehen. Die Normalparabel dagegen kommt zunächst mit viel Elan von oben nach unten an, kehrt dann aber am Nullpunkt plötzlich um und geht direkt wieder nach oben.

Das bedeutet aber, dass jeder positive Wert als Funktionswert zwei Mal auftritt, und genau daran scheitert die Umkehrbarkeit.

Ich fürchte, ich bin vom im Vorwort formulierten hehren Ziel, ein unterhaltsames Mathematikbuch zu schreiben, ein ganz klein wenig abgekommen: Geben Sie’s ruhig zu, Sie langweilen sich. Na ja, ich habe wirklich lange versucht, hier ein Pizza-Beispiel einzubauen, aber ich habe einfach keine Idee, wie man eine Pizza quadrieren sollte.

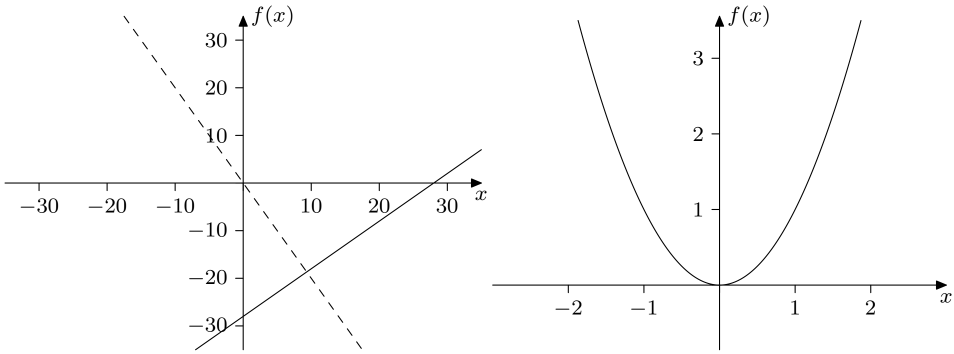

Abb. 2.2 Die Funktionen $f_1(x) = x - 28$ (links, schwarz), $f_2(x) = -2x$ (links, gestrichelt) und $f_3(x) = x^2$ (rechts)

Bei einer Tiefkühlpizza käme das ja auch der Quadratur des Kreises gleich! (O.K., vergessen Sie's, war einer meiner schlechteren Scherze.)

Ich werde daher jetzt einmal entgegen sonstiger Gewohnheit ohne lange Vorrede direkt aufs Ziel zusteuern: Es wird sich gleich herausstellen, dass die Eigenschaft, keine „Schlenker" zu machen, man kann auch sagen: „monoton" zu verlaufen, genau die richtige Voraussetzung für Umkehrbarkeit ist. Da aber in der mathematischen Fachterminologie meines Wissens der Begriff Schlenker nicht vorkommt, muss ich zunächst leider noch mal präzise sagen, was eine monotone Funktion sein soll:

Monotone Funktion

Es sei f eine Funktion und I eine Teilmenge des Definitionsbereichs D von f; das kann, muss aber nicht der ganze Definitionsbereich sein.

Man nennt f **monoton steigend** auf I, wenn gilt: Wenn immer x_2 größer ist als x_1 (also ‚rechts' von x_1 liegt), dann ist $f(x_2)$ größer oder gleich $f(x_1)$. In Kurzform:

Ist $x_1 < x_2$, so ist $f(x_1) \leq f(x_2)$.

Analog nennt man eine Funktion f **monoton fallend** auf I, wenn gilt: Wenn immer x_2 größer ist als x_1 (also ‚rechts' von x_1 liegt), dann ist $f(x_2)$ kleiner oder gleich $f(x_1)$. In Kurzform:

Ist $x_1 < x_2$, so ist $f(x_1) \geq f(x_2)$.

Ein Funktion heißt **monoton**, wenn sie monoton steigend oder monoton fallend ist.

Beispiele für beide Fälle haben wir oben schon gesehen: Die Funktion

$$f(x) \ = \ x - 28$$

ist überall in $\mathbb{R}$ monoton steigend; das sieht man nicht nur an der Abbildung (was kein Beweis wäre), sondern kann es auch nachrechnen: Wenn zwei Eingabewerte x_1 und x_2 so beschaffen sind, dass x_2 größer ist als x_1, dann ist auch $x_2 - 28$ größer als $x_1 - 28$, also $f(x_2)$ größer als $f(x_1)$, und da „größer" natürlich erst recht „größer oder gleich" bedeutet, ist die Funktion f monoton steigend; in Kurzform:

$$\text{Ist } x_1 < x_2, \text{ so ist } x_1 - 28 < x_2 - 28, \text{ also } f(x_1) < f(x_2).$$

Ebenso kann man sich überlegen, dass die Funktion $f(x) = -2x$ überall in $\mathbb{R}$ monoton fallend ist; ich denke, inzwischen haben Sie sich so weit an die Kurzform gewöhnt, dass ich Ihnen diese gleich ohne große Vorrede zumuten kann; wenn nicht, so schreiben Sie mir bitte oder wenden Sie sich vertrauensvoll an Ihren Arzt oder Apotheker – falls der Mann Mathematik gelernt hat. Nun aber:

$$\text{Ist } x_1 < x_2, \text{ so ist } -2x_1 > -2x_2, \text{ also } f(x_1) > f(x_2).$$

Falls Ihnen der „so ist"-Schritt merkwürdig vorkommt, so kann ich Ihnen erst mal nur ein Beispiel anbieten und Sie ansonsten bis zum Kapitel über Ungleichungen vertrösten. Als Beispiel nehme ich einfach mal $x_1 = 1$ und $x_2 = 2$. Sicherlich ist 1 kleiner als 2; wende ich aber die Funktion f auf diese beiden Werte an, so ergibt sich

$$f(1) = -2 \text{ und } f(2) = -4 ,$$

also ist $f(1)$ größer als $f(2)$, denn es liegt weiter rechts auf der Zahlengeraden.

Wenn Monotonie, wie eingangs gesagt, etwas mit Umkehrbarkeit zu tun hat, so müsste ja eigentlich die auf ganz $\mathbb{R}$ definierte Normalparabelfunktion $f(x) = x^2$, von der wir ja inzwischen wissen, dass sie nicht umkehrbar ist, auch bezüglich der Monotonie so ihre Probleme haben. Und tatsächlich stellt sich heraus, dass sie *nicht* monoton ist, weder steigend noch fallend, wenn man sie auf ganz $\mathbb{R}$ betrachtet. Um dies rechnerisch nachzuweisen, wähle ich zunächst die beiden Punkte $x_1 = -1$ und $x_2 = 0$ aus. Sicherlich ist $x_1 < x_2$, und da $f(-1) = 1$ *nicht* kleiner oder gleich $f(0) = 0$ ist, ist die Funktion nicht überall monoton steigend. Sie ist aber auch nicht überall monoton fallend, wie Sie beispielsweise anhand des neuen Punktepaares $x_1 = 0$ und $x_2 = 1$ leicht testen können (und sollten).

Beachten Sie, was hier passiert ist, denn es ist ein allgemeines Prinzip: Wenn man eine Vermutung widerlegen will (hier die Vermutung, dass die Funktion auf ganz $\mathbb{R}$ monoton steigend oder monoton fallend ist), so genügt *ein einziges* Gegenbeispiel.

Übungsaufgabe 2.4:
Prüfen Sie, ob die folgenden Funktionen monoton sind:

$$f : \mathbb{R} \to \mathbb{R} , \quad f(x) = \frac{x}{10} - 17$$

$$g : \mathbb{R} \to \mathbb{R} , \quad g(x) = -\frac{x}{17} + 10$$

$$h : [1, 17] \to \mathbb{R} , \quad h(x) = \frac{1}{x}$$ ■

Vielleicht haben Sie sich schon gewundert, dass ich bei der Definition von Monotonie auch die Gleichheit der Funktionswerte zugelassen, also nicht gefordert habe, dass beispielsweise bei einer monoton steigenden Funktion aus $x_1 < x_2$ folgt: $f(x_1)$ ist echt kleiner als $f(x_2)$.

Der Grund hierfür ist, dass man auch solche Funktionen monoton nennen will, die fast auf ihrem gesamten Definitionsbereich schön brav und tapfer ansteigen und sich nur zwischendurch einmal ganz kurz ausruhen und ein Stückchen konstant bleiben; ein Beispiel einer solchen Funktion sehen Sie in Abbildung 2.3.

Nun gibt es aber leider Funktionen, die diese Großzügigkeit meinerseits (und ich stehe damit innerhalb der Mathematikergemeinde nicht allein) schamlos ausnutzen und sich rein gar nicht vom Fleck bewegen; der wahrscheinlich unverschämteste Vertreter dieser Gattung ist die Funktion

$$f : \mathbb{R} \to \mathbb{R} , \ f(x) = 0 ,$$

die alles und jedes aus der Menge der reellen Zahlen auf den Wert Null abbildet oder bildlich gesprochen, die sich während ihres ganzen Lebens keinen Millimeter nach

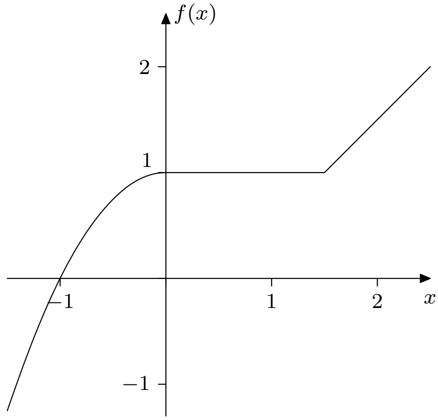

Abb. 2.3 Eine stückweise konstante monotone Funktion

unten oder oben bewegt. Das Dumme ist jetzt, dass diese Funktion nach der obigen Definition *sowohl monoton steigend als auch monoton fallend* ist, denn für je zwei reelle Zahlen x_1 und x_2 ist $f(x_1) = f(x_2)$ und somit ist $f(x_1)$ sowohl kleiner oder gleich als auch größer oder gleich $f(x_2)$.

Um solche pathologischen Fälle auszuschließen, verschärft man die Anforderungen ein wenig und definiert den Begriff der *strengen* Monotonie:

Streng monotone Funktion

Es sei f eine Funktion und I eine Teilmenge des Definitionsbereichs D von f; das kann, muss aber nicht der ganze Definitionsbereich sein.

Man nennt f **streng monoton steigend** auf I, wenn gilt: Wenn immer x_2 größer ist als x_1 (also ‚rechts' von x_1 liegt), dann ist $f(x_2)$ größer $f(x_1)$. In Kurzform:

Ist $x_1 < x_2$, so ist $f(x_1) < f(x_2)$.

Analog nennt man eine Funktion f **streng monoton fallend** auf I, wenn gilt: Wenn immer x_2 größer ist als x_1 (also ‚rechts' von x_1 liegt), dann ist $f(x_2)$ kleiner $f(x_1)$. In Kurzform:

Ist $x_1 < x_2$, so ist $f(x_1) > f(x_2)$.

Ein Funktion heißt **streng monoton**, wenn sie streng monoton steigend oder streng monoton fallend ist.

Diesen Kasten habe ich wahrscheinlich schneller geschrieben als Sie ihn gelesen haben, denn ich konnte mit copy and paste den weiter oben stehenden über einfache Monotonie fast wörtlich übernehmen, lediglich das Wort „streng" war einzusetzen, und an zwei Stellen „größer oder gleich" bzw. „kleiner oder gleich" durch „größer" bzw. „kleiner“ zu ersetzen; womit ich auch gleich darauf hingewiesen hätte, worauf Sie beim Lesen dieses Kastens verschärft achten sollten.

Auf neue Beispiele kann ich an dieser Stelle eigentlich verzichten, denn die obigen Beispiele zur einfachen Monotonie sind sogar welche zur strengen; schauen Sie sich das bitte daraufhin noch einmal genau an, ich warte hier so lange.

Wieder da? Na ja, Sie haben schon Recht, wenn Sie ein wenig unzufrieden sind, ein einziges neues Beispiel sollte ich schon bringen; wie wäre es denn hiermit:

Beispiel 2.5:

$$f : \mathbb{R} \to \mathbb{R}\ , \quad f(x) \ = \ \begin{cases} -x^2, & \text{falls } x < 0 \text{ ist,} \\ x^2, & \text{falls } x \geq 0 \text{ ist.} \end{cases}$$

Bevor ich diese Funktion nun auf (strenge) Monotonie untersuche, sollte ich Ihnen vielleicht noch einmal die Funktion der geschweiften Klammer erklären: Diese dient hier dazu, eine **Fallunterscheidung** zu machen. Für jeden Eingabewert x muss geprüft werden, ob er kleiner als null oder ob er größer oder gleich null ist; im ersten Fall greift die obere Zeile hinter der geschweiften Klammer, und diesem x wird als Funktionswert die Zahl $-x^2$ zugeordnet, wobei das Minuszeichen vor dem x steht und somit nicht mitquadriert wird. Beispielsweise ergibt sich so als Funktionswert von $x = -3$:

$$f(-3) \ = \ -(-3)^2 \ = \ -9\,.$$

Ist der aktuelle Eingabewert von x größer oder gleich null, so muss man zur Ermittlung des Funktionswertes in die untere Zeile der rechten Seite schauen; man erkennt, dass der Funktionswert gerade das Quadrat des Eingabewertes sein soll; beispielsweise ist also $f(4) = 4^2 = 16$.

So weit, so gut, jetzt können wir mit der Funktion ein wenig umgehen, aber ist sie auch monoton?

Ja, das ist sie, sogar streng monoton auf ganz $\mathbb{R}$, und zwar steigend. Um das nachzuweisen, greife ich wieder zwei Punkte x_1 und x_2 mit

$$x_1 \ < \ x_2$$

heraus und vergleiche deren Funktionswerte. Hierzu wiederum unterscheide ich drei Fälle:

1. Fall: x_1 und x_2 sind beide negativ, also $x_1 \ < \ x_2 \ < \ 0$. Zur Ermittlung der Funktionswerte muss ich in beiden Fällen die x-Werte quadrieren und dann ein Minuszeichen davor setzen. Ist aber $x_1 \ < \ x_2 \ < \ 0$, so ist x_1^2 eine größere Zahl als x_2^2; setzt man vor beide nun ein Minuszeichen, so liegt $-x_1^2$ links von $-x_2^2$ und ist somit kleiner. Nicht so ganz klar? Zugegeben, diese negativen Zahlen zu vergleichen ist immer ein wenig unanschaulich, das hatten wir ja weiter oben schon einmal gesehen. Machen wir also ein Beispiel: Ist $x_1 = -5$ und $x_2 = -3$, so ist zweifellos $x_1 \ < \ x_2 \ < \ 0$. x_1^2 ist dann 25 und x_2^2 ist 9, also ist x_1^2 die größere Zahl. Nachdem man aber das Minuszeichen vorne drangesetzt und damit die Funktionswertbildung abgeschlossen hat, gilt $-25 < -9$; somit ist also $f(x_1) < f(x_2)$.

2. Fall: x_1 und x_2 sind beide nicht negativ, also $0 \leq x_1 < x_2$. Na ja, das bedeutet, dass zur Ermittlung der Funktionswerte beide Zahlen gerade quadriert werden müssen, und wenn schon x_1 kleiner ist als x_2, so ist x_1^2 erst recht kleiner als x_2^2. Auch in diesem Fall ist also die strenge Monotonie nachgewiesen.

3. Fall: Geben Sie's zu, Sie haben zumindest einen kleinen Moment lang gedacht, es gäbe keinen weiteren Fall! (Falls nicht, so entschuldige ich mich hiermit in aller Form bei Ihnen und bitte Sie, unverzüglich mit dem Hauptfachstudium Mathematik zu beginnen.) Es gibt aber doch noch einen bisher nicht betrachteten Fall, nämlich den, dass x_1 negativ, x_2 aber nicht negativ ist. Das ist aber nun ganz leicht: Wenn x_1 negativ ist, dann ist auch $f(x_1) = -x_1^2$ negativ, während $f(x_2) = x_2^2$ nicht negativ ist. Insbesondere ist also $f(x_1)$ kleiner als $f(x_2)$.

In jedem der drei Fälle ergab sich somit, dass aus $x_1 < x_2$ stets folgt: $f(x_1) < f(x_2)$. Die Funktion f ist also streng monoton steigend. ■

Die ganze Arbeit muss sich auch gelohnt haben, und die Belohnung kommt in Form des folgenden Satzes, der die Frage, ob eine Funktion umkehrbar ist, auf das recht einfach nachzuprüfende Monotonieverhalten zurückführt:

Strenge Monotonie und Umkehrbarkeit

Es sei $f : D \to W$ eine auf ganz D streng monotone Funktion mit Bildmenge $f(D)$. Dann existiert eine auf $f(D)$ definierte Umkehrfunktion f^{-1} von f, f ist also umkehrbar.

Ist f streng monoton steigend, so ist auch f^{-1} streng monoton steigend, ist f streng monoton fallend, so auch f^{-1}.

Mit diesem schönen Ergebnis (Mathematiker haben manchmal einen etwas eigenen Geschmack, was Schönheit anbetrifft, daran werden Sie sich gewöhnen müssen) beende ich diesen Abschnitt und stelle Ihnen in den nächsten ein paar konkrete Funktionen vor; zuvor allerdings sollten Sie das Gelernte noch ein wenig einüben, und wie der Name schon andeutet, hat man hierfür die Übungsaufgaben erfunden:

Übungsaufgabe 2.5:
Geben Sie die Umkehrfunktionen der Funktionen aus Übungsaufgabe 2.4 an. ■

2.3 Potenz- und Wurzelfunktionen

Ich hatte Ihnen ja weiter oben schon so nebenbei die Normalparabelfunktion $f(x) = x^2$ und die (Quadrat-)Wurzelfunktion $f(x) = \sqrt{x}$ untergejubelt. Beide sind spezielle Fälle der Potenz- bzw. der Wurzelfunktionen, die ich Ihnen in diesem Abschnitt näher bringen will.

Beginnen wir mit den Potenzfunktionen:

Potenzfunktionen

Für jede natürliche Zahl n nennt man die Funktion

$$p_n : \mathbb{R} \to \mathbb{R} \ , \quad p_n(x) \ = x^n$$

Potenzfunktion mit dem Exponenten n oder auch einfach n-te Potenzfunktion.

Bei Licht betrachtet ist das eigentlich nichts Neues für Sie: Bereits im ersten Kapitel wurde das Potenzieren von reellen Zahlen mit beliebigen Exponenten geübt, und daher wissen Sie, dass Sie den Funktionswert der n-ten Potenzfunktion berechnen können, indem Sie den aktuellen Wert von x n-mal mit sich selbst multiplizieren; so ist also beispielsweise

$$p_3(-2) \ = \ -8 \ \text{ und } \ p_4(3) \ = \ 81 \ ,$$

denn $(-2)^3 = -8$ und $3^4 = 81$. Der neue Aspekt ist hier, dass ich das Ganze nun als *Funktion* betrachte und das Augenmerk auf Eigenschaften dieser Funktion wie etwa Monotonie und Umkehrbarkeit legen will.

Die einfachste Potenzfunktion, die diesen Namen eigentlich gar nicht so recht verdient, ist $p_1(x) = x$. Hier wird also gar nicht wirklich potenziert. Der Graph dieser Funktion ist die Winkelhalbierende des Koordinatensystems (siehe Abbildung 2.4), eine ziemlich langweilige Angelegenheit.

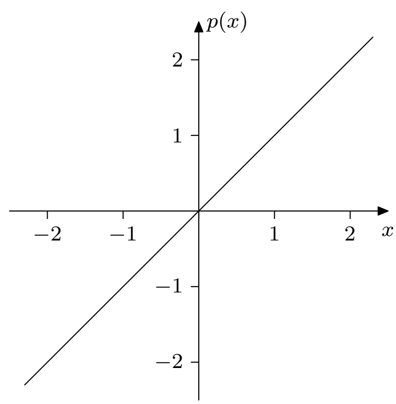

Abb. 2.4 Die Funktion $p_1(x) = x$

Ein anderes Wort für „langweilig“ ist übrigens „monoton“, und da dieser Graph sogar *sehr* langweilig ist, kann ich auch sagen: „streng monoton“. Tatsächlich ist die erste Potenzfunktion auf ganz $\mathbb{R}$ streng monoton steigend und damit, das wissen Sie ja schon, überall umkehrbar. Schlimmer noch: Sie ist ihre eigene Umkehrfunktion, denn

$$p_1(p_1(x)) \;=\; p_1(x) \;=\; x$$

für jede reelle Zahl x. Ich komme auf diese Umkehrbarkeit gleich noch mal zurück, vergessen Sie's nicht!

Die „nächste" Potenzfunktion ist $p_2(x) = x^2$. Diese hatten wir – unter dem Namen Normalparabel – bereits im letzten Abschnitt untersucht und insbesondere gesehen, dass sie *nicht* umkehrbar ist, jedenfalls dann nicht, wenn man sie auf ihrem ganzen Definitionsbereich $\mathbb{R}$ umkehren will.

Abbildung 2.5 zeigt exemplarisch die Graphen einiger Potenzfunktionen. Man erkennt, dass es offenbar zwei Typen von Schaubildern gibt: Ist n gerade, so bewegt sich der Funktionsgraph nur innerhalb des nicht negativen Bereichs und ist symmetrisch zur y-Achse. Ist aber n ungerade, so kommt offenbar jede reelle Zahl als Funktionswert vor und der Funktionsgraph ist symmetrisch zum Nullpunkt.

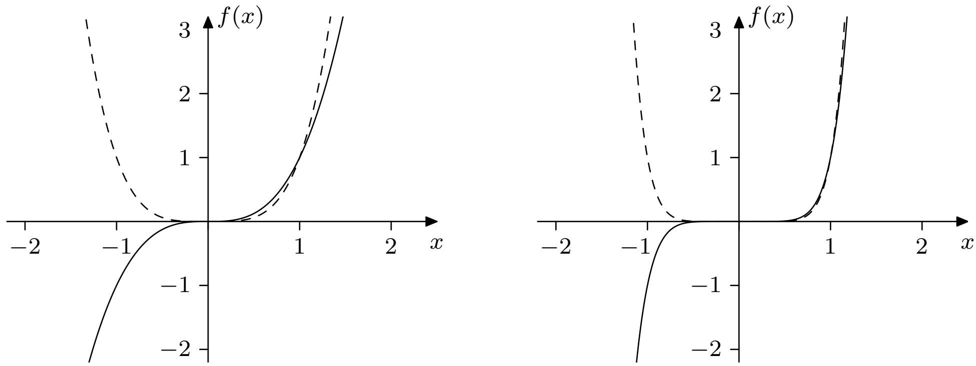

Abb. 2.5 Einige Potenzfunktionen: x^3 (links, schwarz), x^4 (links, gestrichelt), x^7 (rechts, schwarz), x^8 (rechts, gestrichelt)

Es muss jetzt gezeigt werden, dass das nicht nur in den abgebildeten Beispielen, sondern tatsächlich für alle $n \in \mathbb{N}$ gilt.

Zuerst betrachte ich den Fall, dass n eine gerade Zahl ist. Das bedeutet, dass n ohne Rest durch 2 teilbar ist, also den Faktor 2 enthält. Weil Mathematiker etwas so einfach zu Formulierendes nicht einfach so stehen lassen können, sondern es sofort in ihre Formelsprache übersetzen müssen (was sie dann später allerdings dazu befähigt, mit solchen Eigenschaften kurz und prägnant umzugehen), formulieren sie das meist wie folgt:

Wenn n gerade ist, dann kann man es in der Form $n = 2m$ mit einer natürlichen Zahl m schreiben. Beispielsweise ist $14 = 2 \cdot 7$ und $30 = 2 \cdot 15$.

Betrachten wir nun die n-te Potenzfunktion $p_n(x) = x^n$. Wenn n gerade ist, was ich ja zurzeit voraussetze, dann kann ich n durch $2m$ ersetzen, und somit ist die Potenzfunktion darstellbar als

$$p_n(x) \;=\; x^n \;=\; x^{2m}\,.$$

Jetzt kommt die Stelle, wo sich das Herumplagen mit den Potenzregeln im ersten Kapitel endlich mal lohnt. Es wurde dort gezeigt, dass $x^{2m} = (x^2)^m$ ist. Das bedeutet, dass man sich die Auswertung der n-ten Potenzfunktion für *gerades* n wie folgt als Verkettung zweier Funktionen vorstellen kann: Zuerst wird die Funktion x^2 angewandt. Diese Funktion produziert bekanntlich keine negativen Werte, und abgesehen von der Null tritt jeder Funktionswert x^2 genau zwei Mal auf, nämlich als Bild von $+x$ und von $-x$, der Funktionsgraph ist also symmetrisch zur y-Achse. Die anschließende Anwendung der m-ten Potenz (wobei m, wie gesagt, die Hälfte von n ist) ändert an diesen beiden Tatsachen natürlich nichts mehr, denn die Anwendung der m-ten Potenz produziert keine negativen Zahlen aus positiven Eingabewerten, und da sie zwei Mal den Eingabewert x^2 präsentiert bekommt (einmal von $+x$ und einmal von $-x$ herkommend), berechnet sie auch brav zweimal den gleichen Ausgabewert.

Insgesamt ergibt sich also, dass x^n für gerades n einen Graphen besitzt, der symmetrisch zur y-Achse ist und im nicht negativen Bereich verläuft; qualitativ sieht er aus wie derjenige der Normalparabel. Insbesondere ist diese Funktion auf ganz $\mathbb{R}$ nicht umkehrbar.

Da Sie die Argumentationsweise jetzt ja schon kennen und lieben gelernt haben, will ich mich beim verbleibenden Fall, dass n ungerade ist, etwas kürzer fassen (höre ich da etwa ein Aufatmen bei Ihnen?).

Ist n ungerade, so ist $n - 1$, die vor n kommende natürliche Zahl, gerade, somit von der Form $2m$ wie oben gezeigt. Es ist also $n - 1 = 2m$ und damit n von der Form

$$n = 2m + 1 .$$

Die n-te Potenzfunktion kann also für ungerades n wie folgt geschrieben werden:

$$p_n(x) = x^n = x^{2m+1} = x^{2m} \cdot x ,$$

wobei ich im letzten Schritt wieder heimlich still und leise ein Potenzrechengesetzchen benutzt habe.

Aufgrund dieser Darstellung der Funktion $p_n(x)$ kann ich für ungerades n nun wie folgt argumentieren: Wie der Output der Funktion x^{2m} aussieht, haben wir gerade in mühsamer Kleinarbeit ermittelt. Ist nun $n = 2m + 1$, so wird dieser Output noch einmal mit x multipliziert; was dabei passiert, ist aber ziemlich klar: Ist x positiv, so ist auch $x^{2m} \cdot x$ positiv, ist x negativ, so auch $x^{2m} \cdot x$. Und da dieser Ausdruck ja nichts anderes als x^n ist, haben wir somit gezeigt, dass der Graph dieser Funktion für negatives x negativ und für positives x positiv ist. Na ja, und ist $x = 0$, dann ist x^n aber *so was* von null!

Insgesamt erkennt man, dass die Graphen der Funktionen x^n für ungerades n qualitativ alle gleich aussehen; der einfachste und daher ziemlich langweilige Vertreter ist $p_1(x) = x$, und so, wie diese Funktion auf ganz $\mathbb{R}$ umkehrbar ist (erinnern Sie sich noch?), so ist es jede Potenzfunktion für ungerades n.

Wie diese Umkehrfunktion lautet, wissen Sie auch schon spätestens seit dem Lesen des ersten Kapitels, wenngleich dort noch nicht von Funktionen gesprochen wurde:

Um die Wirkung von x^n wieder rückgängig zu machen, brauche ich eine Funktion, nennen wir sie für den Moment $w_n(x)$, die die Eigenschaft

$$w_n(x^n) \;=\; x$$

hat, denn das ist ja gerade die Definition von Umkehrfunktion. Da bietet sich aber nach den Regeln der Potenzrechung die Funktion

$$w_n(x) \;=\; x^{\frac{1}{n}}$$

an, denn es gilt für alle $x \in \mathbb{R}$ und alle ungeraden $n \in \mathbb{N}$:

$$w_n(x^n) \;=\; (x^n)^{\frac{1}{n}} \;=\; x^{n\cdot\frac{1}{n}} \;=\; x$$

und genau danach hatte ich gesucht.

n-te Wurzelfunktion

Die Funktion $w_n(x) = x^{\frac{1}{n}}$, die man ja auch als $w_n(x) = \sqrt[n]{x}$ schreiben kann, nennt man die **n-te Wurzelfunktion**.

Die n-te Wurzelfunktion ist, da sie ja die Umkehrfunktion der streng monoton steigenden Funktion x^n ist, ebenfalls streng monoton steigend, ihren Graphen erhält man, indem man denjenigen von x^n an der ersten Winkelhalbierenden spiegelt.

Ich möchte Sie noch auf ein kleines Problem im praktischen Umgang mit Wurzeln aus negativen Zahlen hinweisen: zwar ist es nach den obigen Überlegungen völlig in Ordnung, einen Ausdruck wie

$$\sqrt[3]{-8}$$

berechnen zu wollen, denn 3 ist eine ungerade Zahl, aber das hat sich zu manchen Programmierern bis heute noch nicht herumgesprochen, und daher melden manche Softwarepakete bei Eingabe von $\sqrt[3]{-8}$ wegen des negativen Radikanden einen Fehler und weigern sich weiterzumachen. (Die guten Programme berechnen dagegen das korrekte Ergebnis als -2.)

Wenn Ihnen so etwas passiert, dann müssen Sie in die Trickkiste greifen: Da $-8 = (-1)\cdot 8$ und $-1 = (-1)^3$ ist, können Sie streng nach den Regeln der Wurzel- bzw. Potenzrechnung den vom Computer ungeliebten Ausdruck $\sqrt[3]{-8}$ wie folgt sukzessive umformen:

$$\sqrt[3]{-8} \;=\; \sqrt[3]{(-1)\cdot 8} \;=\; \sqrt[3]{(-1)^3\cdot 8} \;=\; \sqrt[3]{(-1)^3}\cdot\sqrt[3]{8} \;=\; (-1)\cdot\sqrt[3]{8} \;=\; (-1)\cdot 2 \;=\; -2\,,$$

denn die dritte Wurzel aus der *positiven* Zahl 8 ist nun einmal 2, das bekommt auch das dümmste Computerprogramm noch hin.

Zu schnell gegangen? Schauen Sie sich's noch mal Schritt für Schritt in Ruhe an, es ist jeweils eine ganz einfache Umformung, nur eben mehrere dicht hintereinander;

das wird Ihnen im weiteren Verlauf Ihres Mathematik-Lebens, fürchte ich, noch öfter passieren.

Was ich hier für $n = 3$ und $x = -8$ gemacht habe, geht natürlich für jede ungerade Zahl n und jede negative Zahl x genauso: Die n-te Wurzel $\sqrt[n]{x}$ kann man immer berechnen durch

$$\sqrt[n]{x} \ = \ -\sqrt[n]{|x|}\ .$$

Hierbei ist $|x|$ der Betrag der negativen Zahl x, also ihr „positiver Anteil".

2.4 Polynome und rationale Funktionen

Wenn Sie jetzt glauben, dass ich Sie nach dem Studium der vorangegangenen Kapitel und Abschnitte mit nichts mehr schockieren kann, dann habe ich jetzt vielleicht doch noch eine Überraschung für Sie: Ich werde jetzt ohne lange Vorrede einfach den Begriff „Polynom" definieren; ohne Pizza, ohne Geschwafel, ohne alles; eine wahre Wohltat, nicht wahr?

Polynome

Es sei n eine natürliche Zahl oder null und es seien $a_0, a_1, \dots, a_n$ reelle Zahlen. Eine Funktion $p(x)$, die sich in der Form

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \tag{2.2}$$

darstellen lässt, nennt man ein **Polynom** vom Grad höchstens n.

Die Zahlen $a_0, a_1, \dots, a_n$ heißen die **Koeffizienten** des Polynoms.

Es kann durchaus sein, dass der erste Koeffizient a_n oder, wenn es ganz schlimm kommt, gleich mehrere der ersten Koeffizienten gleich null sind. Beispielsweise ist

$$s(x) \ = \ 0x^5 + 0x^4 + x^3 - x + 1$$

nach obiger Definition ein Polynom vom Grad 5. Natürlich schreibt aber kein Mensch diese führenden Nullen dauernd hin, sondern man schreibt einfach

$$s(x) \ = \ x^3 - x + 1\ ,$$

und so ist aus unserem Polynom fünften Grades eines dritten Grades geworden. Um nicht dauernd irgendwelche derartigen pathologischen Fälle gesondert betrachten zu müssen, hat man das Wörtchen „höchstens" in die Definition eingefügt: $s(x)$ gehört also definitionsgemäß zur Menge der Polynome höchstens fünften Grades, und das stimmt ja auch; dass diese Funktion zwei hohe Potenzen verschenkt, ist ihre eigene Schuld. Will man dagegen betonen, dass ein Polynom einen gewissen Grad, sagen wir n, auch wirklich hat, so sagt man, es sei *vom genauen Grad* n.

Beispiele für Polynome sind:

$$\begin{aligned} p(x) &= 3x^4 - 2x^2 + 5x + 2 \ , \\ q(x) &= -17x^{100} + 1 \ , \\ r(x) &= x^3 + x^2 + x + 1 \ , \end{aligned}$$

die – in dieser Reihenfolge – vom genauen Grad 4, 100 bzw. 3 sind.

Dagegen ist

$$f(x) = 2x^3 - 4x^{\frac{4}{3}} + 1$$

kein Polynom, denn bei einem Polynom dürfen nur natürliche Zahlen im Exponenten auftreten; und $\frac{4}{3}$ kann sich noch so sehr tarnen, es fällt eben doch als Bruch auf, und auch

$$f(x) = 2x^2 - x^{-1}$$

ist *kein* Polynom, da negative Exponenten ebenfalls nicht zugelassen sind.

Apropos „tarnen“: Ein Polynom muss nicht unbedingt in der reinen Form (2.2) hingeschrieben sein, sondern kann sich durchaus zunächst einmal tarnen; wichtig ist, dass es *möglich* ist, es in die Form (2.2) zu bringen.

Beispiel 2.6:
Die Funktion

$$f(x) = (x^2 + 2x)^2$$

hat nicht die in der Definition gewünschte Form eines Polynoms; aber durch Ausmultiplizieren oder durch Anwendung der ersten binomischen Formel stellt sich heraus, dass

$$f(x) = (x^2 + 2x)^2 = (x^2)^2 + 2 \cdot 2x \cdot x^2 + (2x)^2 = x^4 + 4x^3 + 4x^2$$

eben doch ein Polynom ist, an dem es eigentlich nichts auszusetzen gibt. ■

Übungsaufgabe 2.6:
Welche der folgenden Funktionen sind Polynome?

a) $f(x) = x^3 + 3^x$
b) $g(x) = \sqrt{x^2 + 2x + 1}$
c) $h(x) = x^2 - x^{-2}$ ■

An dieser Stelle, da wir uns schon eine ganze Weile so nett über Polynome unterhalten (na ja, zugegeben, in einem solchen Buch ist eine Unterhaltung eher ein Monolog des Autors, aber Sie müssen zugeben, dass ich immerhin versuche, Sie in die Überlegungen einzubeziehen), muss ich doch einmal loswerden, dass ich mich seit 25 Jahren über dieses Wort ärgere. Das liegt daran, dass ich in meiner Jugend einmal sowohl Griechisch als auch Latein gelernt habe; allzu viel ist davon nicht übrig – meine alten Sprachlehrer mögen mir verzeihen – aber es reicht immerhin noch dazu sagen zu können, dass es sich bei dem Wort „Polynom“ um ein eigentlich unmögliches Konglomerat aus diesen beiden Sprachen handelt: ‚polys‘ ($\pi o \lambda \upsilon \sigma$) ist griechisch und heißt ‚viel‘ oder ‚vielfach‘, während der Wortbestandteil ‚nom‘ vom lateinischen Wort ‚nomen‘ kommt, was ‚Ausdruck‘ oder ‚Satzglied‘ bedeutet. Eigentlich müsste man also eher „Multinom“ sagen (weil dann beide Wortbestandteile aus dem Lateinischen stammen), aber ich fürchte, das werden wir wohl alle in diesem Leben nicht mehr ändern. Lassen wir es also bei der Bezeichnung Polynom, und freuen wir uns gemeinsam diebisch, wenn wieder einmal ein Dozent diese Bezeichnung gebraucht, ohne zu wissen, dass sie eigentlich falsch ist. Mein Fremdwörterlexikon erklärt den Begriff Polynom übrigens so: „Ein aus mehr als zwei Gliedern bestehender, durch Plus- und Minuszeichen verbundener mathematischer Ausdruck“. Wenn Sie das besser finden als die oben gegebene mathematische Definition, dann lassen wir es jetzt bleiben und Sie studieren bitte Altphilologie. Sagen Sie den dort tätigen Kollegen aber bitte, dass diese Definition nicht korrekt ist, denn ein Polynom kann durchaus aus nur zwei oder sogar nur einem Glied bestehen.

Wir anderen kommen zurück zur Mathematik: Polynome sind in vielerlei Hinsicht die einfachsten Funktionen, die man sich so vorstellen kann, und werden Ihnen an vielen Stellen Ihres Studiums wieder begegnen. „Einfach“ deswegen, weil Sie ja nur aus Potenzfunktionen zusammengesetzt sind, und diese wiederum, das werden Sie später sehen, lassen sich ganz leicht ableiten, integrieren oder anderweitig mathematisch verarzten. Mit Aussagen zu Polynomen kann man ganze Bücher füllen (und manche tun das auch), aber in diesem Abschnitt, in dem wir die Polynome ja nur beschnuppern wollen, begnüge ich mich damit, Ihnen die Polynome mit kleinen Graden vorzustellen und insbesondere etwas über ihre Nullstellen zu sagen.

Der kleinste für n zugelassene Wert ist sicherlich $n = 0$. Polynome vom Grad null – die einfachsten Polynome also – sind aber eigentlich gar keine so richtigen Polynome, denn es handelt sich um konstante, also nicht von x abhängige Funktionen; ein Polynom vom Grad null ist von der Form

$$p(x) \;=\; a\,,$$

hat also für alle $x \in \mathbb{R}$ denselben Wert, den ich hier einfach a genannt habe. Eine solche Funktion nennt man eine **konstante Funktion**. Das Schaubild einer solchen Funktion ist an Schlichtheit kaum zu überbieten: Er ist einfach eine waagrechte, also parallel zur x-Achse verlaufende, Linie; jedenfalls wenn der Wert a nicht selbst gleich null ist. Ist er das aber, ist also $f(x) = 0$, so ist der Graph von f identisch mit der x-Achse; auch nicht sehr prickelnd.

Viel mehr gibt es über Polynome vom Grad null nicht zu sagen, weswegen ich jetzt auch gleich zu denen vom Grad eins übergehe. Ein Polynom vom Grad eins ist von der Form

$$p(x) \;=\; ax + b\,.$$

Eine solche Funktion nennt man eine **lineare Funktion**. Auch hier habe ich, ähnlich wie bei der konstanten Funktion oben, die Koeffizienten einfach a und b genannt, da es mir bei zweien übertrieben scheint, mit Indizes zu arbeiten.

Den Grund für die Bezeichnung ‚linear' sieht man sofort ein, wenn man beispielsweise das Schaubild einer solchen Funktion in Abbildung 2.6 betrachtet: Es handelt sich um eine Gerade, deren Steigung durch a gegeben wird und die die y-Achse an der Stelle b schneidet, denn für $x = 0$ ergibt sich ja gerade der Wert $p(0) = b$.

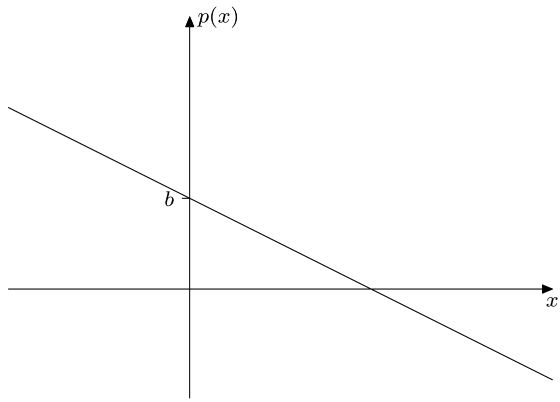

Abb. 2.6 Lineare Funktion mit y-Achsenabschnitt b

Ich will jetzt noch die Nullstellen einer linearen Funktion bestimmen; dazu wäre es vielleicht nicht schlecht, erst mal den Begriff Nullstelle zu definieren. Dieser ist enorm wichtig für die ganze Mathematik, wenn auch nicht so wichtig wie Pizza oder Kartoffelchips, und wird Ihnen mit Sicherheit während des gesamten Studiums immer wieder begegnen. Nun denn:

Nullstelle

Eine reelle Zahl $\overline{x}$ heißt **Nullstelle** einer Funktion $f : \mathbb{R} \to \mathbb{R}$, wenn

$$f(\overline{x}) \;=\; 0$$

gilt.

Beispielsweise ist $\overline{x} = -1$ eine Nullstelle der Funktion

$$f(x) \;=\; (x+1)^3 \;,$$

denn $f(-1) = (-1+1)^3 = 0^3 = 0$.

Es gibt eine Funktion, die besteht überhaupt nur aus Nullstellen, nämlich die konstante Funktion $f(x) = 0$; man bezeichnet sie auch oft als die *Nullfunktion.* Alle anderen konstanten Funktionen haben keine Nullstellen, womit das Nullstellenverhalten dieser Funktionenklasse, also der Polynome vom Grad null, auch schon vollständig geklärt wäre.

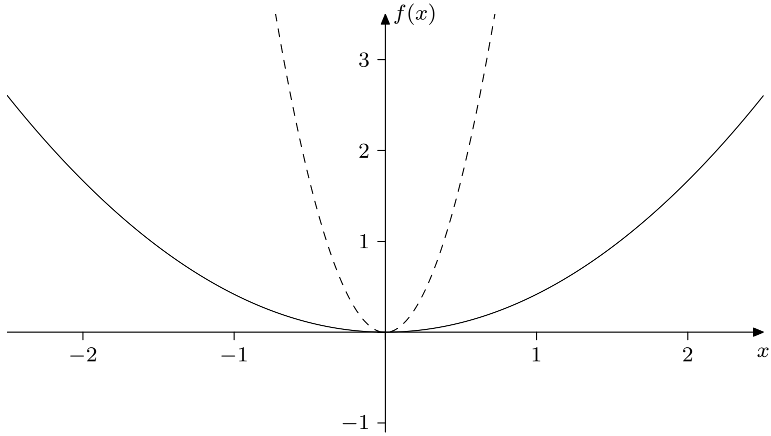

Abb. 2.7 Die Parabeln $\frac{1}{4}x^2$ (schwarz) und $4x^2$ (gestrichelt)

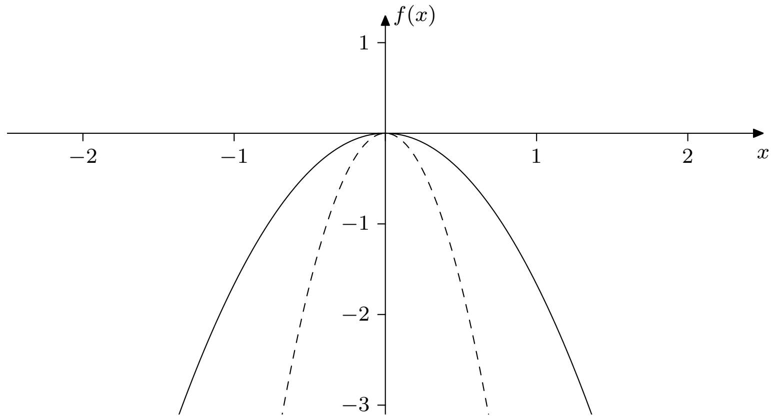

Abb. 2.8 Die Parabeln $-x^2$ (schwarz) und $-4x^2$ (gestrichelt)

Polynome vom *genauen* Grad eins, also lineare Funktionen, haben stets eine Nullstelle, denn eine nicht konstante, also nicht waagrecht verlaufende Gerade schneidet immer irgendwo einmal die x-Achse. Wenn man Lust hat, kann man diese Nullstelle auch ausrechnen, aber da ich gerade keine habe (Lust, nicht Nullstelle), vertröste ich Sie hierfür auf das Kapitel über Gleichungen und Ungleichungen, wo das nachgeholt wird. Für den Moment genügt es festzuhalten, dass ein Polynom höchstens ersten Grades, das nicht gerade die Nullfunktion ist, entweder keine (wenn es konstant ist) oder eine Nullstelle hat.

Wagen wir uns jetzt noch an die Polynome zweiten Grades, die also die Form

$$p(x) \; = \; ax^2 + bx + c \tag{2.3}$$

haben oder zumindest in diese Form gebracht werden können. Eine solche Funktion nennt man auch eine **Parabel**. Den Prototyp dieser Funktionenklasse, die Normalparabel $p(x) = x^2$, hatten wir im Abschnitt über Potenzfunktionen schon kennen gelernt. Worin unterscheidet sich nun eine allgemeine Parabel wie in (2.3) von dieser speziellen? Nun, meine These ist: „Gar nicht! Hast du eine gesehen, hast du alle gesehen.“

Um das zu untermauern, klebe ich an die Normalparabel zunächst einen Vorfaktor a an, betrachte also Funktionen der Form

$$f(x) \; = \; ax^2 \; .$$

Solange a positiv ist, unterscheidet sich diese Funktion von der Normalparabel qualitativ gar nicht, sie wird lediglich etwas breiter (für $a < 1$) oder schlanker (für $a > 1$); wäre übrigens schön, wenn es so etwas für Menschen auch gäbe, denn dank Kartoffelchips und Pizza könnte ich persönlich einen solchen schlank machenden Vorfaktor ab und zu ganz gut gebrauchen. Aber zurück zu Parabeln. Dass das gerade Gesagte stimmt, sieht man am besten an einem Beispiel: Betrachten Sie etwa die Funktion $f(x) = \frac{1}{4}x^2$ (Abbildung 2.7 (schwarz)), so sehen Sie, dass diese für $x = \pm 1$ nicht, wie die Normalparabel, den Wert 1, sondern den Wert $\frac{1}{4}$ hat, den Funktionswert 1 erreicht sie erst für $x = \pm 2$. Und so geht das weiter: Insgesamt kommt diese Funktion sehr viel langsamer hoch als die Normalparabel, was ihr insgesamt ein breiteres Bild verleiht. Anders sieht das beispielsweise bei $f(x) = 4x^2$ in Abbildung 2.7 (gestrichelt) aus: Diese Funktion ist bereits bei $x = \pm 1$ auf 4 angewachsen und bei $x = \pm 2$ ist sie schon bei 16 angekommen. Sie hat daher insgesamt ein schlankeres Aussehen als die Normalparabel.

So viel zu positiven Vorfaktoren; ein *negativer* Vorfaktor spiegelt den Graphen der Parabel mit dem entsprechenden positiven Vorfaktor gerade an der x-Achse, er klappt das Schaubild sozusagen nach unten. Hier sagt ein Bild vermutlich mehr als tausend Worte, daher betrachten Sie bitte Abbildung 2.8.

Damit haben Sie aber im Prinzip auch schon alle Parabeln des allgemeinen Typs gesehen, denn die Koeffizienten b und c in der Darstellung (2.3) bewirken lediglich Verschiebungen des durch den Term ax^2 definierten Graphen; als Beispiel sehen Sie in Abbildung 2.9 den Graphen der Funktion $p(x) = 2x^2 - 3x - 2$.

Die prinzipielle Gestalt einer Parabel kann man also am Koeffizienten von x^2 erkennen: Ist dieser positiv, so sagt man auch, die Parabel sei „nach oben geöffnet“, ist er negativ, so heißt sie „nach unten geöffnet“, und ist er null, so ist die Parabel zu einer Geraden degeneriert, denn dann hat sie ja die Form $bx + c$.

Wie sieht es nun mit den Nullstellen eines Polynoms zweiten Grades, einer Parabel also, aus? Sicherlich gibt es welche ohne Nullstellen, beispielsweise die Funktion $p(x) = x^2 + 1$; da x^2 niemals negativ wird, gibt es keinen Funktionswert, der kleiner als 1 wäre, insbesondere also keinen, der null wird.

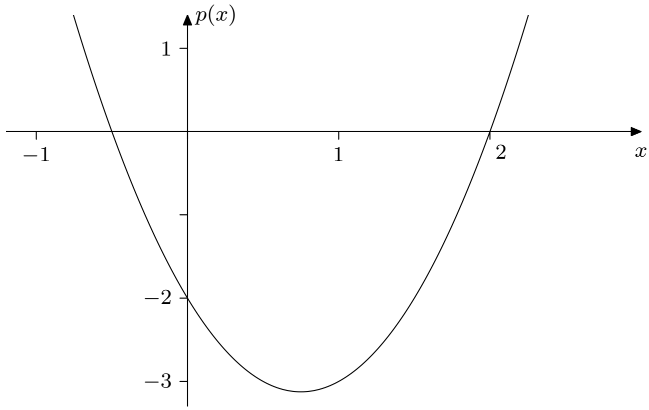

Abb. 2.9 Die Funktion $p(x) = 2x^2 - 3x - 2$

Weiterhin gibt es Parabeln, die genau eine Nullstelle haben, und da brauchen wir gar nicht an den degenerierten Fall der Geraden denken, die gute alte Normalparabel $p(x) = x^2$ tut's auch: Diese hat eine Nullstelle in $x = 0$, und das war es dann auch, alle anderen Funktionswerte sind positiv.

Ziehe ich von der Normalparabel noch 1 ab, lande ich bei der Parabel $p(x) = x^2 - 1$. Diese hat nun zwei Nullstellen, nämlich 1 und -1, wie Sie durch Einsetzen dieser Werte leicht testen können.

Mehr als zwei Nullstellen kann aber keine Parabel der Welt haben, denn wenn Sie sich die Graphen der oben gezeigten Beispiele einmal ansehen, dann erkennen Sie, dass Sie diese so lange umklappen, dehnen oder verschieben können, wie Sie wollen, der Graph wird nie öfter als zweimal die x-Achse schneiden, und das heißt für eine Funktion eben, nicht mehr als zwei Nullstellen zu haben.

Übungsaufgabe 2.7:
Warum muss das Polynom $p(x) = -28x^2 + 117$ zwei Nullstellen haben? ■

Merken Sie etwas? Wenn wir mal den langweiligen Fall der Nullfunktion ausschließen, dann haben wir bisher Folgendes gesehen: Ein Polynom vom Grad null hat keine Nullstelle, ein Polynom vom Grad eins hat höchstens eine Nullstelle, ein Polynom vom Grad zwei hat höchstens zwei Nullstellen. Und ob Sie's glauben oder nicht, auch in der Mathematik geht es manchmal so weiter, wie man sich das wünscht! Soll heißen: Polynome vom Grad drei haben höchstens drei Nullstellen, Polynome vom Grad vier haben höchstens vier Nullstellen, ... ganz allgemein ist Folgendes richtig:

Nullstellen von Polynomen

Für jede natürliche Zahl n gilt: Ein Polynom vom Grad n, das nicht die Nullfunktion ist, hat höchstens n Nullstellen.

Beweisen will ich das hier nicht, irgendetwas müssen Sie im Verlauf Ihres Studiums ja auch noch dazulernen.

Vielleicht haben Sie sich ja schon gefragt, warum ich die ganze Zeit über nichts zum Definitionsbereich der Polynome gesagt habe; nun, das liegt daran, dass Polynome jede beliebige reelle Zahl als Eingabewert akzeptieren und man somit bei der Festlegung des Definitionsbereichs nicht weiter aufpassen muss.

Anders sieht das aus bei der Funktionenklasse, die ich Ihnen jetzt zumindest kurz vorstellen will, den rationalen Funktionen:

Rationale Funktion

Es seien p und q zwei Polynome und D eine Teilmenge der reellen Zahlen, die keine Nullstelle von q enthält.

Man nennt eine Funktion der Form

$$r : D \to \mathbb{R}\ , \quad r(x) \;=\; \frac{p(x)}{q(x)}$$

eine **rationale Funktion**.

Nicht nur der Name, sondern auch die Vorgehensweise erinnert an die Konstruktion der rationalen Zahlen im ersten Kapitel: Diese sind als Quotienten zweier ganzer Zahlen definiert, wobei die Zahl im Nenner nicht null sein darf, während die rationalen Funktionen als Quotienten von zwei Polynomen definiert sind, wobei das Polynom im Nenner nicht null werden darf, weshalb ich ja seine Nullstellen auch aus dem Definitionsbereich der Funktion verbannt habe.

Schauen wir uns erst mal Beispiele an: Die Funktionen

$$r_1(x) \;=\; \frac{3x^3 - 4x^2 + 2}{x^2 + 1}\ , \quad r_2(x) \;=\; \frac{17x^{100} - 2x^{33} + 2x}{x - 1}$$

und

$$r_3(x) \;=\; \frac{1}{x^3 - 2x^2 - x + 2}$$

habe alle drei die richtige Gestalt, stellen also rationale Funktionen dar, denn im Zähler wie auch im Nenner finden sich nur Polynome; allerdings müssen wir uns noch über ihren Definitionsbereich Gedanken machen, denn ohne die Angabe dieses ist die Definition einer Funktion niemals vollständig.

Zuvor jedoch zeige ich Ihnen zwei Beispiele von Funktionsvorschriften, die *keine* rationale Funktion darstellen: Die Funktion

$$g(x) \;=\; \frac{3x^3 - 4x - 2}{x^2 - 3x^{1/2} + 4}$$

ist keine rationale Funktion, da sich im Nenner der Exponent $\frac{1}{2}$ eingeschlichen hat und die Funktion im Nenner somit kein Polynom darstellt. Aber auch ein Kandidat wie

$$h(x) = \frac{3^x - x^3}{x^2 + x - 1}$$

fällt durch, denn der Ausdruck 3^x bewirkt, dass der Zähler kein Polynom ist.

Zurück zu unseren positiven Beispielen von oben; zur Vervollständigung der Funktionsdefinition müssen wir noch den Definitionsbereich bestimmen. Bei rationalen Funktionen geht man übrigens meistens so vor, dass man zunächst die Funktionsvorschrift definiert und sich dann über den Definitionsbereich Klarheit verschafft, mit anderen Worten: die Nullstellen des Nenners untersucht; wenn es Ihnen recht ist, fange ich damit schon mal an.

Der Nenner der ersten rationalen Funktion lautet $x^2 + 1$, und Sie wissen schon, dass dieser niemals null wird, die Funktion $x^2 + 1$ also keine Nullstelle besitzt: Der Ausdruck x^2 wird niemals negativ, das hatten wir jetzt ja schon öfter gesehen, und wenn ich darauf noch 1 addiere, wird das Ganze positiv, also insbesondere nicht null. Der Definitionsbereich der Funktion $r_1(x)$ ist also ganz $\mathbb{R}$.

Der Nenner der Funktion $r_2(x)$ lautet $x-1$, ist also eine lineare Funktion; nun habe ich mich zwar oben erfolgreich darum gedrückt, bereits in diesem Kapitel eine Formel für die Nullstellen linearer Funktionen anzugeben, aber dass diese einfache Funktion die Zahl $x = 1$ als Nullstelle hat, sieht man, denke ich, mit bloßem Auge. Weitere Nullstellen kann sie als Polynom ersten Grades nicht haben, somit ist der größtmögliche Definitionsbereich der Funktion $r_2(x)$ gleich der Menge der reellen Zahlen ohne die Eins; das notiert man in der Kurzform $\mathbb{R} \setminus \{1\}$.

Die dritte Funktion habe ich eigentlich nur hingeschrieben, um Ihnen einen Eindruck davon zu geben, dass das Erkennen der Nullstellen des Nenners und damit der Definitionslücken der rationalen Funktion nicht immer ein reines Vergnügen ist; na ja, soweit Mathematik für Sie überhaupt jemals ein Vergnügen ist, aber auf *die* Debatte lasse ich mich jetzt nicht ein!

An dieser Stelle möchte ich Sie mit drei Statements konfrontieren, von denen Sie die ersten beiden leicht nachprüfen können, und das dritte werden Sie mir ohnehin glauben:

Statement 1: Der Nenner der Funktion $r_3(x)$, also die Funktion x^3-2x^2-x+2, hat die Nullstellen $-1, 1$, und 2; das können Sie durch Einsetzen sofort nachprüfen, worum ich Sie hiermit auch herzlich bitten möchte, da Sie mir ja niemals etwas ungeprüft glauben sollten.

Statement 2: Weitere Nullstellen hat diese Funktion nicht; das wissen Sie aber eigentlich schon, denn der Nenner ist ein Polynom vom Grad drei und ein solches kann nach dem oben Gesagten nicht mehr als drei Nullstellen haben.

Statement 3: Wenn ich mir diese Funktion nicht selbst ausgedacht und den Nenner nach einer Methode, die ich momentan noch in der „Dirty-Tricks“-Kiste für Lehrbuchautoren stecken lassen möchte, konstruiert hätte, dann wäre ich nicht so leicht auf diese drei Nullstellen gekommen; das glauben Sie mir sicherlich, und ich sage es hier nur, um Sie wieder aufzubauen, wenn Sie jetzt verzweifelt denken: „Da wäre ich nie drauf gekommen!“ Wie gesagt, mit diesem Schicksal sind wir schon zu zweit.

Insgesamt haben wir jetzt mit gemeinsamen Kräften herausbekommen, dass der größtmögliche Definitionsbereich der Funktion $r_3(x)$ gleich der Menge $\mathbb{R} \setminus \{-1, 1, 2\}$ ist.

Übungsaufgabe 2.8:
Prüfen Sie, welche der folgenden Funktionen rationale Funktionen sind und geben Sie ggf. den größtmöglichen Definitionsbereich an.

$$\begin{aligned} f(x) &= \frac{2x^2 + x^x}{3x^2 - x + 1} \\ g(x) &= -\frac{x^3 + 5x - 3}{(x-1)^4} \end{aligned} \qquad \blacksquare$$

Wie Sie sich vielleicht erinnern, oder wie Sie sich auch ganz schnell wieder neu überlegen können, kann man ganze Zahlen – wenn man's unbedingt kompliziert haben will – auch als rationale Zahlen mit dem Nenner eins auffassen. Nicht anders ist das bei den Polynomen und rationalen Funktionen: Man kann jedes Polynom als rationale Funktion mit Nenner eins auffassen, beispielsweise ist

$$3x^4 - 3x^3 + 5x^2 - 2 = \frac{3x^4 - 3x^3 + 5x^2 - 2}{1} \; .$$

Ziemlich albern, da haben Sie Recht, aber in diesem Sinne sind Polynome eben die einfachsten rationalen Funktionen. Wenn wir diese links liegen lassen wollen, dann müssen wir also mindestens ein x im Nenner zulassen. Die einfachste „richtige" rationale Funktion ist dann also

$$h(x) = \frac{1}{x} \; .$$

Diese Funktion, genauer gesagt ihr Schaubild, nennt man auch **Hyperbel** oder **Normalhyperbel**. Sie ist offenbar für alle reellen Zahlen außer der Null definiert und bildet jede relle Zahl auf ihren Kehrwert ab.

Wenigstens einmal sollte ich Ihnen wohl ein Schaubild einer rationalen Funktion zeigen und das tue ich jetzt mit der Hyperbel.

Sie erkennen, dass das jetzt ganz anders aussieht als bei den Polynomen, und das liegt eben daran, dass bei rationalen Funktionen irgendwo auf der reellen Achse Nullstellen des Nenners liegen, die man aus dem Definitionsbereich verbannen muss. Tritt nicht der ziemlich unwahrscheinliche Fall ein, dass die Nullstelle des Nenners gleichzeitig auch eine des Zählers ist, so nennt man diese Stelle einen Pol der Funktion.

Polstelle einer rationalen Funktion

Eine reelle Zahl $\overline{x}$, die Nullstelle des Nenners, aber nicht des Zählers einer gegebenen rationalen Funktion $r(x)$ ist, nennt man **Pol** oder auch **Polstelle** von $r(x)$. Eine Polstelle gehört also *nicht* zum Definitionsbereich der rationalen Funktion.

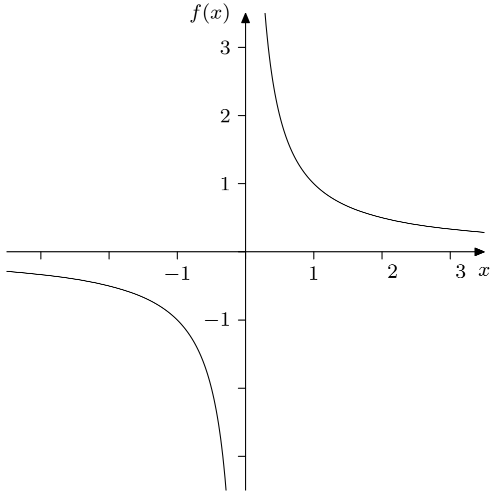

Abb. 2.10 Die Hyperbel $\dfrac{1}{x}$

Nullstellen des Nenners einer rationalen Funktion sollte man meiden wie der Teufel das Weihwasser, denn schon wenn man in ihre Nähe kommt, wird der Funktionsverlauf sehr unangenehm. Am Schaubild der Hyperbel mit der einzigen Polstelle $\overline{x} = 0$ sieht man, was in einem solchen Fall passiert: Die Funktionswerte werden in der Nähe der Polstelle unglaublich groß und können nicht mehr eingezeichnet werden. Zugegeben, „unglaublich groß" ist kein streng mathematischer Begriff, aber Sie werden's mir verzeihen, wenn ich Ihnen die exakte Definition von „über alle Grenzen wachsend" vorenthalte.

Wichtiger ist es, sich darüber klar zu werden, ob die Funktion in den positiven Bereich (also nach oben) oder in den negativen Bereich (also nach unten) strebt. Bei der abgebildeten Hyperbel (Abbildung 2.10) strebt der Funktionsgraph nach oben, wenn man sich von rechts der Polstelle nähert, und nach unten, wenn man dies von links tut. Um das zu verstehen, pirsche ich mich einmal zahlenmäßig von rechts an die Null heran und setze nacheinander die x-Werte 1, $\frac{1}{10}$, $\frac{1}{100}$ und $\frac{1}{1000}$ ein; die entsprechenden Funktionswerte sind 1, 10, 100 und 1000, und man glaubt recht schnell, dass die Funktion hier in den positiven Bereich, also nach oben, entschwindet. Nähert man sich dagegen von links der Null an, setzt also beispielsweise nacheinander die x-Werte -1, $-\frac{1}{10}$, $-\frac{1}{100}$ und $-\frac{1}{1000}$ ein, so erhält man die Funktionswerte -1, -10, -100 und -1000, die Funktion wird also ziemlich schnell *sehr* negativ (was in etwa von der gleichen sprachlichen Qualität ist wie „ziemlich schwanger", aber Sie wissen schon, was ich meine).

Bei den Potenzfunktionen x^n hatten wir gesehen, dass deren Schaubilder prinzipiell in zwei Sorten zerfallen, nämlich die für gerades n und die für ungerades n. Nicht anders ist es bei den Funktionen $\frac{1}{x^n}$, deren ersten Vertreter, die Normalhyperbel, wir gerade untersucht hatten.

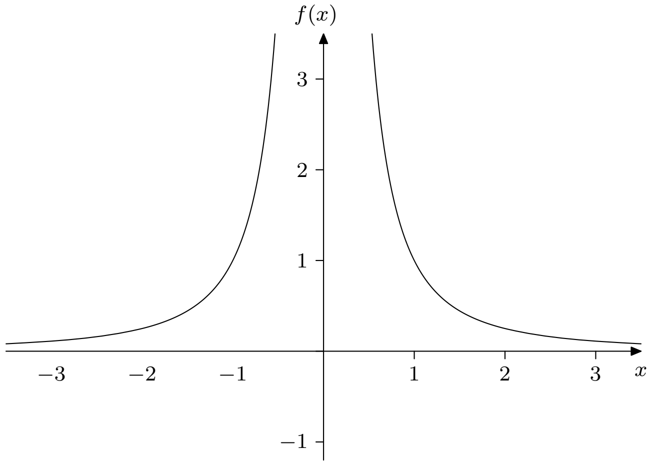

Abb. 2.11 Die Funktion $\dfrac{1}{x^2}$

Wenn Sie sich das Schaubild der Funktion $\frac{1}{x^2}$ in Abbildung 2.11 anschauen, dann sehen Sie einen fundamentalen Unterschied zur Hyperbel, denn dieses neue Schaubild verläuft nur im positiven Bereich, wächst also insbesondere bei Annäherung an die Polstelle null sowohl von links als auch von rechts in den unglaublich großen *positiven Bereich* hinein. Das ist aber nicht weiter verwunderlich, denn dass der Nenner x^2 immer positiv ist, wissen wir ja nun schon lange, und die Kehrwertbildung vermag daran auch nichts zu ändern.

Damit haben Sie aber im Prinzip schon alle Funktionen des Typs $\frac{1}{x^n}$ gesehen, denn qualitativ sehen die alle so aus wie $\frac{1}{x}$, wenn n ungerade, und wie $\frac{1}{x^2}$, wenn n gerade ist. Das ist nicht anders als bei den Potenzfunktionen: Hast du eine gesehen, hast du alle gesehen!

Im nächsten Kapitel befasse ich mich mit Exponentialfunktionen, das sind Funktionen, die sowohl für Kernphysiker als auch für Kapitalanleger von größtem Interesse sind. – Womit ich zwar vermutlich nur 10 % der Leserschaft dieses Buches direkt angesprochen, aber die restlichen 90 % hoffentlich immerhin neugierig gemacht habe.

Zuvor aber ist es meine Pflicht (na ja, ein wenig Kür ist auch dabei), Sie noch mal ein wenig zu verwirren: Ich muss noch erwähnen, dass die hier von mir als Polynome bzw. rationale Funktionen bezeichneten Funktionenklassen gelegentlich, vor allem in der etwas älteren Literatur, auch anders bezeichnet werden. Dort nennt man dann die Polynome „ganzrationale Funktionen“, während man die rationalen Funktionen dann als „gebrochen rationale Funktionen“ bezeichnet.

Ich will darauf nicht näher eingehen und werde diese Bezeichnungen auch nicht weiter benutzen, wollte sie Ihnen aber doch zumindest einmal nennen, damit Sie nicht, sollte Ihnen beispielsweise der Name gebrochen rationale Funktionen einmal im Studium begegnen, verärgert auf dieses Buch schauen und denken: „Na ja, *das* hätte er aber schon erklären können!“ Hat er gerade getan.

2.5 Exponential- und Logarithmusfunktionen

Nehmen Sie einmal an, Sie hätten einen reichen Gönner, der Ihnen etwas Gutes tun will und folgendes Angebot macht: „Also, du kannst wählen: Entweder ich schenke dir auf der Stelle eine Million Euro oder aber ich zahle dir an jedem Tag des nächsten Monats eine gewisse Summe aus; am ersten Tag zwei Euro, am zweiten Tag vier Euro, am dritten Tag acht Euro, am vierten Tag sechzehn Euro und so weiter, also an jedem Tag den doppelten Betrag des Vortages, bis zum dreißigsten Tag des Monats."

Machen Sie jetzt bloß keinen Fehler! Lesen Sie noch ein paar Zeilen weiter, bevor Sie sich entscheiden!

Im Vergleich zu dem Angebot, eine ganze Million bar auf die Tatze zu bekommen, sieht natürlich die zweite Möglichkeit auf den ersten Blick kümmerlich aus: Nach vier Tagen haben Sie insgesamt gerade mal $2 + 4 + 8 + 16 = 30$ Euro in der Hand, und auch wenn Sie eine Woche warten, nennen Sie nur $2+4+8+16+32+64+128 = 254$ Euro Ihr eigen.

Aber keine Angst, es wird besser! Um das zu untermauern, überlege ich jetzt erst mal in Ruhe, wie man den Auszahlungsbetrag an einem beliebigen Tag des Monats berechnen kann. Am ersten Tag sollten es 2 Euro sein, am zweiten das Doppelte, also $2 \cdot 2 = 4$, am Tag darauf wiederum das Doppelte des Vortags, also $2 \cdot 2 \cdot 2 = 8$, tags darauf dann $2 \cdot 2 \cdot 2 \cdot 2 = 16$ usw.

Erinnert Sie das an etwas? Genau um diese lästige Schreibweise der vielen Multiplikationspünktchen loszuwerden, wurde im ersten Kapitel die Potenzschreibweise eingeführt. Mit deren Hilfe kann ich nun ganz einfach formulieren: Die Auszahlung am n-ten Tag beträgt 2^n Euro. Berechnen Sie hiermit die Auszahlung am fünfzehnten Tag, also zur Mitte des Monats, so erhalten Sie immerhin schon $2^{15} = 32.768$ Euro. Auch noch nicht sehr beeindruckend? Nun ja, dann wagen wir jetzt gleich mal den Sprung ans Monatsende: Am dreißigsten Tag erhalten Sie

$$2^{30} = 1.037.741.824 \text{ Euro},$$

also mehr als eine Milliarde und somit mehr als das Tausendfache der als Alternative angebotenen Million. Angesichts solcher Summen kann ich fast großzügig darauf verzichten zu erwähnen, dass Sie ja in unserem Modell an den Tagen zuvor auch schon eine ganz Menge Geld abgestaubt haben, beispielsweise am 29. Tag die Hälfte der gerade berechneten Summe, also mehr als eine halbe Milliarde Euro.

Was hier zum Tragen kommt, ist das fast unvorstellbar schnelle Anwachsen der Exponentialfunktion, in diesem Fall derjenigen zur Basis 2. Wenn Sie nun zögern, weil Sie weit und breit keine Funktion entdecken können, so liegt das wohl daran, dass ich oben die Variable mit n bezeichnet, also 2^n geschrieben habe, und man n eben stets mit einer natürlichen Zahl identifiziert. In der Schreibweise

$$f : \mathbb{R} \to \mathbb{R} \,, \quad f(x) = 2^x$$

sieht das Ganze schon eher nach einer Funktion aus.

Tatsächlich kann man hier für x jede reelle Zahl einsetzen und erhält einen vernünftigen Funktionswert.

Für den Fall, dass x eine rationale, insbesondere also eine ganze Zahl ist, haben Sie im ersten Kapitel gelernt, wie man den Funktionswert 2^x berechnet. Beispielsweise ist $f(\frac{1}{2}) = 2^{1/2} = \sqrt{2}$ und $f(\frac{4}{3}) = 2^{4/3} = \sqrt[3]{2^4} = \sqrt[3]{16}$.

Was aber, wenn x eine richtige reelle Zahl, also keine rationale Zahl ist? Was, um alles in der Welt, soll beispielsweise $f(\sqrt{2}) = 2^{\sqrt{2}}$ sein? Nun, auch dieses Problem haben die Mathematiker im Laufe der Jahrhunderte gelöst, und zwar so, wie das Anna und Otto Normalverbraucher auch tun würden: Man nimmt ein paar (genau genommen unendlich viele, aber lassen wir das) rationale Zahlen $r_1, r_2, r_3, \ldots$ her, die sich immer mehr der Zahl $\sqrt{2}$ annähern, und berechnet die zugehörigen Werte $2^{r_1}, 2^{r_2}, 2^{r_3} \ldots$ Sobald man das Gefühl hat, es ändert sich nicht mehr viel, stoppt man und nimmt den zuletzt berechneten Wert als Näherungswert für $2^{\sqrt{2}}$. (Bevor Sie den Glauben an die Exaktheit der Mathematik ganz verlieren: Natürlich macht man das eigentlich exakter, man benutzt eine so genannte gegen $\sqrt{2}$ konvergente Folge $\{r_i\}$ rationaler Zahlen und bestimmt den Grenzwert der daraus entstehenden Folge $\{2^{r_i}\}$, aber für den Hausgebrauch tut's die oben beschriebene Version durchaus.)

Nun brauchen Sie keine Angst zu haben (die sollten Sie in der Mathematik übrigens niemals haben), dass Sie das tatsächlich selbst durchführen müssten: Wenn Sie jemals das Problem haben sollten, $2^{\sqrt{2}}$ zu berechnen, dann verlassen Sie sich auf die Leute, die Ihren Taschenrechner programmiert haben, und benutzen Sie diesen zur Berechnung. Ich mache das genau so und erhalte

$$2^{\sqrt{2}} = 2{,}66514414\ldots$$

Was ich hier die ganze Zeit mit der Zahl 2 als Basis gemacht habe, kann man natürlich auch mit jeder anderen positiven reellen Zahl machen. Der eingangs als Beispiel zitierte reiche Gönner könnte ja auch sagen, er verdreifacht jeden Tag die Auszahlungssumme, was zur Funktion 3^x führen würde, oder er könnte sich entscheiden, die Summe Tag für Tag zu halbieren, was durch die Funktion $(\frac{1}{2})^x$ modelliert würde. Im letztgenannten Fall würde ich an Ihrer Stelle allerdings die als Alternative genannte Million annehmen, und wenn Sie mal ein paar Werte dieser Funktion berechnen, dann werden Sie mir zustimmen.

Jetzt wird's aber höchste Zeit für eine exakte Definition:

Exponentialfunktion zu allgemeiner Basis

Es sei a eine positive reelle Zahl. Die Funktion

$$\exp_a : \mathbb{R} \to \mathbb{R}\ , \quad \exp_a(x) = a^x$$

heißt **Exponentialfunktion** oder genauer **Exponentialfunktion zur Basis *a***.

Und noch für etwas anderes wird es höchste Zeit:

Übungsaufgabe 2.9:
Berechnen Sie für $a = \frac{1}{2}, 1, 2, 4$ die Funktionswerte von $\exp_a(x)$ an den Stellen $x = -2, 0, 1, 2, 10$. ■

Wenn Sie diese Aufgabe bearbeitet haben, werden Sie einen signifikanten Unterschied beispielsweise im Verhalten von $\exp_{\frac{1}{2}}(x)$ und $\exp_2(x)$ festgestellt haben: Die eine fällt, die andere wächst, und die Funktion $\exp_1(x) = 1^x$ wiederum überzeugt durch Langeweile, denn sie ist konstant gleich 1, da eben 1^x für alle reellen Zahlen x gleich 1 ist.

Tatsächlich ist die Zahl 1 als Basis eine Grenze zwischen den beiden grundsätzlich verschiedenen Typen von Exponentialfunktionen:

Verhalten der Exponentialfunktion zur Basis a

Die Funktion $\exp_a(x) = a^x$ ist auf ganz $\mathbb{R}$ streng monoton wachsend, falls $a > 1$, und auf ganz $\mathbb{R}$ streng monoton fallend, falls $a < 1$ ist. In jedem Fall sind die Funktionswerte stets positive Zahlen, d. h., die Bildmenge der Exponentialfunktion ist die Menge der positiven reellen Zahlen $\mathbb{R}^+$.

Das ist nicht weiter verwunderlich, denn wenn ich eine Zahl, die größer als 1 ist, potenziere, also ein paarmal mit sich selbst multipliziere, dann wird sie natürlich immer größer. Wenn ich dagegen eine Zahl, die zwar positiv, aber kleiner als 1 ist, potenziere, dann verkleinere ich sie ja, und zwar umso mehr, je höher diese Potenz ist; falls Sie das nicht ohnehin schon getan haben, probieren Sie es doch mit $a = \frac{1}{2}$ einfach einmal aus.

Aus den Rechenregeln für das einfache Potenzieren, die Sie im ersten Kapitel noch mal nachlesen können, leitet man direkt die folgenden Gesetzmäßigkeiten ab:

Rechenregeln für die Exponentialfunktion

Für alle reellen Zahlen x und y und für jede positive Basis a gelten folgende Rechenregeln:
1) $\exp_a(x + y) \;=\; \exp_a(x) \cdot \exp_a(y)$
2) $\exp_a(-x) \;=\; \exp_{1/a}(x)$.

Falls Sie die zweite Regel erst einmal umgehauen hat, will ich Sie kurz wieder aufbauen: Das ist nur eine kompakte (und daher auf den ersten Blick nicht sofort einsichtige) Schreibweise für die nach den Potenzgesetzen geltende Gleichung

$$a^{-x} \;==\; \frac{1}{a^x} \;=\; \left(\frac{1}{a}\right)^x \; .$$

So einfach kann Mathematik sein.

Ich hatte Ihnen im Vorspann zu diesem Kapitel ja versprochen zu zeigen, dass die Exponentialfunktion sowohl für Kernphysiker als auch für Kapitalanleger von Interesse ist, und dieses Versprechen will ich jetzt einlösen. Da sich unter der Leserschaft dieses Buches – spätestens nach erfolgreichem Studienabschluss – vermutlich mehr Kapitalanleger als Kernphysiker befinden, beginne ich mit einem Beispiel für Erstere:

Beispiel 2.7:
Nehmen Sie an, Sie legen ein gewisses Kapital K, es muss ja nicht gleich die oben erwähnte Milliarde sein, zu einem jährlichen Zinssatz von p Prozent an. Nach einem Jahr ist Ihr Kapital dann auf

$$K + \frac{p}{100} \cdot K \ = \ K \cdot \left(1 + \frac{p}{100}\right)$$

angewachsen; beispielsweise erhöht sich ein Kapital von 100 Euro bei einem Zinssatz von 3 Prozent auf $100 \cdot (1 + \frac{3}{100}) = 100 \cdot 1{,}03 = 103$ Euro.

Lassen Sie – nun wieder im allgemeinen Fall – die Zinsen auf dem Konto stehen, so werden diese im nächsten Jahr ebenfalls verzinst und am Ende des zweiten Jahres schlagen

$$K \cdot \left(1 + \frac{p}{100}\right) \cdot \left(1 + \frac{p}{100}\right) \ = \ K \cdot \left(1 + \frac{p}{100}\right)^2$$

zu Buche. Und so geht das weiter: Am Ende des n-ten Jahres, wobei n irgendeine natürliche Zahl sein soll, ist das Anfangskapital K auf

$$K \cdot \left(1 + \frac{p}{100}\right)^n$$

angewachsen. Und nun sehen Sie auch, was das mit unserem Thema zu tun hat: Der Faktor, mit dem das Ausgangskapital multipliziert wird, ist gerade der Funktionswert der Exponentialfunktion zur Basis $a = \left(1 + \frac{p}{100}\right)$ an der Stelle n. Und genau wie bei jeder richtigen Funktion muss man sich auch hier nicht auf die natürlichen Zahlen als Definitionsmenge beschränken, sondern darf jede positive reelle Zahl anstelle von n einsetzen. Praktische Bedeutung hat das durchaus, denn wenn Sie oder Ihre Bank wissen wollen, was für einen Wert Ihr Kapital an irgendeinem Tag mitten im Jahr hat (beispielsweise weil Sie Ihr Geld abheben wollen), dann müssen Sie eben nicht ganzzahlige Werte einsetzen; beispielsweise ist der Kapitalwert nach zwei Jahren und fünf Monaten, also nach $\frac{29}{12}$ Jahren, gleich

$$K \cdot \left(1 + \frac{p}{100}\right)^{\frac{29}{12}} \quad .$$

Im Prinzip geht das auch für nicht rationale Exponenten; fragen Sie Ihren Bankberater doch einmal, was für einen Wert Ihr Kapital nach $\sqrt{2} \approx 1{,}4142...$ Jahren hat, aber nehmen Sie bitte Ihren Fotoapparat mit und schicken Sie mir ein Bild seines Gesichtsausdrucks. ■

Übungsaufgabe 2.10:
Frau Pleite möchte 8 000, – Euro optimal anlegen. Bank A bietet ihr an, das Geld mit 4,5 % jährlich zu verzinsen. Bank B bietet ihr an, das Geld vier Jahre lang festzulegen; am Ende des vierten Jahres erhält sie 9 431, 07 Euro ausbezahlt. Welche Bank sollte Frau Pleite wählen, wenn sie am Ende des vierten Jahres ein maximales Vermögen haben will? ■

Beispiel 2.8:
In der Zwischenzeit wende ich mich an die zweite momentane Zielgruppe, die Kernphysiker. Hier gibt es beim radioaktiven Zerfall das Phänomen der **Halbwertszeit**: Nach einer jeweils genau gleichen Zeitspanne, eben der Halbwertszeit, ist gerade die Hälfte des vorher noch vorhandenen radioaktiven Materials zerfallen. Ist also zu Beginn des Zerfallsprozesses eine Masse M des Materials vorhanden, so ist nach Ablauf einer Halbwertszeitspanne noch die Hälfte, also $\frac{M}{2}$, nach einer weiteren wiederum die Hälfte davon, also $\frac{M}{4}$ übrig. Sie sehen, wie das läuft: Nach n Halbwertszeiten ist noch

$$M \cdot \left(\frac{1}{2}\right)^n$$

Material vorhanden. Auch jetzt ist es also eine Exponentialfunktion, in diesem Falle diejenige zur Basis $\frac{1}{2}$, die den entscheidenden Beitrag zu dieser Massenfunktion liefert. Auch hier kann man wieder nicht ganzzahlige Werte einsetzen und so beispielsweise berechnen, wie viel Material noch nach $2,3451$ Halbwertszeitspannen vorhanden ist, was Kernphysiker im Gegensatz zu Bankberatern allerdings im Allgemeinen keineswegs verblüfft. ■

Übungsaufgabe 2.11:
Radioaktiver Kohlenstoff ^{14}C, den man zur Datierung fossiler Funde benutzt, hat eine Halbwertszeit von etwa 5 776 Jahren. Berechnen Sie, wie viel von einem Gramm ^{14}C nach 10 000 Jahren noch vorhanden ist. ■

Zum Abschluss dieses kleinen Ausflugs in die Welt der Exponentialfunktionen möchte ich Sie nochmals mit einer speziellen Bezeichnung in der Mathematik vertraut machen: Sie werden nämlich sicherlich irgendwann einmal in einem Mathematikbuch dem Begriff „Exponentialfunktion“ ohne weiteren Zusatz, also ohne Angabe der Basis, begegnen. In diesem Fall meint man eine ganz spezielle Exponentialfunktion, nämlich diejenige zur Basis e.

Alles klar? Nein, kann ja nicht sein, ich muss wohl schon sagen, was dieses e bedeuten soll. Es ist innerhalb der Mathematik eine ebenso fest stehende Bezeichnung wie beispielsweise die Notation π für die Kreiszahl. e bezeichnet die **eulersche Zahl**, benannt nach Leonhard Euler, einem der Giganten der Mathematik.

Nach vorsichtigen Schätzungen meinerseits gibt es etwa 117 verschiedene Definitionen bzw. Herleitungen dieser Zahl, und alle führen natürlich auf denselben Wert. Das Problem ist, dass keine einzige dieser Definitionen so kurz und knackig ist, dass ich Sie Ihnen hier zumuten könnte; belassen wir es also bei der Namensgebung und der Angabe des zahlenmäßigen Werts von e; es ist

$$e \;=\; 2{,}718\,281\,828\,459\,045\ldots\ldots\,,$$

wobei die Folgen dieser Nachkommazahlen niemals abbricht, da e keine rationale Zahl ist. Redet oder schreibt also jemand von *der* Exponentialfunktion, so ist damit die Funktion $f(x) = e^x$ mit e wie gerade definiert gemeint.

So weit hatte ich den Text geschrieben, als der Verlag, der per definitionem immer Recht hat, Einspruch einlegte und mich bat, die Definition der Zahl e zumindest ein klein wenig zu motivieren. Da ich der Letzte bin, der etwas gegen Motivation von Mathematik-Interessierten hat, will ich dem gerne nachkommen. Nun denn:

Ein Stück weiter oben sind wir gemeinsam in die Tiefen der Zinseszinsrechnung eingestiegen. Nehmen Sie einmal an, Sie hätten eine Bank gefunden, die Ihnen den sagenhaften Jahreszinssatz von 100 % gewährt. Da Sie angesichts dieses Angebots zurecht misstrauisch sind, legen Sie nur einen Euro an. Am Ende des Jahres erhalten Sie dann

$$K_1 \;=\; \left(1 + \frac{100}{100}\right)^1 \;=\; (1+1)^1 \;=\; 2$$

ausbezahlt.

Im nächsten Schritt zur Verbesserung Ihrer Finanzlage geht die Bank dazu über, Ihren Einsatz am Ende jeden Monats, also 12-mal im Jahr, zu verzinsen, dabei allerdings natürlich immer nur mit 1/12 des Jahreszinses. Am Ende des Jahres erhalten Sie in diesem Fall also

$$K_{12} \;=\; \left(1 + \frac{1}{12}\right)^{12} \;=\; 2{,}613\,03$$

ausbezahlt.

Warum sollte eine Bank, die am Ende jeden Monats verzinst, nicht auch am Ende jeder Woche, also 52-mal im Jahr, verzinsen? Täte sie das, so erhielten Sie am Ende des Jahres

$$K_{52} \;=\; \left(1 + \frac{1}{52}\right)^{52} \;=\; 2{,}692\,59$$

ausbezahlt.

Das Spielchen kennen Sie nun schon: Ich gehe jetzt zur täglichen Verzinsung mit 1/365 des Jahreszinses über: Sie erhielten dann am Ende des Jahres

$$K_{365} \;=\; \left(1 + \frac{1}{365}\right)^{365} \;=\; 2{,}714\,56$$

ausbezahlt.

Und wenn wir schon eine Bank gefunden haben, die verrückt genug ist, Ihr Kapital täglich zu verzinsen, dann finden wir auch eine, die stündlich verzinst, die also

$24 \cdot 365 = 8\,760$ Zinstermine zu jeweils $1/8\,760$ des Jahreszinses zulässt; hier erhalten Sie am Ende des Jahres

$$K_{8\,760} \;=\; \left(1 + \frac{1}{8\,760}\right)^{8\,760} \;=\; 2{,}718\,13$$

ausbezahlt.

Ich bitte Sie nun, die berechneten Zahlenwerte $K_2, \ldots, K_{876\,0}$ nochmals der Reihe nach anzuschauen. Sicherlich fällt Ihnen auf, dass sich diese Werte immer mehr der oben definierten eulerschen Zahl e annähern; tatsächlich kann man beweisen, dass sie für immer kleiner werdende Zinsabstände beliebig nahe an die eulersche Zahl herankommen, und man kann diese Zahl auch als so genannten Grenzwert der Folge K_n für n gegen unendlich definieren – man spricht dann auch von *stetiger Verzinsung.*

Ich habe oben mehr so nebenher aufgeschrieben, dass die Exponentialfunktion für jede positive Basis außer 1 eine streng monotone Funktion ist. Das (nicht die Tatsache, dass ich es nebenher geschrieben habe, sondern die Tatsache, dass es überhaupt da steht) bedeutet aber, dass eine Umkehrfunktion der Exponentialfunktion existiert, die das gleiche Monotonieverhalten hat wie diese.

Diese Umkehrfunktion will ich Ihnen jetzt zum Abschluss dieses Kapitels noch vorstellen, es handelt sich um die **Logarithmusfunktion**. Den meisten Menschen graust es schon allein bei der Nennung dieses Namens, aber ich will versuchen zu vermitteln, dass es sich hierbei um etwas ganz Normales handelt (soweit Mathematik etwas ganz Normales sein kann, aber auf *die* Diskussion lasse ich mich jetzt nicht ein). Vielleicht beginne ich am besten damit, dieses etwas merkwürdige Wort zu erklären: Es ist aus zwei griechischen Wortbestandteilen zusammengesetzt, nämlich *logos* ($\lambda o \gamma o \varsigma$) und *arithmos* ($\alpha \varrho \iota \vartheta \mu o \varsigma$). „Arithmos“ ist das griechische Wort für ‚Zahl‘ und steckt beispielsweise auch in ‚Arithmetik‘ drin; „Logos“ ist ein stark philosophisch angehauchter Begriff und hat sehr viele Bedeutungen, beispielsweise ‚Wort‘, ‚Lehre‘, aber auch ‚Bedeutung‘. „Logarithmus“ könnte man also auf zahlreiche verschieden Arten und Weisen übersetzen, aber da ich Mathematiker bin und nicht Altphilologe, will ich das gar nicht erst versuchen, sondern gehe direkt zur mathematisch-präzisen Definition des Logarithmus über:

Logarithmus zur Basis a

Es seien a und x positive reelle Zahlen und $a \neq 0$. Diejenige reelle Zahl y, die die Gleichung

$$a^y \;=\; x$$

löst, nennt man **Logarithmus** oder auch **Logarithmusfunktion von $\boldsymbol{x}$ zur Basis $\boldsymbol{a}$** und bezeichnet sie mit $\log_a(x)$.

Beispielsweise ist also $\log_2(8) = 3$, denn $2^3 = 8$, und $\log_5(\frac{1}{5}) = -1$, denn $5^{-1} = \frac{1}{5}$. Ein anderes charakteristisches Beispiel ist $\log_4(2) = \frac{1}{2}$, denn $4^{1/2} = 2$. Schließlich will ich noch darauf hinweisen, dass für jede Basis a gilt: $\log_a(1) = 0$, denn für alle positiven reellen Zahlen a gilt $a^0 = 1$.

An schwierigere Dinge traue ich mich nicht heran, das würde ich gerne Ihnen überlassen:

Übungsaufgabe 2.12:
Berechnen Sie ohne Taschenrechner die folgenden Logarithmen:

$$\log_{10}(0,001)\ ,\quad \log_7(\sqrt[4]{7^3}).$$ ■

Es gilt nach obiger Definition also $a^{\log_a(x)} = x$, weswegen man die Logarithmusfunktion mit Fug und Recht als Umkehrung der Exponentialfunktion auffassen kann. Das wiederum impliziert aufgrund der Monotonieeigenschaften der Exponentialfunktion die folgende Feststellung:

Monotonieverhalten der Logarithmusfunktion zur Basis a

Der Definitionsbereich der Logarithmusfunktion $\log_a(x)$ ist die Menge der positiven reellen Zahlen, bezeichnet mit $\mathbb{R}^+$. Die Funktion $\log_a(x)$ ist auf ganz $\mathbb{R}^+$ streng monoton wachsend, falls $a > 1$, und auf ganz $\mathbb{R}^+$ streng monoton fallend, falls $a < 1$ ist.

Nicht nur das Monotonieverhalten, sondern auch die Rechenregeln für die Exponentialfunktion haben direkte Auswirkungen auf die Logarithmusfunktion, denn sie führen direkt auf die folgenden Regeln:

Rechenregeln für die Logarithmusfunktion

Für alle reellen Zahlen x, y und p und für jede positive Basis a gelten folgende Rechenregeln:
1) $\log_a(x \cdot y) \;=\; \log_a(x) + \log_a(y)$
2) $\log_a(\frac{x}{y}) \;=\; \log_a(x) - \log_a(y)$
3) $\log_a(x^p) \;=\; p \cdot \log_a(x)$

„Ziemlich trocken und weltfremd das Ganze“, denken Sie gerade, stimmt’s? Das liegt vermutlich an mir (na ja, an wem sonst?), denn bisher habe ich nur mit Monotonie und Rechenregeln um mich geworfen; aber das wird jetzt anders, jetzt kommen die praktischen Beispiele:

Beispiel 2.9:
Erinnern Sie sich noch an die Kapitalanleger und Kernphysiker? Die können, wie wir oben gesehen haben, sehr gut die Exponentialfunktion gebrauchen; aber auch der Logarithmus ist für beide Berufsgruppen von großem Nutzen. Für die Kapitalanleger

demonstriere ich Ihnen das jetzt einmal, die Kernphysiker überlasse ich dann Ihnen!

Wieder einmal haben Sie ein Kapital K zur Verfügung, das Sie bei einer Bank Ihres Vertrauens zum Zinssatz von 4 Prozent pro Jahr anlegen. Frage: Wann hat sich das Kapital verdoppelt?

Um diese Frage zu beantworten, erinnern wir uns mal kurz gemeinsam an die Formel zur Zinseszinsrechnung, die ich oben im Zusammenhang mit der Exponentialfunktion angegeben hatte: Das Kapital K ist beim Zinssatz von 4 Prozent nach p Jahren auf $K \cdot (1{,}04)^p$ angewachsen, und da wir danach fragen, wann es sich verdoppelt hat, müssen wir also berechnen, wann (also für welches p)

$$(1{,}04)^p \;=\; 2$$

gilt. Das ist aber ein klarer Fall für den Logarithmus, denn nach Definition ist

$$p \;=\; \log_{1{,}04}(2) \;\approx\; 17{,}673\;.$$

Sie müssen also etwa 17 Jahre und acht Monate warten, bis sich Ihr Kapital verdoppelt hat. ■

Den Zahlenwert 17,673 habe ich natürlich mit dem Taschenrechner bestimmt. Bevor Sie jetzt verzweifelt Ihren Taschenrechner nach der Taste für den Logarithmus zur Basis 1,04 absuchen: Meiner hat auch keine solche Taste und die allerwenigsten haben so etwas. Meist sind nur ein oder zwei prominente Logarithmen implementiert, beispielsweise diejenigen zur Basis 2 oder 10, und diese tragen auch besondere Bezeichnungen, die ich Ihnen dann gleich noch verraten will.

Zunächst aber müssen wir das Problem lösen, einen Logarithmus beispielsweise zur Basis 1,04 zu berechnen, wenn unser Taschenrechner nur den zur Basis 2 oder 10 beherrscht. Das geht zum Glück recht einfach, denn man kann mittels der folgenden Formel Logarithmen zwischen zwei Basen (das bezeichnet hier keine Verwandten, sondern ist der Plural von Basis) einfach umrechnen:

Umrechnung von Logarithmen zu verschiedenen Basen

Für beliebige positive und von 1 verschiedene Zahlen a und b und positive Werte x gilt die Umrechnungsformel

$$\log_a(x) \;=\; \frac{\log_b(x)}{\log_b(a)}\;.$$

Um diese Formel auf unser obiges Beispiel anzuwenden, setze ich jetzt $a = 1{,}04$ und b auf einen Wert, zu dem der Taschenrechner einen Logarithmus bereithält, sagen wir $b = 10$. Beachten Sie, dass ich auf der rechten Seite der Umrechnungsformel nur

Zehnerlogarithmen berechnen muss, was nach Annahme mein Taschenrechner auch kann; es ergibt sich

$$\log_{1,04}(2) = \frac{\log_{10}(2)}{\log_{10}(1{,}04)} = \frac{0{,}301029995}{0{,}017033339} \approx 17{,}673\,,$$

also der oben angegebene Wert. Wenn Sie das nachrechnen wollen, müssen Sie evtl. noch mal in die Betriebsanleitung Ihres Taschenrechners schauen, denn die Taste für den Zehnerlogarithmus ist wahrscheinlich nicht mit „$\log_{10}$“ beschriftet, sondern mit „log“ oder „lg“. Warum das so ist, verrate ich Ihnen gleich... – nein, nicht nach der Werbung, sondern nach der nächsten Übungsaufgabe:

Übungsaufgabe 2.13:
Eine radioaktive Substanz hat eine Halbwertszeit von 500 Jahren. Nach wie vielen Jahren ist noch
a) ein Viertel
b) ein Fünftel
des ursprünglich vorhandenen Materials vorhanden? ■

Auch zum Schluss dieses Abschnitts will ich Sie mit ein paar speziellen Bezeichnungen bewaffnen, damit Sie im späteren Verlauf Ihrer Mathematik-Karriere nicht unnötigerweise darüber stolpern, denn für die sehr häufig auftretenden Basen 2, 10 und e (die eulersche Zahl) hat man spezielle Bezeichnungen eingeführt, die Sie vermutlich auch auf Ihrem Taschenrechner so finden: Den Logarithmus zur Basis 2 nennt man **Logarithmus dualis** und bezeichnet ihn mit $\mathrm{ld}(x)$, also

$$\mathrm{ld}(x) = \log_2(x) ,$$

den Logarithmus zur Basis 10 nennt man **dekadischen Logarithmus** und bezeichnet ihn mit $\lg(x)$, also

$$\lg(x) = \log_{10}(x) .$$

Für den wahrscheinlich am häufigsten vorkommenden Logarithmus, denjenigen zur Basis e, hat man zwei Namen gefunden, man nennt ihn den **Logarithmus naturalis** oder den **natürlichen Logarithmus** und bezeichnet ihn mit $\ln(x)$, also

$$\ln(x) = \log_e(x) .$$

Wieder einmal könnte ich mich darüber aufregen, dass man hier einen Mix aus Griechisch (Logarithmus) und Latein (naturalis und dualis) fabriziert hat, aber was soll's, das Kapitel ist erfolgreich zu Ende gebracht, und da sollte die schiere Freude überwiegen. Bei mir ist das auch so, wie geht's Ihnen?

3 Gleichungen und Ungleichungen

Neben dem Begriff der Funktion, mit dem Sie sich im letzten Kapitel herumgeschlagen haben, ist derjenige der Gleichung bei vielen Menschen fast als Synonym für „Mathematik“ zu finden. Man spricht beispielsweise in Politik und Verwaltung von „Gleichungen mit mehreren Unbekannten“, von „Gleichungen, die nicht aufgehen“ und so weiter. Den bei mir obligatorischen Seitenhieb auf Politiker und ihre Fähigkeiten zum mathematischen, also logischen Denken, verkneife ich mir jetzt, wir müssen ja weiterkommen.

Was *ist* denn nun aber eine Gleichung? Nun, zum Beispiel ist $2 + 2 = 4$ eine Gleichung, eine ziemlich übersichtliche sogar. Über solche rede ich hier aber nicht, vielmehr über Gleichungen, die (mindestens) eine Variable enthalten; ein Beispiel für so etwas ist

$$\frac{2x^2}{1+x^3} - 7(x-1)^2 = \frac{\sqrt{3+x}}{2x^5} \tag{3.1}$$

Ganz allgemein kann man den Begriff der Gleichung wie folgt definieren:

Gleichung

Eine Gleichung besteht aus zwei mathematischen Ausdrücken (Termen), die durch ein Gleichheitszeichen verbunden sind.

Enthält mindestens einer der Ausdrücke eine Variable, so sucht man nach **Lösungen der Gleichung**. Dies sind Zahlen, die, wenn man sie anstelle der Variablen einsetzt, bewirken, dass die beiden durch das Gleichheitszeichen verbundenen Terme denselben Wert haben. Die Menge aller Lösungen, die auch leer sein kann, bezeichnet man als **Lösungsmenge** der Gleichung.

Man setzt also überall dort, wo eine Variable steht – oft wird die im Weiteren x heißen –, dieselbe Zahl ein und berechnet die Werte der Ausdrücke rechts und links des Gleichheitszeichens; ergibt sich beide Male derselbe Wert, ist die eingesetzte Zahl eine Lösung der Gleichung. Im obigen Beispiel (3.1) ist übrigens $x = 1$ eine Lösung, denn dies ergibt auf der linken wie auf der rechten Seite des Gleichheitszeichens den Wert 1.

Im Folgenden werde ich Ihnen einige Lösungsverfahren für Gleichungen, also Methoden zur Berechnung ihrer Lösung(en), vorstellen; keine Sorge, dabei befassen wir uns nicht mit solchen Monstern wie (3.1), das diente nur dazu, Sie wachzurütteln.

Fast alle Lösungsverfahren für Gleichungen beruhen darauf, dass es erlaubt ist, Gleichungen nach gewissen Regeln umzuformen, ohne ihre Lösungsmenge zu ändern. Man wendet diese Umformungen dann – möglicherweise in mehreren Schritten – so an, dass man am Ende eine Lösung direkt ablesen kann.

Ich zeige Ihnen das einmal an zwei ganz einfachen Beispielen: Die Gleichung

$$x - 2 = 4$$

soll gelöst werden. Hierzu addiere ich *auf beiden Seiten* der Gleichung 2; das ergibt links $x - 2 + 2 = x$ und rechts $4 + 2 = 6$, also

$$x = 6 .$$

Und schon steht's da. Wenn Sie mir nicht trauen, dann setzen Sie die so ermittelte Lösung in die Ausgangsgleichung ein und sehen, dass Ihr Misstrauen ausnahmsweise einmal unbegründet war.

Als zweites Beispiel nehme ich mir die ebenfalls höchst einfache Gleichung

$$3x = 12$$

vor. Hier multipliziere ich *beide Seiten* der Gleichung mit $\frac{1}{3}$ und erhalte die Lösung

$$x = 4 .$$

Mit diesen beiden Beispielen habe ich Ihnen prinzipiell schon alle erlaubten Umformungen vorgestellt (die man üblicherweise natürlich mehrfach anwendet und somit wesentlich komplexere als die gerade beispielhaft genannten Gleichungen lösen kann); ich schreibe das nun noch mal formal auf.

Umformung von Gleichungen

Die folgenden Umformungen einer Gleichung ändern nicht deren Lösungsmenge und werden **erlaubte Umformungen** genannt:
- Addition bzw. Subtraktion derselben Zahl auf beiden Seiten der Gleichung
- Multiplikation beider Seiten mit derselben von null verschiedenen Zahl

Vielleicht wundern Sie sich darüber, dass die Addition von null erlaubt ist, die Multiplikation mit null aber nicht. Nun, wenn Sie auf beiden Seiten null addieren, passiert ja nichts, aber auch gar nichts mit der Gleichung, und das ändert zweifellos auch nichts an Ihrer Lösungsmenge, ist also erlaubt. Dass es uns bei der Suche nach der Lösung nicht weiterbringt, steht auf einem anderen Blatt.

Wie sieht es mit der Multiplikation mit null aus? Betrachten Sie hierzu mal die Gleichung

$$\sin(x^{2/7} + 2^{x^2}) = 528x^{-3/5} - 3^{1/x} . \tag{3.2}$$

Ich habe nicht die geringste Ahnung, wie die Lösungsmenge dieser Gleichung aussieht, aber wenn ich beide Seiten (verbotenerweise) mit null multipliziere, erhalte ich die Gleichung

$$0 \;=\; 0\,.$$

Das hat wie fast alles im Leben einen Vorteil und einen Nachteil: Der Vorteil ist, dass die solchermaßen umgeformte Gleichung zweifellos wahr ist, der Nachteil, dass es bei der Suche nach der Lösungsmenge von (3.2) keinen Schritt weiterhilft. Und da das bei jeder anderen Gleichung genauso wäre, sehen Sie, dass es eine gute Idee ist, die Multiplikation mit null nicht als erlaubte Umformung zuzulassen.

Ich zeige Ihnen nun Methoden zur Lösung von einigen sehr häufig auftretenden Gleichungen und beginne sinnigerweise mit den einfachsten, den linearen.

3.1 Lineare Gleichungen

Nehmen wir an, Sie haben gerade genüsslich ein Wannenbad genommen und hierfür die Wanne mit 120 Liter Wasser befüllt; nun lassen Sie das Wasser ablaufen, wobei pro Minute 8 Liter Wasser durch den Abfluss gehen (altes Haus, verkalkte Rohre ...). Bezeichnet x die Anzahl der Minuten nach Öffnung des Abflusses, so berechnet die Badewannenabflussfunktion

$$b(x) \;=\; 120 - 8x$$

den verbleibenden Inhalt nach x Minuten. Wann ist die Wanne leer? Nun, „leer sein“ bedeutet mathematisch formuliert, dass 0 Liter in der Wanne sind (Mathematiker können *alles* kompliziert ausdrücken!), ich muss also die Gleichung

$$120 - 8x \;=\; 0$$

lösen. Eine solche Gleichung nennt man *linear*, da hier keine Potenz von x oder sonst etwas Gefährliches zu sehen ist. Um sie zu lösen, addiere ich zunächst auf beiden Seiten $8x$, das liefert $120 = 8x$, und multipliziere dann beide Seiten mit $\frac{1}{8}$, was auf die Lösung $x = 15$ führt.

Fasst man diese beiden Schritte zusammen, so gewinnt man die Erkenntnis, dass man diese Gleichung durch die Rechnung

$$x \;=\; -\frac{120}{-8} \;=\; 15$$

lösen kann. Das schreibe ich jetzt einmal als allgemeines Prinzip auf:

Lösung linearer Gleichung

Eine Gleichung der Form

$$ax + b \; = \; 0 \tag{3.3}$$

mit $a \neq 0$ nennt man **lineare Gleichung**, genauer **lineare Gleichung in Normalform**. Eine solche Gleichung hat genau eine Lösung, diese lautet

$$x \; = \; -\frac{b}{a} \; . \tag{3.4}$$

Dass a nicht null sein darf, hat hier gleich zwei Konsequenzen, die sich auch wunderbar ergänzen: Zum einen wäre die lineare Gleichung für $a = 0$ entweder trivial (falls $b = 0$) oder schlicht falsch (falls $b \neq 0$). Zum anderen dürfte man durch a nicht dividieren, wenn es null wäre. Sie sehen also: alles in bester Ordnung. Bis auf die Tatsache, dass ich noch nicht nachgerechnet habe, dass $-\frac{b}{a}$ tatsächlich Lösung der Gleichung ist. Machen wir das kurz noch vor dem Essen: Setzt man den angegebenen Wert in die linke Seite der Gleichung ein, so erhält man

$$a \cdot \left(-\frac{b}{a}\right) + b \; = \; -b + b \; = \; 0 \; ,$$

also den Wert der rechten Seite.

Hat man also eine lineare Gleichung in Normalform vorliegen, so braucht man nicht viel zu tun, sondern kann die Lösung direkt angeben. Allerdings liegt bei weitem nicht jede lineare Gleichung unmittelbar in Normalform vor; als Einstimmung hierfür schauen Sie sich doch einmal Aufgabe 3.1 an.

Übungsaufgabe 3.1:

Hier war schon lange nicht mehr von Pizza die Rede. Also dann: Sie wollen genau kalkulieren, was es Sie kostet, wenn Sie ein paar Freunde zur Pizza zu sich nach Hause einladen. Dazu haben Sie erfragt, dass eine Pizza 6 Euro kostet, wenn Sie also x Pizzen kaufen, kostet Sie das $6x$ Euro. Außerdem entstehen bei der Einladung von der Anzahl der Gäste unabhängige feste Kosten, beispielsweise durch das Heizen des Backofens und die Beleuchtung der Räume, von 12 Euro. Damit haben Sie insgesamt die so genannte Pizzapartykostenfunktion $P(x) \; = \; 12 + 6x$ aufgestellt. Wie viele Gäste können Sie mit einem Budget von 60 Euro einladen, wenn jeder eine Pizza isst?

■

In den folgenden Zeilen will ich Ihnen einige Beispiele von Gleichungen zeigen, die zwar lineare Gleichungen sind, aber z. T. so weit weg von der Normalform oben, dass man sie auf den ersten Blick nicht als solche erkennt. Gleichzeitig erfahren Sie natürlich, wie man diese Gleichungen in Normalform bringt und damit löst.

Beginnen wir mit einem ganz einfachen Fall und betrachten den Ausdruck

$$2x - 3(5 - x) \; = \; 3(2x - 5) + 9 \; . \tag{3.5}$$

Linear ist das ganz sicher, aber nicht in Normalform. Um diese zu erhalten, muss ich zunächst nach den im ersten Kapitel gelernten Regeln die Klammern ausmultiplizieren und erhalte

$$2x - 15 + 3x \;=\; 6x - 15 + 9 \;.$$

Nun fasse ich auf jeder Seite die mit x behafteten Terme und ebenso die nicht mit x behafteten zusammen, was auf

$$5x - 15 \;=\; 6x - 6$$

führt. Beachten Sie: Bisher habe ich jede Seite der Gleichung separat behandelt, insbesondere also keine der Regeln zur Umformung von Gleichungen benutzen müssen. Das kommt jetzt, denn wenn ich die Normalform erhalten will, muss ich auf der rechten Seite eine null erzwingen; dazu addiere ich das Negative der rechten Seite *auf beiden Seiten* und erhalte

$$-x - 9 \;=\; 0 \;.$$

Die Lösung der Gleichung (3.5) ist also $x = -9$, was ich mithilfe von (3.4) berechnet habe.

Als nächste Steigerung betrachten wir mal die Gleichung

$$(x-1)^2 + (x+1)(x-3) \;=\; 2(x-2)^2 \;. \tag{3.6}$$

Sieht nicht sehr linear aus, meinen Sie? Sie haben völlig Recht, aber das *sieht eben nur so aus.* Wenn man nämlich getreu dem Motto „zuerst ausmultiplizieren und zusammenfassen, danach erst denken“ vorgeht, erhält man

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

also

$$2x^2 - 4x - 2 \;=\; 2x^2 - 8x + 8 \;.$$

Jetzt addiere ich auch hier wieder das Negative der rechten Seite auf beiden Seiten und erhalte

$$4x - 10 \;=\; 0 \;,$$

der x^2-Term ist also verschwunden und die Lösung der Gleichung berechnet man mittels (3.4) als

$$x \;=\; \frac{10}{4} \;=\; \frac{5}{2} \;.$$

Setzt man dies nun zur Kontrolle wieder in (3.6) ein, so ergibt sich auf beiden Seiten der Wert $\frac{1}{2}$, also eine Gleichung.

Bevor Sie nun aber in falsche Euphorie verfallen und glauben, dass jede Gleichung nach dem Ausmultiplizieren eine lineare ist, zeige ich Ihnen diese:

$$x^2 - 4x + 1 \;=\; -2\;.$$

Hier kann man sortieren, so lange man will, der x^2-Term verschwindet nicht und die Gleichung ist nicht linear; was man in diesem Fall macht, erfahren Sie (wenn Sie es nicht ohnehin schon wissen) ein paar Zeilen weiter unten.

„Schlimmer geht's nimmer", meinen Sie? Falsch! Auch die Gleichung

$$\frac{2(x+1)}{x-2} \;=\; \frac{2(2-x)}{1-x} \tag{3.7}$$

ist in Wirklichkeit ein lineare. Um das zu zeigen, multipliziere ich mit dem Hauptnenner (erinnern Sie sich?) $(x-2)(1-x)$ durch und erhalte

$$2(x+1)(1-x) \;=\; (x-2)2(2-x)\;.$$

Das oben geübte Ausmultiplizieren und Zusammenfassen wandelt diese Gleichung um in

$$2 \;=\; 8x - 8\,,$$

also eine lineare Gleichung reinsten Wassers, deren Lösung man mithilfe von (3.4) berechnen kann; sie lautet $x = \frac{5}{4}$.

Nun kommt aber noch ein wichtiger Punkt: Die Multiplikation beider Seiten einer Gleichung mit derselben Zahl ist ja nur dann erlaubt, wenn diese Zahl nicht null ist. Wer sagt uns aber, dass der Term $(x-2)(1-x)$, mit dem wir im Laufe der Umformungen durchmultipliziert hatten, nicht gleich null ist? Zunächst einmal weiß man das tatsächlich nicht, man muss nach dem Motto „Augen zu und durch!" die gerade gezeigten Umformungen durchführen und die Lösung ermitteln. *Nun* setzt man diese Lösung (von der man noch nicht weiß, ob sie wirklich eine ist, weshalb sie manche Leute auch als **Scheinlösung** bezeichnen) in die Ausgangsgleichung ein und testet, ob sie diese löst, und ob insbesondere der Nenner nicht null wird. Wenn ja, so ist es eine Lösung, wenn nein, so ist die Gleichung unlösbar. Übrigens werden wir uns später, bei den so genannten Wurzelgleichungen, noch des Öfteren mit solchen Scheinlösungen herumplagen müssen.

Um das vorliegende Beispiel nun endlich abzuschließen, setze ich $x = \frac{5}{4}$ in die Ausgangsgleichung (3.7) ein: Beide Seiten erhalten dann den Wert -6, und insbesondere sind die Nenner weit davon entfernt, null zu sein.

Um Ihnen auch einmal ein Beispiel einer Bruchgleichung zu zeigen, bei der das schiefgeht, habe ich mir

$$\frac{x^2-1}{x-1} \;=\; x \tag{3.8}$$

ausgedacht. Die Multiplikation mit dem Nenner $x - 1$ ergibt hier

$$x^2 - 1 \;=\; x^2 - x \;,$$

und wenn ich nun auf beiden Seiten x^2 subtrahiere, kann ich die Scheinlösung $x = 1$ direkt ablesen. Setze ich diese aber in (3.8) ein, so ist die linke Seite nicht definiert, da der Nenner null würde; $x = 1$ ist also wirklich nur eine Scheinlösung und die Gleichung selbst hat keine Lösung.

Mehr möchte ich von meiner Seite aus nicht zu linearen Gleichungen sagen, sondern Sie ein wenig zum Einüben auffordern:

Übungsaufgabe 3.2:
Berechnen Sie die Lösungen der folgenden Gleichungen:

a) $2(x-3)+4 \;=\; 3(1-x)$

b) $3(x+1)(1-x) \;=\; 1 - x - 3x^2$

c) $\dfrac{x-1}{2(x+3)} \;=\; \dfrac{x}{2x-1}$ ■

3.2 Quadratische Gleichungen

Tritt in einer Gleichung die Variable nicht mehr nur linear, sondern mit dem Exponenten 2, also „im Quadrat“ auf, so nennt man auch die Gleichung selbst quadratisch:

Quadratische Gleichung

Eine Gleichung der Form

$$ax^2 + bx + c \;=\; 0$$

nennt man **quadratische Gleichung**.

Eine Gleichung der Form

$$x^2 + px + q \;=\; 0$$

nennt man **quadratische Gleichung in Normalform**.

Dass ich bei den beiden Gleichungen die Koeffizienten unterschiedlich bezeichnet habe, sollte Sie nicht weiter stören, das ist reine Konvention. Das Wichtige an der Normalform ist, dass der Koeffizient von x^2 gleich 1 ist. Ich werde mich im Folgenden nur mit quadratischen Gleichungen in Normalform befassen; wenn Sie jetzt denken: „Na, der macht sich's aber einfach!“, dann haben Sie Recht, aber ich darf das tun, und zwar nicht nur, weil ich der Autor bin, sondern weil damit tatsächlich auch der allgemeine Fall abgedeckt ist: Haben Sie es nämlich mit einer quadratischen Gleichung der Form $ax^2+bx+c=0$ zu tun, so können Sie, falls a nicht null ist, das Ganze zunächst durch a dividieren und erhalten

$$x^2 + \frac{b}{a}x + \frac{c}{a} \;=\; 0\,,$$

also eine Gleichung in Normalform. Ist aber $a = 0$, so brauchen Sie gar nichts zu tun, denn dann lautet die Gleichung ja einfach $bx + c = 0$, ist also eine lineare.

Konzentrieren wir uns also auf die Lösungen von quadratischen Gleichungen in Normalform.

Ich will jetzt ausnahmsweise einmal kein große Vorrede halten, sondern Ihnen direkt die Lösungsformeln für quadratische Gleichungen, die Sie wahrscheinlich ohnehin schon kennen, angeben.

Lösung quadratischer Gleichungen

Um die quadratische Gleichung

$$x^2 + px + q \;=\; 0 \tag{3.9}$$

zu lösen, berechnet man die beiden Zahlen

$$x_1 \;=\; -\frac{p}{2} + \sqrt{\frac{p^2}{4} - q}$$

und

$$x_2 \;=\; -\frac{p}{2} - \sqrt{\frac{p^2}{4} - q}\,.$$

Ist der Ausdruck unter der Wurzel, der *Radikand*, negativ, so sind beide Ausdrücke nicht definiert, und die Gleichung (3.9) hat keine Lösung.

Ist der Radikand gleich null, so ist auch der Wert der Wurzel gleich null; in diesem Fall stimmen die beiden Zahlen x_1 und x_2 überein und stellen die einzige Lösung der Gleichung (3.9) dar.

Ist der Radikand positiv, so ist auch der Wert der Wurzel eine positive Zahl; in diesem Fall sind x_1 und x_2 zwei verschiedene reelle Zahlen und stellen die beiden Lösungen der Gleichung (3.9) dar.

Ich sollte noch hinzufügen, dass man die beiden Formeln zur Berechnung von x_1 und x_2, die sich ja nur im Vorzeichen der Wurzel unterscheiden, meist in der Kurzform

$$x_{1/2} \;=\; -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$$

zusammenfasst. Man nennt sie auch p-q-Formel(n), da man so gut wie immer die Koeffizienten dieser Gleichung mit diesen Buchstaben bezeichnet, und ich wette darauf, dass das in jedem Lehrbuch, das Sie im Laufe Ihres Studiums in die Hand nehmen werden, auch so sein wird.

Ich hoffe, Sie glauben mir, wenn ich sage, dass ich diese Formeln im Ernstfall auch herleiten bzw. ihre Korrektheit beweisen könnte; da ich hier aber keine zwingende Notwendigkeit dafür sehe, lasse ich es einfach bleiben und illustriere das Ganze lieber an ein paar Beispielen.

Ich beginne mit der quadratischen Gleichung

$$x^2 + 3x + 2 = 0,$$

die praktischerweise schon in Normalform vorliegt. Die Anwendung der genannten Formeln (in der Kompaktversion) liefert

$$x_{1/2} = -\frac{3}{2} \pm \sqrt{\frac{9}{4} - 2} = -\frac{3}{2} \pm \sqrt{\frac{1}{4}} = -\frac{3}{2} \pm \frac{1}{2},$$

also

$$x_1 = -\frac{3}{2} + \frac{1}{2} = -1 \text{ und } x_2 = -\frac{3}{2} - \frac{1}{2} = -2.$$

Diese Gleichung hat also zwei verschiedene Lösungen.

Als zweites Beispiel untersuche ich die Gleichung

$$-2x^2 - 3x + 2 = 0$$

Um diese in Normalform zu bringen, dividiere ich sie durch -2, den Koeffizienten von x^2, und erhalte

$$x^2 + \frac{3}{2}x - 1 = 0.$$

Nun geht alles seinen Gang, ich wende die Lösungsformel an, und diese liefert

$$x_{1/2} = -\frac{3}{4} \pm \sqrt{\frac{9}{16} - (-1)} = -\frac{3}{4} \pm \sqrt{\frac{9}{16} + \frac{16}{16}} = -\frac{3}{4} \pm \sqrt{\frac{25}{16}} = -\frac{3}{4} \pm \frac{5}{4},$$

also

$$x_1 = -\frac{3}{4} + \frac{5}{4} = \frac{1}{2} \text{ und } x_2 = -\frac{3}{4} - \frac{5}{4} = -2.$$

Auch diese Gleichung hat also zwei verschiedene Lösungen.

Mein drittes Beispiel in diesem Kontext ist

$$x^2 - 4x + 4 = 0.$$

Anwendung der Lösungsformel auf diesen Fall liefert

$$x_{1/2} = 2 \pm \sqrt{4 - 4} = 2.$$

Der Radikand ist hier also null, und konsequenterweise sind die beiden Lösungen x_1 und x_2 identisch, es gibt also nur eine Lösung der Gleichung.

Bis hierhin ging ja alles gut, und vielleicht denken Sie ja, dass es eigentlich gar keine unlösbaren Gleichungen gibt bzw., dass diese so pathologisch aussehen, dass ich Ihnen keine vorsetzen will. Weit gefehlt, hier ist eine und die sieht doch eigentlich ganz harmlos aus:

$$x^2 + 2x + 3 = 0 .$$

Wenn ich aber hierauf die Lösungsformel anwende, erhalte ich

$$x_{1/2} = -1 \pm \sqrt{1-3} = -1 \pm \sqrt{-2} .$$

Die Wurzel aus der negativen Zahl -2 ist aber im Reellen nicht definiert und somit existiert keine Lösung der hier untersuchten quadratischen Gleichung.

Damit haben wir für jeden der drei möglichen Fälle mindestens ein Beispiel gesehen, und es wird Zeit, das Ganze noch weiter einzuüben; Sie ahnen schon, was kommt? Sie haben Recht:

Übungsaufgabe 3.3:
Bestimmen Sie die Lösungen der folgenden quadratischen Gleichungen:

a) $x^2 + \sqrt{8}x + 2 = 0$

b) $x^2 + (3 - \sqrt{3})x - \sqrt{27} = 0$ ■

Ebenso wie die linearen treten auch die quadratischen Gleichungen selten gleich in der Normalform auf, sondern müssen zur Lösung erst in diese gebracht werden. Auch hier zeige ich Ihnen dies anhand einiger Beispiele; ich denke, ich kann gleich mit einem etwas komplexeren beginnen:

$$x(x+2)(x-1) + 2x^3 - x = 3(x^2-1)(x+2) + 1 \tag{3.10}$$

Ausmultiplizieren der linken und der rechten Seite macht hieraus:

$$x^3 + x^2 - 2x + 2x^3 - x = 3x^3 - 3x + 6x^2 - 6 + 1$$

und Zusammenfassen ergibt

$$-5x^2 = -5 .$$

Es muss also

$$x^2 = 1$$

sein und entweder durch Anwendung der Lösungsformeln oder durch bloßes Hinsehen erkennt man, dass die Lösungen hiervon

$$x_1 \;=\; -1 \quad \text{und} \quad x_2 \;=\; 1$$

sind. Um das zu testen, können Sie nun diese beiden Zahlen in die Ausgangsgleichung (3.10) einsetzen; Sie werden sehen, dass sich auf beiden Seiten jeweils derselbe Wert ergibt.

Bei linearen Gleichungen hatte ich Ihnen gezeigt, dass sich diese manchmal sogar als Bruchgleichungen tarnen, und auch das ist bei den quadratischen Gleichungen nicht anders. Als Beispiel hierfür untersuche ich die Bruchgleichung

$$\frac{x}{x+1} \;=\; \frac{2-x}{2x} \;. \tag{3.11}$$

Durchmultiplizieren mit dem Hauptnenner $(x+1)2x$ ergibt die Form

$$2x^2 \;=\; (x+1)(2-x) \;.$$

Ausmultiplizieren der rechten Seite und Zusammenfassen der mit x^2, der mit x behafteten und der konstanten Terme kann ich Ihnen inzwischen wohl auf einen Schritt zumuten. Ich erhalte die quadratische Gleichung

$$3x^2 - x - 2 \;=\; 0\,,$$

also in Normalform

$$x^2 - \frac{1}{3}x - \frac{2}{3} \;=\; 0 \;.$$

Hierauf wende ich natürlich auch wieder die Lösungsformel an und erhalte

$$x_{1/2} \;=\; \frac{1}{6} \pm \sqrt{\frac{1}{36} + \frac{2}{3}} \;=\; \frac{1}{6} \pm \sqrt{\frac{25}{36}}\;,$$

also

$$x_1 \;=\; \frac{1}{6} + \frac{5}{6} \;=\; 1 \;\; \text{und}\; x_2 \;=\; \frac{1}{6} - \frac{5}{6} \;=\; -\frac{2}{3}\;.$$

Gleichung (3.11) hat also möglicherweise zwei verschiedene Lösungen, wobei sich das „möglicherweise“ darauf bezieht, dass es sich ja auch um Scheinlösungen handeln könnte. Das ist aber nicht der Fall, wie Sie durch Einsetzen von x_1 und x_2 feststellen können.

Bevor ich Sie mit der nächsten Übungsaufgabe alleine lasse – vielleicht sollte ich lieber sagen: ungestört lasse – noch ein letztes Beispiel zu versteckten quadratischen Gleichungen: Man bestimme die Lösungen von

$$\frac{1}{x} + \frac{1}{x-1} \;=\; \frac{1}{x+1} \;. \tag{3.12}$$

Sie wissen ja schon, was jetzt kommt: um die unangenehmen Brüche loszuwerden, multipliziere ich mit dem Hauptnenner $x(x-1)(x+1)$ durch und erhalte die umgeformte Gleichung

$$(x-1)(x+1) + x(x+1) \;=\; x(x-1) \;.$$

Wieder multipliziere ich aus und fasse zusammen, was auf

$$x^2 + 2x - 1 = 0$$

führt. Diese Gleichung ist praktischerweise schon in Normalform, sodass ich unmittelbar die Lösungsformeln anwenden kann und die Lösungen

$$x_1 = -1 + \sqrt{2} \text{ und } x_2 = -1 - \sqrt{2}$$

erhalte. Für diese beiden Zahlen wird sicherlich keiner der Nenner in (3.12) null, es handelt sich also um echte Lösungen dieser Gleichung und nicht um Scheinlösungen. Vielleicht wundert es Sie ja, dass die zunächst so glatt und ganzzahlig aussehende Gleichung (3.12) so „krumme" Lösungen hat, aber das kann bei quadratischen Gleichungen wegen der Wurzelbildung in der Lösungsformel eben jederzeit passieren und es gibt leider keinen Rechtsanspruch auf Ganzzahligkeit der Lösungen von Übungs- und Klausuraufgaben.

Übungsaufgabe 3.4:
Bestimmen Sie die Lösungen der folgenden Gleichungen:

a) $(x-1)^2 + x(x+2) = 2(x-1)(x+2)$

b) $\dfrac{x-1}{x+2} + \dfrac{x}{x-1} = \dfrac{2x^2+1}{(x+2)(x-1)}$ ■

Ausdrücke der Form $ax^2 + bx + c$, die ich hier als „linke Seite" von quadratischen Gleichungen untersuche, hatten Sie ja schon im Kapitel über Funktionen kennen gelernt, dort wurden sie als Polynome zweiten Grades oder auch Parabeln betitelt; und dort waren Sie auch dem Begriff der Nullstelle einer Funktion begegnet; jedenfalls wenn Sie dieses Buch in der vorgegebenen Reihenfolge der Kapitel durchlesen; falls nicht, so sollten Sie zum Verständnis der folgenden Zeilen einen kurzen Blick in das Funktionen-Kapitel werfen.

Wenn Sie das getan haben, dann ist Ihnen vielleicht aufgefallen, dass das Lösen einer quadratischen Gleichung nichts anderes ist als das Bestimmen der Nullstellen des Polynoms zweiten Grades, also der Funktion, das die linke Seite der Gleichung darstellt. Und die Tatsache, dass eine quadratische Gleichung keine, eine oder zwei Lösungen haben kann, deckt sich auch prächtig mit der Feststellung, dass ein Polynom zweiten Grades keine, eine oder zwei Nullstellen haben kann.

Damit wäre also eine Brücke zwischen dem Lösen von Gleichungen und dem Behandeln von Funktonen geschlagen; und wozu? Nun, beispielsweise um das Folgende, nämlich das so genannte **Faktorisieren von Polynomen**, besser zu verstehen. Ich beginne auch hier wieder mit einem Beispiel und untersuche das Polynom

$$p(x) = x^2 + 2x - 3$$

auf Nullstellen. Das bedeute also, dass ich die Gleichung

$$x^2 + 2x - 3 \;=\; 0$$

lösen muss, was aber mithilfe der gerade hinlänglich geübten p-q-Formeln kein Problem ist: Es existieren zwei verschiedene Nullstellen, nämlich $x_1 = 1$ und $x_2 = -3$. Nun mache ich einmal etwas zunächst recht Verwegenes und definiere eine neue Funktion, nennen wir sie $q(x)$, indem ich diese beiden Zahlen jeweils von der Variablen x abziehe und dann diese beiden Faktoren multipliziere; das haben Sie noch nicht ganz verstanden, stimmt's? Keine Sorge, das liegt nur an meiner verquasten Ausdrucksweise, hier sagt eine Formel tatsächlich mehr als tausend Worte. Ich definiere also die Funktion

$$q(x) \;=\; (x - x_1)(x - x_2) \;,$$

also

$$q(x) \;=\; (x - 1)(x + 3) \;.$$

Offenbar hat $q(x)$ die beiden Nullstellen $x_1 = 1$ und $x_2 = -3$, denn wenn ich diese beiden Zahlen einsetze, wird jeweils einer der beiden Faktoren null und somit auch die ganze Funktion. Da ich x_1 und x_2 ja als Nullstellen von p berechnet hatte, heißt das aber, dass p und q *dieselben* Nullstellen haben. Und jetzt kommt's: Damit sind sie auch schon komplett identisch.

Im vorliegenden Beispiel können Sie das nachrechnen, indem Sie $(x-1)(x+3)$ ausmultiplizieren: Sie sollten dann den eingangs formulierten Term von p wieder erhalten. Machen Sie das doch einfach einmal, und lesen Sie danach den folgenden Textkasten, der das gerade beispielhaft Erläuterte allgemein formuliert.

Zerlegung von Polynomen zweiten Grades in Linearfaktoren

Hat ein Polynom zweiten Grades $p(x) = ax^2 + bx + c$ zwei reelle Nullstellen x_1 und x_2, so hat es die Darstellung

$$p(x) \;=\; a(x - x_1)(x - x_2) \;.$$

Dies gilt auch, wenn diese beiden Nullstellen identisch sind, der Radikand in der p-q-Formel also null ist.

Man nennt dies die **Zerlegung von Polynomen zweiten Grades in Linearfaktoren**.

Das verlangt nicht nur, das schreit förmlich nach mindestens einem weiteren Beispiel. Ich untersuche hierfür das Polynom

$$p(x) \;=\; 2x^2 + 32x - 34$$

auf Nullstellen, löse also die Gleichung

$$2x^2 + 32x - 34 \;=\; 0 \;.$$

Entsprechend meiner eingangs gemachten Empfehlung, jede Gleichung in Normalform zu bringen, dividiere ich diese Gleichung erst mal durch den Koeffizienten von x^2, also 2. Das ergibt

$$x^2 + 16x - 17 \;=\; 0 \;,$$

und die Lösungsformeln, die ich nun schon des öfteren strapazieren musste, liefern hier die Nullstellen

$$x_1 \;=\; -8+\sqrt{64+17} \;=\; -8+9 \;=\; 1 \quad \text{und} \quad x_2 \;=\; -8-\sqrt{64+17} \;=\; -8-9 \;=\; -17.$$

Somit ist also

$$x^2 + 16x - 17 \;=\; (x-1)(x+17) \;.$$

Da ich durch 2 dividiert hatte, um die Nullstellenformel anwenden zu können, muss ich jetzt wieder mit 2 multiplizieren, um das Ausgangspolynom wieder zu erhalten. Es gilt also

$$p(x) \;=\; 2x^2 + 32x - 34 \;=\; 2(x-1)(x+17) \;.$$

Wieder einmal möchte ich Sie auffordern, die Richtigkeit dieser Gleichung probehalber durch Ausmultiplizieren der rechten Seite zu prüfen.

Diese Zerlegung in Linearfaktoren dient nicht nur der schöneren Darstellung des Polynoms (was ohnehin Geschmackssache wäre), sondern hat beispielsweise ganz praktischen Nutzen bei der vereinfachten Darstellung von rationalen Funktionen. Ich zeige Ihnen das am besten einmal an einem Beispiel: Die Funktion

$$r(x) \;=\; \frac{x^2 - 6x + 5}{x^2 + 4x - 45}$$

macht zunächst nicht den Eindruck, dass man sie in irgendeiner Weise kürzen könnte; schließlich gibt es den alten Merkspruch „Differenzen und Summen kürzen nur die Dummen“, also lassen Sie bloß die Finger von dem x^2-Term!

Viel besser ist es, den oben eingeübten Weg einzuschlagen und sowohl Zähler als auch Nenner in Linearfaktoren zu zerlegen. Ich denke, ich kann jetzt so langsam auf die ausführliche Darstellung der Lösungsformeln verzichten und direkt die Zerlegung angeben: Der Zähler hat die Nullstellen 1 und 5, also ist

$$x^2 - 6x + 5 \;=\; (x-1)(x-5) \;,$$

und der Nenner hat die Nullstellen 5 und -9, somit ist

$$x^2 + 4x - 45 \;=\; (x-5)(x+9) \;.$$

Insgesamt haben wir damit herausgefunden, dass die Funktion $r(x)$ auch geschrieben werden kann als

$$r(x) \;=\; \frac{(x-1)(x-5)}{(x-5)(x+9)} \;,$$

und das schreit förmlich danach, den Faktor $(x-5)$ herauszukürzen: man erhält

$$r(x) \;=\; \frac{x-1}{x+9} \;.$$

Übungsaufgabe 3.5:
Kürzen Sie die folgenden rationalen Funktionen so weit wie möglich:

a) $r_1(x) \;=\; \dfrac{x^3 + 7x^2 - 60x}{3x^2 - 27x + 60}$.

b) $r_2(x) \;=\; \dfrac{x^2 - 4}{2x - 4}$. ■

3.3 Polynomgleichungen höherer Ordnung

Im bisherigen Verlauf war ich immer stillschweigend davon ausgegangen, dass man bei der jeweils vorgelegten Gleichung bereits wusste, ob es sich um eine lineare oder eine quadratische Gleichung handelte, und einer der beiden Fälle trat ja auch auf jeden Fall ein, so verquast die jeweilige Gleichung auch zunächst ausgesehen haben mag; denken Sie nur an die Gleichung in (3.12), die sich als quadratische entpuppte.

Das ist aber im wahren Leben – soweit es mit Mathematik zu tun hat – leider sehr selten der Fall: Hier trifft man oft Gleichungen an, die zwar, wenn man Glück hat, nur Potenzen von x enthalten, aber ob es sich um eine lineare, eine quadratische oder eine Gleichung höheren Grades handelt, verrät einem leider kein Mensch, und man kann es oft auch nicht ohne Weiteres erkennen.

Da hilft nichts: Man muss zunächst einmal evtl. vorhandene Klammerausdrücke ausmultiplizieren, evtl. auftretende Nenner durch Erweitern verschwinden lassen und schließlich durch Sortieren nach x-Potenzen die Gleichung in die Normalform

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \;=\; 0$$

bringen; da die linke Seite die Form eines Polynoms hat, nennt man solche Gleichungen auch **Polynomgleichungen**. Wenn man viel Glück hat, verschwinden dabei alle höheren Potenzen von x, es bleibt also höchstens x^2 stehen, und man kann eine der oben geübten Lösungsmethoden anwenden. Wenn nicht, dann kann man eigentlich nur hoffen, dass einer der gleich noch vorzustellenden Spezialfälle vorliegt, denn Polynomgleichungen, deren Grad größer als zwei ist, sind im Allgemeinen praktisch nicht zu lösen. Für Grad drei und Grad vier gibt es zwar noch Lösungsformeln, aber diese sind so kompliziert, dass ich eigentlich keinen lebenden Menschen kenne, der sie jemals angewandt hätte. Falls es Sie dennoch interessiert, wie etwa diejenigen für Grad drei aussehen, dann finden Sie diese in der Literatur oder im Internet unter dem Namen *cardanische Formeln*; sie sind benannt nach und vermutlich auch gefunden von Geronimo Cardano, der auch das nach ihm benannte Cardan-Gelenk erfunden hat.

Ist der Grad der Gleichung fünf oder größer, so existieren keine Lösungsformeln, und das nicht etwa, weil die Mathematiker zu ungeschickt wären, sie zu finden, sondern weil der norwegische Mathematiker Niels Henrik Abel bewiesen hat, dass es keine solche Formel geben *kann*. Fragen Sie mich bitte nicht, wie er das bewiesen hat, schlimm genug, dass es ihm gelungen ist.

Aber ich bin nun schon zu stark abgeschweift, ich wollte Ihnen zunächst anhand von zwei Beispielen zeigen, wie man eine Gleichung in Polynomform bringen kann. Zum Warmlaufen nehme ich mir die Gleichung

$$(x^2+x-2)(x^2-2)-2 \;=\; x^3(x+1)-3(x^2-1)$$

vor. In dieser Form kann man sicherlich keine Lösungen erkennen, aber das hat ja auch niemand behauptet. Nach dem Ausmultiplizieren sieht die Sache schon freundlicher aus, es ergibt sich

$$x^4+x^3-4x^2-2x+2 \;=\; x^4+x^3-3x^2+3\ ,$$

und wenn man nun nach x-Potenzen sortiert, sieht man, dass sich sowohl der x^4- als auch der x^3-Term weghebt und eine simple quadratische Gleichung übrig bleibt, nämlich

$$x^2+2x+1 \;=\; 0\,,$$

welche die einzige Lösung $x=-1$ besitzt. Folglich hat auch die Ausgangsgleichung nur diese eine Lösung.

Als zweites Beispiel schaue ich mir die Gleichung

$$\frac{x^2+2}{x+5}+\frac{1}{x^3+x+1} \;=\; \frac{x}{x+1}$$

an. Hier wird man mit dem Hauptnenner, also dem Produkt der drei Nenner, durchmultiplizieren müssen; das liefert

$$(x^2+2)(x^3+x+1)(x+1)+(x+5)(x+1) \;=\; x(x+5)(x^3+x+1)\ ,$$

und wenn man dies ausmultipliziert und nach x-Potenzen sortiert, erhält man

$$x^6-2x^4+3x^3-2x^2+5x+7 \;=\; 0\ .$$

Sie sehen hier nicht, wie es weitergehen sollte? Ehrlich gesagt: ich auch nicht. Es hat ja auch niemand behauptet, dass sich jede Gleichung in eine quadratische oder lineare verwandelt, wenn man sie nur lange genug umformt, und hier haben wir nun eben einmal den Fall, dass eine Gleichung sechsten Grades herauskommt, die einfachen Lösungsmethoden unzugänglich ist. So ist das Leben, jedenfalls manchmal.

Übungsaufgabe 3.6:
Bestimmen Sie alle Lösungen der Gleichung

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

indem Sie sie zunächst in eine möglichst einfache Polynomgleichung verwandeln. ■

Ich hatte oben schon angedeutet, dass man in speziellen Situationen auch die Lösung von Polynomgleichungen höheren als zweiten Grades finden kann, obwohl die Lösungsformeln für Gleichungen dritten und vierten Grades äußerst kompliziert sind und es für Gleichungen höheren Grades keine allgemeinen Formeln gibt. Zum Abschluss dieses Abschnitts werde ich Ihnen nun exemplarisch einige solcher Situationen vorstellen, um Ihr Auge zu schärfen, wenn Ihnen so etwas einmal in einer Übungs- oder gar Klausuraufgabe begegnen sollte.

Der erste Fall hat eigentlich weniger mit mathematischen Überlegungen zu tun, sondern damit, dass auch die Autoren von Übungs- oder Klausuraufgaben einerseits von der Nichtexistenz höherer Lösungsformeln wissen und andererseits meist freundliche Menschen sind. Das bewirkt, dass man bei vorgelegten Gleichungen dritten (oder auch höheren) Grades eine Lösung meist mit bloßem Auge sehen, also *erraten* kann.

Beispielsweise springt bei der Gleichung

$$x^3 - x^2 + x - 1 \;=\; 0$$

eigentlich die Lösung $x = 1$ ins Auge, denn $1 - 1 + 1 - 1 = 0$.

Beim zweiten Beispiel

$$x^3 - 2x^2 + x - 2 \;=\; 0$$

findet man relativ schnell heraus, dass $x = 2$ eine Lösung ist. Ebenso wie bei der in Abschnitt 3.2 gezeigten Zerlegung in Linearfaktoren kann man auch in diesem Fall einen Linearfaktor abspalten, nämlich $(x - 2)$, und man erhält

$$x^3 - 2x^2 + x - 2 \;=\; (x - 2)(x^2 + 1)\ .$$

Da $x^2 + 1$ nicht null werden kann, hat die Gleichung

$$x^3 - 2x^2 + x - 2 \;=\; 0$$

außer 2 keine Lösung.

Ich gebe Ihnen jetzt einen gut gemeinten und durch langjährige Erfahrung motivierten Rat: Müssen Sie in einer Übung oder Klausur eine Gleichung dritten oder höheren Grades lösen, so testen Sie zunächst einmal die Werte $x = \pm 1$ und $x = \pm 2$ sowie, sollte das noch nicht zum Ziel geführt haben, die ganzzahligen Teiler des Absolutglieds (also des nicht mit x behafteten Koeffizienten) der Gleichung.

Meist ist die gesuchte Lösung hierunter zu finden. Wohlgemerkt: Das ist in keinster Weise ein mathematischer Satz oder so etwas, sondern fällt in die Kategorie Klausurautorenpsychologie; aber das kann ja auch manchmal ganz nützlich sein.

Der nächste Spezialfall, den ich Ihnen vorstellen will, ist nun wieder von streng mathematisch beweisbarer Natur. Als erstes Beispiel hierzu zeige ich Ihnen die Gleichung

$$23x^7 + 77x^4 - 17x^2 + 3x \;=\; 0\ .$$

Schon an den nicht ganz alltäglichen Koeffizienten sieht man, dass es auf diese gar nicht ankommen wird. Vielmehr ist bei dieser Gleichung entscheidend, dass sie kein

Absolutglied a_0 hat bzw., dass dieses null ist. Das bedeutet, dass jeder vorkommende Term mit x multipliziert wird und damit ist $x = 0$ eine Lösung der Gleichung.

Ich glaube, das brauche ich nicht mit noch mehr Beispielen unterlegen, vielmehr möchte ich es als allgemeine Regel formulieren:

Polynomgleichungen ohne Absolutglied

Enthält eine Polynomgleichung kein Absolutglied, so ist stets $x = 0$ eine Lösung dieser Gleichung.

Eine Gleichung ohne Absolutglied hat aber noch einen, mit dem gerade genannten eng verwandten Vorteil: Man kann x ausklammern und dadurch den Grad der verbleibenden Gleichung reduzieren. Ist dieser dann zwei oder eins, so kann man die oben vorgestellten Lösungsformeln anwenden.

Ein erstes einfaches Beispiel: Die Gleichung

$$x^3 - 9x = 0$$

hat $x = 0$ als Lösung und demnach kann man auch x ausklammern; es ergibt sich

$$x(x^2 - 9) = 0 .$$

Der verbleibende Faktor $x^2 - 9$ ist nun aber ein quadratisches Polynom, und die zugehörige Gleichung

$$x^2 - 9 = 0$$

besitzt die beiden Lösungen $x_1 = -3$ und $x_2 = 3$, was man entweder mithilfe der p-q-Formel oder der dritten binomischen Formel erkennt.

Ein zweites Beispiel ist die Gleichung vierten Grades

$$x^4 + 4x^3 + 3x^2 = 0 .$$

Hier kann man sogar x^2 ausklammern (streng genommen müsste ich nach derzeitigem Stand zunächst x ausklammern und dann ‚erstaunt' feststellen, dass ich gleich noch ein weiteres x ausklammern kann, aber das wäre albern) und man erhält

$$x^2(x^2 + 4x + 3) = 0 .$$

Die verbliebene quadratische Gleichung hat, wie Sie schnell nachrechnen können, die beiden Lösungen $x_1 = -1$ und $x_2 = -3$, womit wir auch in diesem Beispiel alle Lösungen gefunden hätten. Mehr Beispiele halte ich hier für unnötig.

Schließlich gibt es noch eine spezielle Situation, in der eine Gleichung höheren, in diesem Fall vierten, Grades einer recht einfachen Lösungsmöglichkeit zugänglich ist; ich rede von so genannten biquadratischen Gleichungen, das sind Gleichungen vierten Grades, bei denen die ungeraden Potenzen von x fehlen:

Biquadratische Gleichung

Eine Gleichung der Form

$$ax^4 + bx^2 + c = 0 \tag{3.13}$$

heißt **biquadratische Gleichung**.

Um der Lösung einer solchen Gleichung näher zu kommen, führe ich zunächst die *Substitution*

$$u = x^2$$

durch. Ich vergesse also sozusagen für den Moment, dass der Ausdruck x^2 durch Quadrieren einer gewissen Variablen entsteht, und fasse ihn zunächst einmal einfach als feste Größe auf, und um das zu dokumentieren, taufe ich ihn um, in diesem Fall in u. Ersetze ich nun in der biquadratischen Gleichung x^2 durch u, so erhalte ich die neue Gleichung

$$au^2 + bu + c = 0 \ , \tag{3.14}$$

da ja $x^4 = (x^2)^2 = u^2$ ist.

Diese neue Gleichung ist aber eine ganz simple quadratische Gleichung, für deren Lösung wir weiter oben eine handliche Formel gesehen haben. Hat man nun diese Lösungen – Existenz vorausgesetzt – berechnet, und handelt es sich dabei um nicht negative Zahlen, so kann man daraus wieder die Wurzel ziehen und erhält die gewünschten Lösungen der biquadratischen Gleichung (3.13). Genauer gesagt stellt sowohl die positiv als auch die negativ genommene Wurzel aus u eine Lösung von (3.13) dar, denn es gilt

$$(\pm\sqrt{u})^2 = (\sqrt{u})^2 = u = x^2 \ .$$

Da es maximal zwei verschiedene Lösungen u_1 und u_2 der Gleichung (3.14) gibt und hieraus jeweils wieder bis zu zwei verschiedene Wurzeln zu ziehen sind, gewinnt man auf diese Weise bis zu vier Lösungen der biquadratischen Gleichung, und da es sich dabei ja um eine Gleichung vierten Grades handelt, ist das auch gut so.

Ich weiß, dass es höchste Zeit für ein Beispiel ist, aber ich will das zunächst noch mal in Form einer merkfähigen Regel aufschreiben; sollten Sie das nicht für notwendig halten, so können Sie den nachfolgenden Kasten auch gerne überspringen und gleich zu den Beispielen übergehen.

Lösung biquadratischer Gleichungen

Um die Lösungen der biquadratischen Gleichung

$$ax^4 + bx^2 + c = 0$$

zu bestimmen, führt man die Substitution $u = x^2$ durch und berechnet die Lösungen u_1 und u_2 der quadratischen Gleichung

$$au^2 + bu + c = 0$$

in gewohnter Weise. Ist u_1 nicht negativ, so berechnet man

$$x_{11} = \sqrt{u_1} \quad \text{und} \quad x_{12} = -\sqrt{u_1} \ ,$$

und ist u_2 nicht negativ, so berechnet man ebenso

$$x_{21} = \sqrt{u_2} \quad \text{und} \quad x_{22} = -\sqrt{u_2} \ .$$

Die so berechneten Zahlen x_{11}, x_{12}, x_{21} und x_{22} sind dann Lösungen der biquadratischen Gleichung.

Schon gut, schon gut, Beispiele kommen ja schon. Als Erstes betrachten wir die Gleichung

$$x^4 - 6x^2 + 8 = 0 \ .$$

Schon wieder ist es passiert, ich bin in den typischen Mathematiker-Slang hineingeraten: Betrachten alleine reicht natürlich nicht, ich will die Gleichung sogar lösen. Dazu substituiere ich $u = x^2$ und löse die entstehende Gleichung

$$u^2 - 6u + 8 = 0$$

mithilfe der Lösungsformeln aus Abschnitt 3.2. Dies liefert

$$u_1 = 2 \quad \text{und} \quad u_2 = 4 \ .$$

Beides sind positive Zahlen, sodass ich munter Wurzeln ziehen darf und die folgenden vier Lösungen erhalte:

$$x_{11} = \sqrt{2}\,, \ x_{12} = -\sqrt{2}\,, x_{21} = 2\,, \ x_{22} = -2 \ .$$

Sie dürfen mir ruhig wieder misstrauen und diese vier Werte in die ursprüngliche biquadratische Gleichung einsetzen.

Als zweites Beispiel nehme ich mir die Gleichung

$$-2x^4 + 16x^2 - 32 = 0$$

vor. Diese muss ich zunächst in Normalform bringen, indem ich durch -2 durchdividiere; das liefert die Gleichung

$$x^4 - 8x^2 + 16 = 0 \ ,$$

die durch die nunmehr bereits gewohnte Substitution in

$$u^2 - 8u + 16 \; = \; 0$$

überführt wird. Wendet man hierauf die Lösungsformeln an, so erhält man

$$u_1 = u_2 = 4 \; .$$

Diese Gleichung hat also nur *eine* Lösung, aber das macht eigentlich gar nichts, immerhin ist diese positiv, sodass man die Wurzel daraus ziehen kann. Als Endergebnis erhält man die beiden Lösungen

$$x_{11} = 2 \quad \text{und} \quad x_{12} = -2$$

der Ausgangsgleichung, die Sie natürlich wahlweise auch mit x_{21} und x_{22} bezeichnen dürfen.

Als drittes und – das sei Ihnen zum Trost und zur Aufmunterung gesagt – vorletztes Beispiel nenne ich die Gleichung

$$x^4 - 8x^2 - 9 \; = \; 0 \; .$$

Substitution und Lösung der entstehenden quadratischen Gleichung führt auf

$$u_1 = -1 \quad \text{und} \quad u_2 = 9 \; .$$

Nun ist Vorsicht geboten, denn u_1 ist negativ, und bei allem, was oben gesagt wurde, stand stets als Vorbedingung, dass das betreffende u nicht negativ sein darf. Aus negativen Zahlen darf ich aber nun mal keine Wurzeln ziehen und daher lassen sich aus diesem u_1 auch keine x-Lösungen der Ausgangsgleichung gewinnen.

Anders sieht es aus mit u_2, das ist positiv, und somit haben wir das Ergebnis, dass die Ausgangsgleichung nur zwei Lösungen besitzt, nämlich

$$x_{21} = 3 \quad \text{und} \quad x_{22} = -3 \; .$$

Damit Sie in Ihrem vermutlich existierenden Vorurteil, dass in der Mathematik auch alles schief gehen kann, bestätigt werden, untersuche ich zum Schluss noch die Gleichung

$$x^4 + 2x^2 + 4 \; = \; 0 \; .$$

Substitution führt hier auf die Gleichung

$$u^2 + 2u + 4 \; = \; 0 \; ,$$

und die Lösungsformel liefert

$$u_{1/2} \; = \; -1 \pm \sqrt{1-4} \; .$$

Die substituierte Gleichung besitzt also keine einzige Lösung, aus der man evtl. noch Wurzeln ziehen könnte und somit besitzt die Ausgangsgleichung keine reelle Lösung.

Übungsaufgabe 3.7:
Bestimmen Sie alle Lösungen der folgenden biquadratischen Gleichungen:

a) $x^4 + x^2 + 1 \ = \ 0$

b) $2x^4 + 32x^2 - 450 \ = \ 0$

c) $x^4 - 13x^2 + 36 \ = \ 0$ ■

3.4 Wurzel- und Exponentialgleichungen

In der Mathematik, wie in jeder wissenschaftlichen Disziplin, wimmelt es von zunächst unverständlichen und zum Teil sogar irreführenden Bezeichnungen, das ist leider nicht zu leugnen. Manchmal schreibt das mathematische Leben aber auch schöne Geschichten und eine Bezeichnungsweise ist leicht verständlich und selbst erklärend; so ist es im Falle der Wurzelgleichungen, die ich Ihnen jetzt näher bringen will:

> **Wurzelgleichung**
> Eine Gleichung, bei der die Variable unter einer Wurzel steht, heißt **Wurzelgleichung**.

Ein erstes Beispiel ist die Gleichung

$$\sqrt{x-1} \ = \ 3 \ .$$

Um diese zu lösen, also ein x zu finden, für das beide Seiten der Gleichung denselben Wert haben, muss ich irgendwie die Wurzel wegbekommen. Dazu bietet es sich an, beide Seiten der Gleichung zu quadrieren, also mit 2 zu potenzieren; das ergibt

$$x - 1 \ = \ 9 \ ,$$

also eine lineare Gleichung, deren Lösung $x = 10$ Sie sofort erkennen können. Setzt man dies in die linke Seite der Wurzelgleichung ein, so erhält man

$$\sqrt{10-1} \ = \ \sqrt{9} \ = \ 3 \ ,$$

also den Wert der rechten Seite, und somit ist $x = 10$ als Lösung der Wurzelgleichung bestätigt.

Dieses Einsetzen in die Ausgangsgleichung ist im Falle von Wurzelgleichungen keineswegs nur eine Probe, um auszuschließen, dass man sich verrechnet hat, sondern ein unabdingbarer *Teil des Lösungsprozesses.* Warum das so ist, sehen Sie gleich:

Um die Gleichung

$$\sqrt{x-1} = -1$$

zu lösen, quadriere ich ebenfalls wieder beide Seiten und erhalte dadurch die lineare Gleichung

$$x - 1 = 1$$

mit der Lösung $x = 2$. Setze ich dies nun aber in die Ausgangsgleichung ein, erhalte ich auf der linken Seite

$$\sqrt{2-1} = \sqrt{1} = 1 ,$$

was leider nicht mit der rechten Seite -1 übereinstimmt.

Unschöne Sache, das Ganze: Die Lösung der umgeformten Gleichung hat uns auf eine Zahl geführt, die leider die Ausgangsgleichung *nicht* löst.

So etwas kommt bei Wurzelgleichungen häufig vor, man spricht dann wie bei den Bruchgleichungen von **Scheinlösungen**, und um solche auszuschließen, muss man die durch den Umformungsprozess gewonnenen „Lösungen" stets zur Kontrolle in die Ausgangsgleichung einsetzen.

Keine Sorge, zahlreiche Beispiele warten schon, aber Sie wissen ja inzwischen, dass ich gerne zunächst das Wichtigste in einem kleinen Merkkasten zusammenfasse:

Lösung von Wurzelgleichungen

Um eine Wurzelgleichung zu lösen, quadriert bzw. potenziert man sie – eventuell mehrfach –, um sie in eine Polynomgleichung umzuformen.

Dann löst man – falls möglich – die entstandene Polynomgleichung und setzt anschließend zum Ausschluss von Scheinlösungen die Lösungen der Polynomgleichung in die Ausgangsgleichung (Wurzelgleichung) ein.

Was bei Wurzelgleichungen so alles passieren kann, zeige ich Ihnen nun, wie schon gewohnt, anhand einiger Beispiele.

Als Erstes löse ich die Wurzelgleichung

$$\sqrt{17-x} = \sqrt{2x+14} . \tag{3.15}$$

Das empfohlene Quadrieren beider Seiten führt auf

$$17 - x = 2x + 14 .$$

Das ist nun zweifellos eine Polynomgleichung, und zwar eine ziemlich einfach zu lösende, denn sie ist linear. Sortiert man diese nach x-Termen und Konstanten, erhält man

$$-3x = -3,$$

also $x = 1$.

Dies ist aber zunächst einmal nur ein Kandidat für die Lösung der Ausgangsgleichung. Um zu verifizieren, dass es sich tatsächlich um eine Lösung und nicht etwa nur um eine Scheinlösung handelt, muss ich $x = 1$ in (3.15) einsetzen; es ergibt sich auf beiden Seiten der Wert $\sqrt{16} = 4$, und somit habe ich mit $x = 1$ tatsächlich eine Lösung gefunden.

Sie ahnen es sicher schon, jetzt kommt ein Beispiel, bei dem nicht alles glatt geht; ich versuche mich an der Gleichung

$$\sqrt{2x-3} = \sqrt{x^2-2x}\ . \tag{3.16}$$

Auch hier wieder führt das Quadrieren beider Seiten direkt auf eine Polynomgleichung, in diesem Fall eine quadratische:

$$2x - 3 = x^2 - 2x\ ,$$

also

$$x^2 - 4x + 3 = 0\ .$$

Um diese zu lösen, bemühe ich wieder einmal die p-q-Formel und erhalte

$$x_{1/2} = 2 \pm \sqrt{4-3} = 2 \pm 1,$$

also

$$x_1 = 3 \text{ und } x_2 = 1\ .$$

Setzt man nun $x_1 = 3$ in die Ausgangsgleichung ein, so ergibt sich auf beiden Seiten der Wert $\sqrt{3}$, es ist also alles in bester Ordnung und $x_1 = 3$ ist eine Lösung der Gleichung.

Setzt man aber $x_2 = 1$ ein, so erhält man auf der linken Seite einen negativen Radikanden, nämlich -1, und da die Wurzel aus einer negativen Zahl nicht definiert ist, ist es auch die linke Seite dieser Gleichung nicht und somit ist $x_2 = 1$ nur eine Scheinlösung.

Als drittes Beispiel dient die Wurzelgleichung

$$\sqrt{x - \sqrt{x-2}} = \sqrt{2}\ . \tag{3.17}$$

Auch hier wieder beginne ich in nunmehr schon gewohnter Weise damit, beide Seiten zu quadrieren und erhalte

$$x - \sqrt{x-2} = 2\ .$$

Hier ist es also mit einfachem Quadrieren nicht mehr getan, denn es steht immer noch ein Wurzelausdruck im Weg herum, den ich wegschaffen muss. Wenn ich die Gleichung aber in der hier vorliegenden Form quadriere, muss ich auf der linken Seite die binomische Formel anwenden und es bleibt auch nach dem Quadrieren ein Wurzelausdruck stehen (probieren Sie's mal aus!). Besser ist es, zunächst den Term $\sqrt{x-2}$ auf einer Seite der Gleichung zu *isolieren;* ich mache das, indem ich auf beiden Seiten $\sqrt{x-2}$ addiere, was mir

$$x \;=\; 2+\sqrt{x-2}$$

liefert, und anschließend 2 subtrahiere. Das Ergebnis ist

$$x-2 \;=\; \sqrt{x-2}\ .$$

Nun steht dem erfolgreichen Quadrieren nichts mehr im Wege, und dieses führt auf die quadratische Gleichung

$$x^2-4x+4 \;=\; x-2$$

oder in Normalform

$$x^2-5x+6 \;=\; 0\ .$$

Mit den p-q-Formeln will ich Sie nun wirklich nicht mehr anöden, ich wende sie einfach still und heimlich an und erhalte als Lösungen der quadratischen Gleichung die Zahlen

$$x_1=2 \quad \text{und} \quad x_2=3\ .$$

Damit bin ich aber noch *nicht fertig,* denn ich muss noch durch Einsetzen in (3.17) prüfen, ob es sich nicht um Scheinlösungen handelt.

Setze ich $x_1 = 2$ auf der linken Seite ein, so wird die innere Wurzel gleich null, und es verbleibt die äußere Wurzel, angewandt auf $x_1 = 2$, also ist der Wert der linken Seite gleich $\sqrt{2}$, was erstens definiert ist und zweitens mit der rechten Seite übereinstimmt. Das Einsetzen des zweiten Kandidaten, $x_2 = 3$, in die linke Seite führe ich nun ausnahmsweise etwas formelhaft durch: es wird

$$\sqrt{3-\sqrt{3-2}} \;=\; \sqrt{3-1} \;=\; \sqrt{2}\ .$$

Auch x_2 ist also eine echte Lösung dieser Wurzelgleichung, diese hat keine Scheinlösungen.

Als letztes Beispiel im Kontext Wurzelgleichungen schauen wir uns die folgende an:

$$\sqrt{x+1} \;=\; \sqrt{x}+3 \tag{3.18}$$

Auch hier empfehle ich, zunächst einmal zu quadrieren, um wenigstens eine der lästigen Wurzeln loszuwerden; es ergibt sich

$$x+1 \;=\; x+6\sqrt{x}+9\ .$$

(Wenn Sie an dieser Stelle ein anderes Ergebnis auf der rechten Seiten haben, schauen Sie noch mal unter „binomische Formeln“ nach!) Hier drängt es sich natürlich auf, auf beiden Seiten x zu subtrahieren, wodurch dieser Term schlicht und ergreifend verschwindet und die Restgleichung

$$-6\sqrt{x} \;=\; 8$$

übrig bleibt. Erneutes Quadrieren ergibt

$$36x \;=\; 64 \;,$$

also $x = \frac{16}{9}$. Wiederum darf ich nicht vergessen, diesen Kandidaten durch Einsetzen in die Ausgangsgleichung zu überprüfen; übrigens sollte man besser auch in der Politik die Kandidaten des Öfteren einmal genauer überprüfen, aber das nur am Rande.

Setze ich $\frac{16}{9}$ in der linken Seite von (3.18) ein, so erhalte ich

$$\sqrt{\frac{16}{9}+1} \;=\; \sqrt{\frac{25}{9}} \;=\; \frac{5}{3}\;,$$

immerhin ein glatter Wert. Auf der rechten Seite ergibt sich aber

$$\sqrt{\frac{16}{9}}+3 \;=\; \frac{4}{3}+3 \;=\; \frac{13}{3}\;,$$

und das ist auch bei großzügigster Auslegung algebraischer Regeln nicht identisch mit $\frac{5}{3}$. Ich habe also nur eine Scheinlösung ermittelt, die Ausgangsgleichung (3.18) besitzt gar keine Lösung.

Vielleicht denken Sie ja schon die ganze Zeit über, dass das ja alles gut und schön ist, dass es sich aber um rein innermathematische Denksportaufgaben handelt, die aber keinerlei Bezug zum wirklichen Leben haben. Das ist aber falsch, wie ich jetzt kurz zeigen will:

Nehmen Sie einmal an, Sie stehen auf einer hohen Klippe oder einem Leuchtturm und schauen aufs offene Meer hinaus. In welcher Entfernung können Sie noch einen Gegenstand auf der Wasseroberfläche erkennen (bessere Augen als meine vorausgesetzt), d. h., in welcher Entfernung verschwindet die Wasseroberfläche hinter dem Horizont?

Entgegen allem, was ich mir so beim Schreiben eines Lehrbuches vorgenommen habe, knalle ich Ihnen die Formel, die dieses löst, zunächst einmal hin und sage hinterher etwas Erläuterndes dazu. Sie lautet

$$l \;=\; \sqrt{12\,740\,000 \cdot h + h^2}\;.$$

Hierbei ist l die gefragte Entfernung und h die Höhe, in der Sie (genauer gesagt Ihre Augen) sich über dem Meeresspiegel befinden. Der merkwürdige Koeffizient 12 740 000 ist näherungsweise der Erddurchmesser, hier ebenfalls in Metern angegeben.

Zum Beweis bzw. zur Herleitung dieser Formel müssten wir den Kollegen Pythagoras und seinen Satz bemühen; den kennen Sie zwar aller Wahrscheinlichkeit schon, aber da er in diesem Buch bisher noch nicht auftauchte, will ich ihn hier auch nicht strapazieren und bitte Sie, die Formel so zu akzeptieren. Mit ihrer Hilfe können Sie beispielsweise berechnen, dass Sie von einer 20 Meter hohen Klippe etwa

$$\sqrt{12\,740\,000 \cdot 20 + 20^2} \;=\; 15\,962{,}47$$

Meter, also knapp 16 Kilometer weit sehen können.

Sehr hübsch, aber mindestens ebenso interessant ist die umgekehrte Fragestellung: Wie hoch müssen Sie, beispielsweise auf einem Aussichtsturm, klettern, um eine gewünschte Entfernung l, sagen wir $l = 20$ Kilometer, weit sehen zu können?

Zur Beantwortung dieser Frage müssen Sie die Gleichung

$$20\,000 \;=\; \sqrt{12\,740\,000 \cdot h + h^2}$$

nach h auflösen, das heißt also, eine Wurzelgleichung lösen!

Ich mache das mal für Sie, schließlich bin ich hier der Autor. Dazu quadriere ich natürlich wieder beide Seiten der Gleichung, um die ungeliebte Wurzel loszuwerden, und erhalte die quadratische Gleichung

$$400\,000\,000 \;=\; 12\,740\,000 \cdot h + h^2 \;,$$

deren Lösungen ich mit der p-q-Formel berechne; sie lauten

$$h_1 = 31{,}40 \quad \text{und} \quad h_2 = -12\,740\,031{,}40 \;.$$

Da ich nicht annehme, dass Sie sich durch die Erde hindurchgraben wollen, kommt nur die Lösung h_1 infrage; Sie müssen also etwa 31,40 Meter hoch steigen, um 20 Kilometer weit sehen zu können.

Übungsaufgabe 3.8:
Bestimmen Sie alle Lösungen der folgenden Wurzelgleichungen:

a) $\sqrt{x+4} \;=\; \sqrt{x^2+x}$

b) $\sqrt{x} + \sqrt{5-x} \;=\; \sqrt{2x+7}$

c) $\sqrt{x - \sqrt{x+2}} \;=\; 2$ ■

Ich komme nun zum zweiten Teilthema dieses Kapitels, den Exponentialgleichungen; hier ist es mit der Namensgebung ähnlich schön einfach wie bei den Wurzelgleichungen, sie erklärt sich quasi von selbst.

Exponentialgleichung

Eine Gleichung, bei der die Variable im Exponenten auftritt, heißt **Exponentialgleichung**.

Der einfachste Fall einer Exponentialgleichung ist die Gleichung der Form

$$a^x = b ,$$

wobei a und b positive reelle Zahlen sein sollen und a ebenfalls nicht gleich 1 ist.

Eine solche Gleichung kann man durch Anwendung des Logarithmus' unmittelbar lösen: es ist

$$x = \log_a(b) .$$

Wenn Sie den Umgang mit dem Logarithmus nicht mehr ganz so gewohnt sind, dann schauen Sie doch noch mal im zweiten Kapitel nach. Spätestens danach werden Sie mir zustimmen, dass die Lösung der Exponentialgleichung

$$3^x = 9$$

$x = 2$ lautet, denn $\log_3(9) = 2$, und ebenso, dass die Exponentialgleichung

$$10^x = 117$$

durch $x = \log_{10}(117) = 2{,}06818586...$ gelöst wird.

Am Rande möchte ich nur bemerken, dass es nicht nötig ist, den Logarithmus zu bemühen, wenn die rechte Seite der Gleichung, also b, selbst eine Potenz von a ist. Beispielsweise liest man an der Exponentialgleichung

$$17^{x+1} = 17^5$$

sofort ab, dass $x + 1 = 5$, x also gleich 4 sein muss.

Ich denke nicht, dass ich zu diesem einfachen Fall weitere Beispiele machen sollte, vielmehr können wir, wie Sie es ja schon gewohnt sind, sukzessive den Schwierigkeitsgrad der Beispiele erhöhen. Sollten Sie im Verlauf Ihres Studiums oder bei sonstigen Berührungen mit der Mathematik einmal auf eine Exponentialgleichung treffen, so finden Sie unter den folgenden sicherlich ein passendes Beispiel, an dem Sie sich orientieren können.

Die erste, noch winzige Schwierigkeit tritt auf, wenn die linke Seite a^x noch mit einem Vorfaktor behaftet ist, die Gleichung also beispielsweise

$$2 \cdot 5^x = 50$$

lautet. Sie sehen aber sofort, dass man in einem solchen Fall nur durch den Vorfaktor durchdividieren muss, was in diesem Fall zur Gleichung $5^x = 25$ führt, deren Lösung $x = 2$ einem wiederum sofort ins Auge springt.

Weiterhin kann es vorkommen, dass die Variable x bei mehreren Summanden im Exponenten vorkommt. Wenn man Glück hat, und das will ich jetzt erst einmal annehmen (zum Glück kann ich das Glück hier ja steuern), dann handelt es sich dabei jeweils um die gleiche Basis (also das gleiche a). Ein Beispiel für eine solche Gleichung ist

$$2 \cdot 3^x + 3^{x+1} = 135 \ . \tag{3.19}$$

In diesem Fall macht es sich bezahlt, wenn man beim Potenzrechnen gut aufgepasst hat; man weiß dann nämlich, dass $3^{x+1} = 3 \cdot 3^x$ ist, und somit kann ich die Gleichung auch schreiben als

$$2 \cdot 3^x + 3 \cdot 3^x = 135$$

oder

$$5 \cdot 3^x = 135 \ .$$

Nun sind wir aber im Fahrwasser, denn diese Form einer Exponentialgleichung haben wir oben bereits studiert. Man dividiert nun beide Seiten durch 5, was auf

$$3^x = 27$$

führt, und liest hieran die Lösung $x = 3$ ab. Setzen Sie diese doch probehalber mal in die Ausgangsgleichung (3.19) ein, Sie werden sehen, dass $x = 3$ diese tatsächlich zu einer wahren Aussage macht.

Das allgemeine Prinzip, welches Sie an diesem Beispiel vielleicht schon erkennen konnten, ist folgendes: Man sucht den kleinsten aller Summanden, welche x im Exponenten enthalten (hier ist das 3^x), und bringt die anderen (hier ist das nur 3^{x+1}) durch Zerlegung mithilfe der Potenzregeln in dieselbe Form.

Ein weiteres Beispiel hierfür ist die Exponentialgleichung

$$2^{x+1} - 3 \cdot 2^x + 5 \cdot 2^{x-1} = 48 \ .$$

Hier ist der kleinste der mit x behafteten Summanden 2^{x-1}, also zerlege ich die anderen beiden wie folgt:

$$4 \cdot 2^{x-1} - 3 \cdot 2 \cdot 2^{x-1} + 5 \cdot 2^{x-1} = 48 \ .$$

Zusammenfassen der linken Seite ergibt

$$3 \cdot 2^{x-1} = 48 \ ,$$

also

$$2^{x-1} = 16 \ .$$

Hieraus folgt, dass $x - 1 = 4$ und schließlich, dass $x = 5$ ist.

Leider hat man im Leben, um das Obige aufzugreifen, nicht immer Glück, und es kann vorkommen, dass x im Exponenten *verschiedener* Basen auftritt. In diesem Fall hilft nichts, da muss man logarithmieren. Schauen wir uns hierzu gleich ein typisches Beispiel an, nämlich die Exponentialgleichung

$$2^{x-1} \;=\; 3^{x+1}\ .$$

Ich wende nun auf beiden Seiten den Logarithmus an und benutze – das ist der entscheidende Trick – die Rechenregel $\log(a^b) = b \cdot \log(a)$. Es folgt

$$(x-1) \cdot \log(2) \;=\; (x+1) \cdot \log(3)\ .$$

Plötzlich ist aus dem ungeliebten Logarithmus ein simpler Koeffizient von x geworden. Jetzt multipliziere ich beide Seiten aus und sortiere nach x-Termen und Konstanten; das liefert

$$x(\log(2) - \log(3)) \;=\; (\log(2) + \log(3)).$$

Es ist also

$$x \;=\; \frac{\log(2) + \log(3)}{\log(2) - \log(3)} \;=\; \frac{\log(6)}{\log(\frac{2}{3})} \;=\; -4,419\ ,$$

wobei ich auch hier wieder von einer Logarithmenregel, in diesem Fall $\log(a)+\log(b) = \log(ab)$, Gebrauch gemacht habe.

Vielleicht ist Ihnen aufgefallen, dass ich noch gar nicht gesagt habe, *welchen* Logarithmus ich überhaupt verwende. Der Grund ist nicht so sehr meine fortschreitende Senilität, sondern vielmehr die Tatsache, dass es *völlig egal* ist, welchen Logarithmus man benutzt, wichtig ist nur, dass es auf beiden Seiten der Gleichung derselbe ist. Ich empfehle hier aus rein praktischen Erwägungen heraus, immer einen Logarithmus zu nehmen, der auch auf dem Taschenrechner implementiert ist und meist ist dies derjenige zur Basis 10. So habe ich auch die obige Lösung berechnet, Sie können ja zur Kontrolle das Ergebnis unter Benutzung eines anderen Logarithmus' einmal nachrechnen.

Den entscheidenden Trick zur Lösung von Exponentialgleichungen mit verschiedenen Basen habe ich Ihnen anhand dieses Beispiels schon gezeigt, und viel mehr gibt es hier eigentlich auch nicht zu sagen; außer vielleicht der Tatsache, dass man diesen Trick nicht unmittelbar anwenden kann, wenn auf mindestens einer Seite der Gleichung noch eine Summe steht. In einem solchen Fall muss man immer durch Umformungen, meist Ausklammern, erreichen, dass nur noch Produkte und Potenzen übrig bleiben.

Ein Beispiel ist die Gleichung

$$3^x + 3^{x+1} \;=\; 5^{x-1}\ .$$

Hier muss ich zunächst die linke Seite durch Ausklammern wie folgt umformen:

$$3^x + 3^{x+1} \;=\; 3^x + 3 \cdot 3^x \;=\; 4 \cdot 3^x\ .$$

Dies wiederum in die Gleichung eingesetzt liefert

$$4 \cdot 3^x \; = \; 5^{x-1} \; ,$$

logarithmieren formt dies um in

$$\log(4) + x \cdot \log(3) \; = \; (x-1) \cdot \log(5) \; ,$$

und nachdem man noch ein wenig Ordnung gemacht hat, lautet diese Gleichung

$$x \cdot (\log(3) - \log(5)) \; = \; -\log(4) - \log(5) \; ,$$

also

$$x \; = \; \frac{-\log(4) - \log(5)}{\log(3) - \log(5)} \; = \; 5{,}865 \; .$$

Ebenso wie bei den Wurzelgleichungen wäre es auch bei den Exponentialgleichungen falsch, anzunehmen, dass diese keinen Praxisbezug hätten; der Bezug kann sogar sehr nahe am täglichen Leben sein, jedenfalls wenn es um Geld geht. Im Zusammenhang mit den Logarithmen wurde dies schon einmal dargestellt, ich will es hier in der Sprache der Gleichungen nochmals kurz aufgreifen.

Im zweiten Kapitel wurde gezeigt, dass ein Kapital K, das zum Zinssatz von jährlich p Prozent angelegt wird, nach einem Zeitraum von x Jahren auf

$$K_x \; = \; K \cdot \left(1 + \frac{p}{100}\right)^x \tag{3.20}$$

angewachsen ist. Den Exponenten habe ich hier mit Absicht nicht mit n, sondern mit x bezeichnet, um deutlich zu machen, dass es sich hierbei nicht unbedingt um eine ganze Zahl handeln muss, denn diese Formel benutzen Banken auch, wenn das Kapital an einem beliebigen Tag des Jahres entnommen wird.

Sehr praxisrelevant, das werden Sie zugeben, ist in diesem Zusammenhang die folgende Fragestellung: Wie lange muss ein Kapital K angelegt werden, bis es auf eine vorgegebene Größe K_x angewachsen ist? In dieser Situation sind also K, K_x und natürlich p gegeben, gesucht ist x. Dadurch wird die Gleichung (3.20) zu einer Exponentialgleichung reinsten Wassers.

Vermutlich ist einmal mehr ein gutes Beispiel das Mittel der Wahl: Angenommen, Sie hätten 12 000 Euro zur Verfügung, und Ihre Bank bietet Ihnen an, diese Summe mit 3 % jährlich zu verzinsen. Wie lange müssen Sie das Geld liegen lassen, um eine Auszahlungssumme von 15 000 Euro zu erhalten?

Die Exponentialgleichung (3.20) lautet in diesem Fall:

$$15\,000 \; = \; 12\,000 \cdot 1{,}03^x \; .$$

Division durch 12 000 und anschließendes Kürzen verwandelt diese Gleichung in

$$\frac{5}{4} \; = \; 1{,}03^x$$

und anschließendes Logarithmieren liefert

$$x = \log_{1,03}\left(\frac{5}{4}\right) \approx 7{,}55 \,. \tag{3.21}$$

Sie müssen das Geld also etwa siebeneinhalb Jahre lang festlegen. Genauso rechnen das die Banken auch. Sollten Sie übrigens auf Ihrem Taschenrechner vergebens nach der Taste „Logarithmus zur Basis 1,03“ suchen, um dies nachzurechnen, so erinnern Sie sich bitte an die Umrechnungsformel

$$\log_a(x) = \frac{\log_b(x)}{\log_b(a)} \,.$$

Irgendeinen Logarithmus werden Sie auf Ihrem Rechner schon haben, vermutlich den zur Basis 10, und mit dessen Hilfe können Sie die Zahl x in (3.21) wie folgt ausrechnen:

$$x = \log_{1,03}\left(\frac{5}{4}\right) = \frac{\log_{10}\left(\frac{5}{4}\right)}{\log_{10}(1{,}03)} \,.$$

Damit verlasse ich – hoffentlich immer noch gemeinsam mit Ihnen – den Bereich der Gleichungen und wende mich zumindest kurz den Ungleichungen zu. Zuvor allerdings möchte ich Sie noch ein wenig zum Selbstdenken anregen:

Übungsaufgabe 3.9:
Bestimmen Sie alle Lösungen der folgenden Exponentialgleichungen:

a) $7^x = 17^x$

b) $2^{x+1} + 16 - 12 \cdot 2^{x-1} = 0$

c) $2^{x+1} + 2^{x+2} = 2 \cdot 3^{x-1}$ ■

Übungsaufgabe 3.10:
Wie lange muss man ein Kapital von einem Euro zu 4 % Zinsen anlegen, um 1 Million Euro ausgezahlt zu bekommen? ■

3.5 Ungleichungen

Ungleichungen heißen so, weil sie *ungleich* schwieriger zu behandeln sind als Gleichungen.

Nein, nicht gleich weiterblättern, war nur ein Scherz! Vielmehr tragen Ungleichungen diesen Namen, weil sie aus zwei Termen bestehen, die nicht wie Gleichungen durch ein Gleichheitszeichen, sondern durch ein Ungleichheitszeichen verbunden sind. Eine **Ungleichung** hat also die Form

$$a \ < \ b \ ,$$

gesprochen „a ist kleiner als b“, was eben bedeuten soll, dass a eine kleinere Zahl als b ist, oder

$$a \ > \ b \ ,$$

gesprochen „a ist größer als b“ im umgekehrten Fall. Da man durch Vertauschen der beiden Seiten einer Ungleichung stets eine ‚kleiner als‘-Ungleichung in eine ‚größer als‘-Ungleichung verwandeln kann und umgekehrt, genügt es, sich mit einem der beiden Fälle zu befassen; ich werde mich im Folgenden auf den erstgenannten Fall beschränken.

Neben den beiden gerade eingeführten echten Ungleichheitszeichen gibt es auch noch die Ihnen sicherlich schon bekannten Zeichen $\leq$ und $\geq$, gesprochen „kleiner oder gleich “ bzw. „größer oder gleich “, die es auch zulassen, dass die beiden durch sie verbundenen Terme gleich sind. Da die Behandlung solcher Ungleichungen identisch ist mit derjenigen der echten Ungleichungen, beschränke ich mich auch in dieser Hinsicht auf den Fall der echten Ungleichung.

Wie schon im Fall der Gleichungen betrachtet man auch bei Ungleichungen meist solche, die eine Variable enthalten, und versucht sie zu *lösen:*

Ungleichung

Eine Ungleichung besteht aus zwei mathematischen Ausdrücken (Termen), die durch ein Ungleichheitszeichen verbunden sind.

Enthält mindestens einer der Ausdrücke eine Variable, so sucht man nach **Lösungen der Ungleichung**. Dies sind Zahlen, die, wenn man sie anstelle der Variablen einsetzt, bewirken, dass die Werte der beiden durch das Ungleichheitszeichen verbundenen Terme gemäß diesem in Relation stehen. Die Menge aller Lösungen, die auch leer sein kann, bezeichnet man als **Lösungsmenge** der Ungleichung und symbolisiert sie mit $\mathbb{L}$.

Im Gegensatz zu Gleichungen hat eine Ungleichung meist unendlich viele Zahlen als Lösung, weshalb ich ab jetzt auch verstärkt von der Schreibweise als Lösungs*menge* Gebrauch machen werde.

Ein allererstes Beispiel soll Sie gleichzeitig auf eine kleine Tücke hinweisen: Die Ungleichung

$$x + 1 \ < \ 2$$

wird – dazu braucht man noch keine Lösungsverfahren – durch alle Zahlen x gelöst, die kleiner als 1 sind, die Lösungsmenge dieser Ungleichung ist also

$$\mathbb{L} = \{x \in \mathbb{R} \,|\, x < 1\} \ .$$

Die Tücke, von der ich gerade schrieb, besteht nun darin, dass dies keineswegs bedeutet, dass die Elemente von $\mathbb{L}$ *betragsmäßig* kleiner als 1 sein müssen, sie müssen vielmehr nur irgendwo links von der Eins auf der Zahlengeraden liegen. Mit anderen Worten: In $\mathbb{L}$ liegen nicht nur beispielsweise $\frac{1}{2}, 0, -\frac{3}{4}$, sondern auch -2 und $-1\,000\,000$. Und wenn Sie jetzt gerade denken: „Was soll das, das ist doch klar!“, so glauben Sie mir bitte, dass das im Eifer des Gefechts sehr oft falsch gemacht wird, so klar es auch sein mag.

Nun aber zu den Lösungsverfahren für Ungleichungen; auch hier gibt es eine große Übereinstimmung mit der Vorgehensweise bei Gleichungen, denn auch hier unterwirft man die zu lösende Ungleichung gewissen erlaubten Umformungen:

Umformung von Ungleichungen

Die folgenden Umformungen einer Ungleichung ändern nicht deren Lösungsmenge und werden **erlaubte Umformungen** genannt:
- Addition bzw. Subtraktion derselben Zahl auf beiden Seiten der Ungleichung
- Multiplikation beider Seiten mit derselben *positiven* Zahl

Haben Sie den Unterschied zu Gleichungen bemerkt? Ich habe ihn extra kursiv gedruckt: Bei Ungleichungen genügt es nicht mehr zu fordern, dass die Zahl, mit der durchmultipliziert wird, von null verschieden ist, sie darf auch nicht negativ sein!

Was nun aber, wenn man doch gezwungen ist, mit einer negativen Zahl zu multiplizieren? Keine Sorge, hier kommt Abhilfe:

Multiplikation einer Ungleichung mit einer negativen Zahl

Die Multiplikation beider Seiten einer Ungleichung mit derselben negativen Zahl ändert nicht ihre Lösungsmenge, ist also erlaubt, wenn man gleichzeitig das Ungleichheitszeichen, das die beiden Seiten verbindet, umkehrt, also „$<$“ in „$>$“ verwandelt und umgekehrt.

Das ist keineswegs so verrückt, wie es zunächst klingen mag: Beispielsweise ist die Menge aller Zahlen x, für die $x > 1$ gilt, identisch mit der Menge aller Zahlen, für die $-x < -1$ gilt; probieren Sie es ruhig einmal mit ein paar Werten aus.

Nun aber endlich zur Vorstellung der Lösungsverfahren, bei denen ich mich (und Sie) auf lineare Ungleichungen beschränken werde, da man an diesen alles Notwendige bereits erklären kann: Zur Bestimmung der Lösungsmenge der Ungleichung

$$3x + 4 \ < \ 1 \tag{3.22}$$

subtrahiere ich zunächst auf beiden Seiten 4, was ergibt

$$3x \ < \ -3 \ ,$$

und multipliziere anschließend mit der positiven Zahl 1/3. Die Lösungsmenge der Ungleichung (3.22) ist also

$$\mathbb{L} = \{x \in \mathbb{R} \,|\, x < -1\} \ .$$

War doch gar nicht so schlimm, oder? Ich nutze diese positive Grundstimmung und führe Sie an die etwas kompliziertere Ungleichung

$$x^2 - 2x + 3 \ < \ (x-1)(x+2) \tag{3.23}$$

heran. Eingangs hatte ich versprochen, mich auf lineare Ungleichungen zu beschränken, und nun so etwas! Aber Sie sehen wahrscheinlich schon, dass man hier natürlich auch so vorgeht wie bei Gleichungen und zunächst beide Seiten der Ungleichung vereinfacht; das ergibt in diesem Fall

$$x^2 - 2x + 3 \ < \ x^2 + x - 2 \ .$$

Wenn ich nun auf beiden Seiten x^2 subtrahiere, bleibt die lineare Ungleichung

$$-2x + 3 \ < \ x - 2$$

übrig, die ich noch in die handlichere Form

$$-3x \ < \ -5$$

bringe. Nun muss ich noch durch -3 dividieren bzw. mit $-\frac{1}{3}$ multiplizieren und dabei ist Vorsicht geboten. Da dies eine negative Zahl ist, muss ich, wie oben gesagt, das Ungleichheitszeichen umkehren, die Ungleichung (3.23) nimmt also die Form

$$x \ > \ \frac{5}{3}$$

an, ihre Lösungsmenge ist somit

$$\mathbb{L} = \left\{ x \in \mathbb{R} \,|\, x > \frac{5}{3} \right\} \ .$$

Im Prinzip scheint also alles – mit diesem kleinen Holperer bei der Multiplikation mit negativen Zahlen – ebenso glatt zu laufen wie bei den Gleichungen. Na ja, das ist auch so, allerdings kann dieser Holperer ziemlich komplizierte Auswirkungen haben, was ich Ihnen nun noch in einem letzten instruktiven Beispiel verdeutlichen will, bevor ich Sie, nicht ohne Ihnen noch ein paar Übungsaufgaben ans Herz gelegt zu haben, aus der Welt der Gleichungen und Ungleichungen wieder entlasse.

Das angesprochene Beispiel ist die Ungleichung

$$\frac{x-1}{2x+1} \ < \ 1 \ . \tag{3.24}$$

Um diese in die Standardform linearer Ungleichungen zu bringen, will ich mit dem Nenner der linken Seite durchmultiplizieren. Das Frage ist nun: „Ist dieser Nenner positiv oder negativ?“ Denn davon hängt ja ab, was mit dem Ungleichheitszeichen passiert. Die Antwort auf diese Frage ist: „Ich habe nicht die geringste Ahnung!“

Demnach bleibt nichts anderes übrig, als beide Möglichkeiten getrennt zu untersuchen und die Ergebnisse am Ende zusammenzuführen.

Fall 1: $2x + 1 > 0$, also $x > -\frac{1}{2}$: In diesem Fall kann ich die Ungleichung (3.24) durchmultiplizieren, ohne am Ungleichheitszeichen etwas zu ändern, und erhalte

$$x - 1 \ < \ 2x + 1 \ ,$$

also die Bedingung $-x < 2$ oder

$$x > -2 \ .$$

Zum guten Schluss muss man sich nun noch einmal gut konzentrieren, um das Folgende zu verstehen: In diesem ersten Fall, den wir gerade untersuchen, muss x also zwei Bedingungen erfüllen: Es muss aufgrund der Fallunterscheidung größer als $-\frac{1}{2}$ sein und es muss als Ergebnis der Umformung größer als -2 sein. Im Allgemeinen sind dies also zwei Bedingungen, die zur Formulierung der Lösungsmenge beitragen werden, in diesem Beispiel allerdings – und das wird meist so sein – ist die erste Bedingung schärfer als die zweite (da jede Zahl, die größer als $-\frac{1}{2}$ ist, automatisch größer als -2 ist), weshalb hier in Wirklichkeit nur eine Bedingung übrig bleibt, nämlich

$$x \ > \ -\frac{1}{2} \ . \tag{3.25}$$

Fall 2: $2x + 1 < 0$, also $x < -\frac{1}{2}$: Auch hier multipliziere ich die Ungleichung (3.24) durch, muss nun aber das Ungleichheitszeichen umkehren und erhalte

$$x - 1 \ > \ 2x + 1 \ ,$$

also $x < -2$. Wie in Fall 1 haben wir auch hier zwei Bedingungen an x, nämlich $x < -\frac{1}{2}$ und $x < -2$, von denen eine – hier ist es die zweitgenannte – die andere automatisch nach sich zieht. Im Fall 2 ergibt sich also die Bedingung

$$x \ < \ -2 \ . \tag{3.26}$$

Mehr Fälle gibt es nicht, da $2x + 1$ sicher nicht null sein darf, so dass ich aus (3.25) und (3.26) die Lösungsmenge der Ungleichung (3.24) gewinnen kann; sie ist

$$\mathbb{L} \ = \ \left\{ x \in \mathbb{R} \,|\, x < -2 \ \text{ oder } \ x > -\frac{1}{2} \right\} \ .$$

So, jetzt aber Sie:

Übungsaufgabe 3.11:
Bestimmen Sie die Lösungsmengen der folgenden Ungleichungen:

a) $2x + 1 \ < \ 3x - 5$

b) $2(x + 1)(x - 2) \ < \ 2x^2 + 5$

c) $\dfrac{1 - 2x}{x + 1} \ < \ -1$ ■

4 Geometrie

Schon seit alters her haben sich die Menschen mit Dreiecken, Vierecken und allgemeineren Figuren wie zum Beispiel Kreisen und Ellipsen intensiv beschäftigt. Der Grund hierfür ist sicherlich *nicht*, dass man damals sehr viel mehr Freizeit hatte und an die Erfindung des Fernsehens noch lange nicht zu denken war. Vielmehr traten solche Figuren in sehr vielen Situationen des täglichen Lebens auf und die Beherrschung dieser Objekte erlaubte es, mit den eingeschränkten Ressourcen hauszuhalten, um mit halbwegs vertretbarem Arbeitseinsatz ein damals im Allgemeinen karges Leben zu bestreiten. Insbesondere entwickelten die alten Griechen Kenntnisse über die Längen- und Winkelverhältnisse solcher, **ebene Figuren** genannten, Objekte, um beispielsweise die Länge eines zu bauenden Kanals, Distanzen zwischen weiter entfernten Orten, die Höhe eines zu erstellenden Gebäudes oder aber auch die Fläche eines zu bewirtschaftenden Grundstücks geeignet einschätzen zu können. Inzwischen hat sich die Menschheit – auch, und insbesondere, aufgrund dieser Kenntnisse – entscheidend weiterentwickelt. Heute benötigen wir diese, wenn wir beispielsweise einen Teppichboden effizient verlegen möchten, wissen möchten, wie groß die Teile einer Pizza sind, die wir uns mit unseren Freunden teilen, oder aber dem Phänomen von auf dem Grill zusammenschmurgelnder Hamburger auf die Spur kommen möchten. Spaß beiseite – in diesem Kapitel werde ich Ihnen die Grundlagen der Berechnungen am **Dreieck** im Zusammenhang mit **trigonometrischen Funktionen** erklären. Hier werde ich Ihnen ein paar zentrale Aussagen, die Sie aus Ihrer Schulzeit sicher noch kennen, in Erinnerung bringen. Dann werde ich Ihnen zeigen, wie man allgemeinere, geradlinig begrenzte Figuren, so genannte **Vielecke**, behandeln kann – der Haupttrick hierbei ist, dass ich diese in Dreiecke zerlege und die Dinge anwende, die ich über Dreiecke weiß. Schließlich werde ich mit Ihnen auch noch ein bisschen über **Kreise** und **Ellipsen** plaudern.

4.1 Dreiecke und trigonometrische Funktionen

In diesem Abschnitt beabsichtige ich, Ihnen Grundlegendes über besonders einfache geometrische Figuren in Erinnerung zu bringen. Diese Figuren sind **Dreiecke**. Gebe ich mir drei Punkte A, B und C in der Ebene vor, die nicht auf einer gemeinsamen Gerade liegen, so kann ich daraus ein Dreieck Δ mit den **Eckpunkten** A, B und C bilden, indem ich die drei Geraden durch je zwei dieser Punkte miteinander schneide. Die Strecken zwischen je zwei Eckpunkten von Δ mit den Längen a, b und c heißen dann **Seiten von** Δ, während die Winkel α, β und γ innerhalb von Δ an den Eckpunkten **Innenwinkel von** Δ genannt werden. Da zwei sich in einem Punkt schneidende Geraden maximal zwei verschiedene Winkel einschließen, treten diese Winkel α, β und γ auch außerhalb des Dreiecks auf. **Außenwinkel von** Δ nennt man jedoch die entsprechenden Winkel α', β' und γ', die jeweils zweimal außerhalb von Δ auftreten. Schauen Sie sich Abbildung 4.1 an – Sie erkennen, dass die dem Eckpunkt A gegenüberliegende Seite die Länge a hat und ich die dortigen Winkel mit α und α' bezeichne.

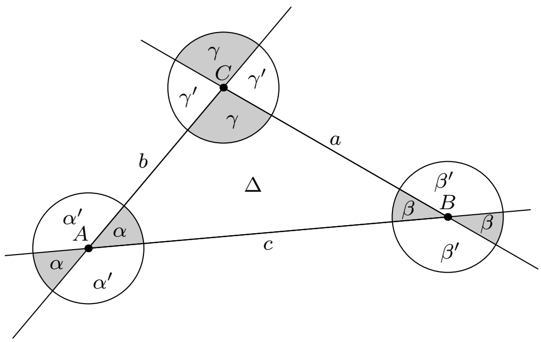

Abb. 4.1 Die Punkte, Seitenlängen und Innen- und Außenwinkel eines Dreiecks.

In Abbildung 4.2 zeige ich Ihnen verschiedene Typen von Dreiecken. Die dort vorgenommene Klassifizierung unterscheidet in **spitzwinklige**, **rechtwinklige**, **stumpfwinklige**, **gleichschenklige** und **gleichseitige** Dreiecke. Im spitzwinkligen Dreieck sind alle drei Innenwinkel kleiner als 90° (Gradmaß), während für rechtwinklige Dreiecke ein Innenwinkel 90° hat (dies kennzeichnet man durch einen Punkt) und für stumpfwinklige Dreiecke ein Innenwinkel mit mehr als 90° existiert. Für gleichschenklige Dreiecke sind zwei Seiten gleich lang und gleichseitige Dreiecke sind ein Sonderfall gleichschenkliger Dreiecke, bei dem sogar alle drei Seiten gleich lang sind.

Abhängig von der Art eines vorliegenden Dreiecks kann ich, wie ich Ihnen weiter unten zeigen werde, Aussagen über den Zusammenhang verschiedener Größen des Dreiecks machen. So stimmen beispielsweise zwei der drei Innenwinkel eines gleichschenkligen Dreiecks überein und für gleichseitige Dreiecke sind sogar alle drei Innen-

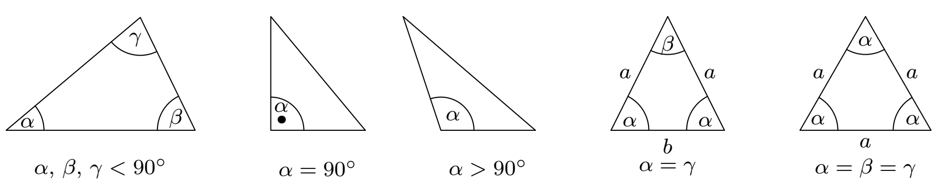

Abb. 4.2 Klassen von Dreiecken: spitzwinklig, rechtwinklig, stumpfwinklig, gleichschenklig und gleichseitig (von links nach rechts).

winkel gleich. Warum das so ist, werde ich Ihnen später ganz einfach zeigen können. Ich werde Ihnen jetzt zunächst ein paar grundlegende Beziehungen nennen, die für *jedes* Dreieck gültig sind.

Schauen Sie sich zur Klärung einer ersten solchen Beziehung noch einmal Abbildung 4.1 an: Um vom Eckpunkt A zum Eckpunkt B auf möglichst kurzem Weg zu gelangen, muss ich eine Strecke der Länge c zurücklegen. Jede andere Wahl stellt einen Umweg da, weil die kürzeste Verbindung zwischen zwei Punkten durch die Strecke zwischen diesen Punkten festgelegt ist. So ist es zum Beispiel ein Umweg, wenn ich von A aus erst nach C gehe und danach von C nach B, denn dann habe ich insgesamt die größere Strecke $a + b$ zurückgelegt. Diese Beobachtung halte ich als **Dreiecksungleichung** fest:

Dreiecksungleichung

Die Summe der Länge zweier Seiten a und b eines Dreiecks ist stets größer als die Länge der verbleibenden Seite c: $a + b > c$.

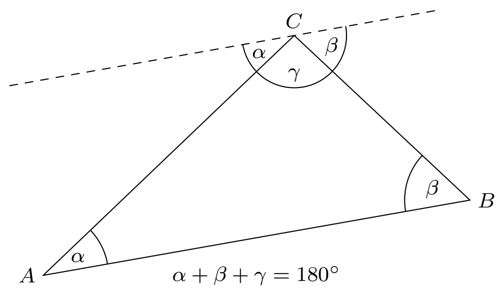

Abb. 4.3 Die Summe der Innenwinkel im Dreieck ist 180°.

Als Nächstes schaue ich mir Abbildung 4.3 an. Hier zeige ich Ihnen ein Dreieck zusammen mit einer (gestrichelten) Gerade, die parallel zu einer Seite des Dreiecks ist und durch den gegenüberliegenden Eckpunkt C geht. Der Innenwinkel α am Punkt A tritt hier ein *zweites Mal* als **Wechselwinkel** auf – dieser ist nämlich auch der Winkel zwischen dieser Gerade und der Seite des Dreiecks, die A und C verbindet. Analog verhält es sich mit dem Innenwinkel β bei B. Abbildung 4.3 zeigt Ihnen jetzt, dass die Addition des verbleibenden Innenwinkel γ zu der Summe dieser beiden Duplikate von α und β immer gleich 180° ist. Beachte ich zudem, dass die Außenwinkel in Relation mit den Innenwinkel stehen:

$$\alpha' = 180° - \alpha,\ \beta' = 180° - \beta \text{ und } \gamma' = 180° - \gamma,$$

so erhalte ich zusätzlich

$$\alpha' + \beta' + \gamma' = 3 \cdot 180° - \overbrace{(\alpha + \beta + \gamma)}^{180°} = 360°.$$

Das kann man sich merken.

Summe der Winkel am Dreieck

Die Summe $\alpha + \beta + \gamma$ der Innenwinkel α, β, γ eines Dreiecks ist stets 180°. Die Summe $\alpha' + \beta' + \gamma'$ der Außenwinkel α', β', γ' eines Dreiecks ist stets 360°.

Insbesondere zeigen mir diese Aussagen, dass für jedes Dreieck alle Innenwinkel und Außenwinkel kleiner als 180° sind. Für spezielle Dreiecke erkenne ich zudem das Folgende: In rechtwinkligen Dreiecken summieren sich die beiden Innenwinkel, die sich vom **rechten Winkel** (das heißt der Innenwinkel, der 90° hat) unterscheiden, stets zu 90° auf. In stumpfwinkligen Dreiecken summieren sich die beiden Innenwinkel, die sich vom **stumpfen Winkel** (das heißt der Innenwinkel, der mehr als 90° hat) unterscheiden, zu weniger als 90° auf. Ist ein gleichschenkliges Dreieck gleichzeitig ein rechtwinkliges Dreieck, so sind die beiden verbleibenden Winkel jeweils 45°. In gleichseitigen Dreiecken ist jeder Innenwinkel 60° und jeder Außenwinkel 120°.

Ich fahre nun fort, indem ich Ihnen ein wenig über **Transversale** und deren Eigenschaften berichte. Mit diesem Begriff bezeichnet man gemeinhin charakteristische Geraden, die ein vorgegebenes Dreieck schneiden. Die wichtigsten Vertreter solcher Geraden am Dreieck habe ich für Sie in Abbildung 4.4 zusammengestellt: Mittelsenkrechte, Winkelhalbierende, Seitenhalbierende und Höhe.

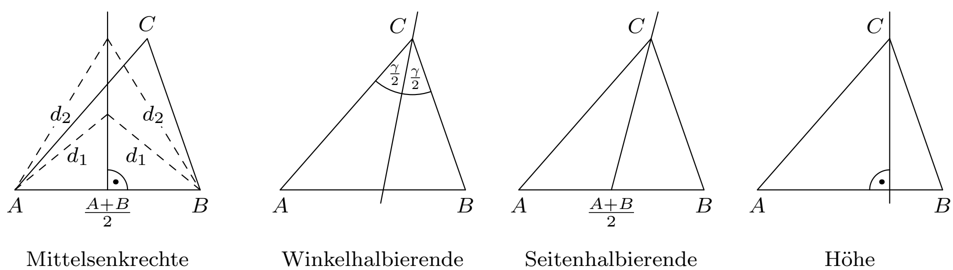

Abb. 4.4 Transversale – von links nach rechts: Mittelsenkrechte gehen durch den Mittelpunkt einer Seite des Dreiecks und stehen senkrecht auf diese Seite. Winkelhalbierende gehen durch einen Eckpunkt des Dreiecks und halbierend den dortigen Innenwinkel. Seitenhalbierende gehen durch den Mittelpunkt einer Seite des Dreiecks und den gegenüberliegenden Eckpunkt. Höhen gehen durch einen Eckpunkt des Dreiecks und stehen senkrecht auf der diesem Punkt gegenüberliegenden Seite.

Eine Gerade wird **Mittelsenkrechte** eines Dreiecks genannt, falls sie durch den Mittelpunkt einer Seite des Dreiecks geht und **senkrecht** (auch: **orthogonal**) auf dieser Seite steht, das heißt, der Winkel zwischen der Gerade und dieser Seite ist 90°.

Offenbar hat jeder Punkt auf einer Mittelsenkrechten die Eigenschaft, dass der Abstand zu den beiden Eckpunkten der zugehörigen Seite des Dreiecks stets gleich ist – ich habe dies in Abbildung 4.4 (links) durch gestrichelte Strecken mit Längen d_1 und d_2 angedeutet. Der Schnittpunkt S *zweier* Mittelsenkrechten desselben Dreiecks ist somit von allen drei Eckpunkten des Dreiecks *gleich weit entfernt* und S liegt somit auch auf der verbleibenden, dritten Mittelsenkrechten. Da der Abstand r von S zu den drei Eckpunkten eines vorgegebenen Dreiecks Δ somit jeweils gleich ist, kann ich nun sehen, dass S der Mittelpunkt eines Kreises mit Radius r durch die drei Eckpunkte von Δ ist. Dieser Kreis wird **Umkreis** des Dreiecks genannt, weil es der kleinste Kreis ist, den ich *um* das Dreieck herum legen kann. Ich habe das für Sie in Abbildung 4.5 (links) illustriert. Wie man den Umkreis aus drei vorgegebenen Eckpunkten explizit angeben und ausrechnen kann, werde ich Ihnen später in Beispiel 4.15 zeigen. Dort finden Sie weitere Informationen über Kreise.

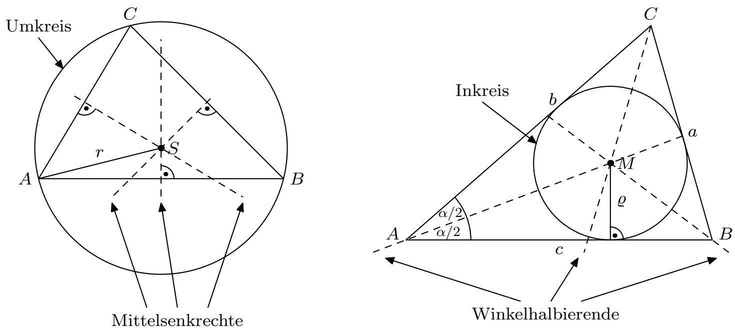

Abb. 4.5 Die Mittelsenkrechten eines Dreiecks schneiden sich im Mittelpunkt des Umkreises (links), während sich die Winkelhalbierenden eines Dreiecks im Mittelpunkt des Inkreises schneiden (rechts).

> **Schnitt der Mittelsenkrechten und Umkreis**
>
> Die drei Mittelsenkrechten eines Dreiecks schneiden sich im Mittelpunkt des Umkreises des Dreiecks.

Winkelhalbierende sind Transversale, die durch einen Eckpunkt des Dreiecks gehen und den dortigen Innenwinkel halbieren. Abbildung 4.5 (rechts) motiviert, dass sich die drei Winkelhalbierenden eines vorgegebenen Dreiecks ebenfalls in einem Punkt schneiden. Dies ist der Mittelpunkt des **Inkreises** des Dreiecks. Man nennt diesen Kreis so, weil es der größte Kreis ist, den ich *in* das Dreieck hineinlegen kann.

> **Schnitt der Winkelhalbierenden und Inkreis**
>
> Die drei Winkelhalbierenden eines Dreiecks schneiden sich im Mittelpunkt des Inkreises des Dreiecks.

Ich verrate Ihnen an dieser Stelle noch, dass der Radius ϱ (sprich: „ro“) des Inkreises (auch: **Inkreisradius**) eines beliebigen Dreiecks mit Seitenlängen a, b und c gemäß

$$\varrho = \sqrt{\frac{(s-a)(s-b)(s-c)}{s}},$$

wobei $s = (a+b+c)/2$ berechnet werden kann.

Im Folgenden werde ich gleich ein bisschen mit Punkten in der Ebene „rechnen“. Das mache ich wie folgt:

Sind $P = (p_1|p_2)$ und $Q = (q_1|q_2)$ zwei Punkte, so kann ich diese addieren, indem ich deren Komponenten addiere

$$P + Q = (p_1 + q_1|p_2 + q_2)$$

und ich kann P mit einer reellen Zahl λ (**Skalar** genannt) multiplizieren, indem ich dies ebenfalls für die Komponenten so mache

$$\lambda P = (\lambda p_1|\lambda p_2)\ .$$

Diese Operationen bilden die Basis der **Analytischen Geometrie**. Hier reicht es aus, zu erkennen, dass der Punkt

$$\tfrac{1}{2}(P + Q) = \tfrac{1}{2}\,P + \tfrac{1}{2}\,Q = \left(\tfrac{1}{2}(p_1 + q_1) \mid \tfrac{1}{2}(p_2 + q_2)\right)$$

in der Mitte der Strecke von P nach Q liegt und

$$\tfrac{1}{3}\,P + \tfrac{2}{3}\,Q = \left(\tfrac{1}{3}\,p_1 + \tfrac{2}{3}\,q_1 \mid \tfrac{1}{3}\,p_2 + \tfrac{2}{3}\,q_2\right)$$

der Punkt auf der Strecke von P nach Q ist, der diese im Verhältnis $2 : 1$ teilt.

Seitenhalbierende sind Transversale, die durch den Mittelpunkt einer Seite des Dreiecks und den gegenüberliegenden Eckpunkt gehen (siehe Abbildung 4.4 (Mitte, rechts)). Ist $M_{AB} = (A + B)/2$ ein solcher Mittelpunkt, so ist C der entsprechende gegenüberliegende Eckpunkt und ich sehe, dass der (mit 1/3 und 2/3) gewichtete Punkt

$$P = \tfrac{1}{3}\,C + \tfrac{2}{3}\,M_{AB}$$

auf dieser Seitenhalbierenden liegt und die Strecke von C nach M_{AB} im Verhältnis $2 : 1$ teilt. Wegen $M_{AB} = (A + B)/2$ kann ich P aber auch wie folgt hinschreiben:

$$P = \tfrac{1}{3}\,C + \tfrac{2}{3}\,(A + B)/2 = \tfrac{1}{3}\,C + \tfrac{1}{3}\,(A + B) = \tfrac{1}{3}\,(A + B + C).$$

Denselben Punkt bekomme ich somit heraus, wenn ich die Mittelpunkte der verbleibenden Seiten und die entsprechenden gegenüberliegenden Eckpunkte betrachte und in derselben Weise entlang der beiden verbleibenden Seitenhalbierenden gewichte.

Schnitt der Seitenhalbierenden

Die drei Seitenhalbierende eines Dreiecks Δ schneiden sich im **Schwerpunkt** $P = (A + B + C)/3$ des Dreiecks. Dieser teilt jede Strecke eines Eckpunkts von Δ zum Mittelpunkt von dessen gegenüberliegender Seite im Verhältnis $2 : 1$.

Schließlich sind **Höhen** Transversale durch die Eckpunkte, die jeweils senkrecht auf der diesem Punkt gegenüberliegenden Seite des Dreiecks stehen (siehe Abbildung 4.4

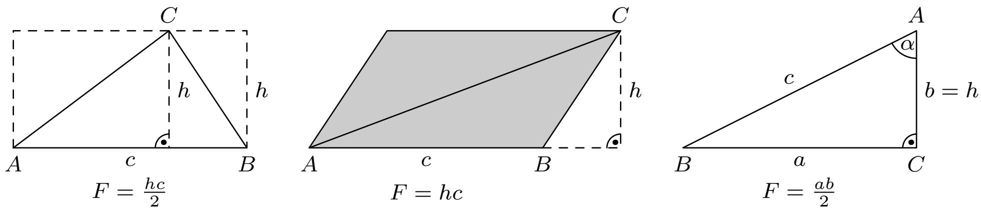

Abb. 4.6 Der Flächeninhalt jedes Dreiecks Δ mit Eckpunkten A, B und C ist die Hälfte des Flächeninhalts eines Rechtecks mit Seitenlängen h und c (links). Die Formel gilt auch, wenn die Höhe außerhalb von Δ liegt, und somit ist der Flächeninhalt eines Parallelogramms hc (Mitte). Für rechtwinklige Dreiecke vereinfacht sich die Anwendung dieser Formel, denn man muss die Länge der Höhe nicht berechnen: Der Flächeninhalt des Dreiecks ist dann die Hälfte des Produkts der Längen der Katheten des Dreiecks (rechts).

(rechts)). Zu Ihrer Information verrate ich Ihnen, dass diese sich ebenfalls in einem Punkt schneiden.

Kenne ich die **Länge einer Höhe** h, das heißt den Abstand des zu einer Höhe gehörigen Eckpunkts zum Schnittpunkt auf der gegenüberliegende Seite des Dreiecks, und die Länge dieser Seite c, so kann ich den Flächeninhalt dieses Dreiecks ausrechnen. Abbildung 4.6 (links) zeigt Ihnen nämlich, dass dieser gerade die Hälfte des Flächeninhalts eines Rechtecks mit Seitenlängen h und c ist.

Flächeninhalt des Dreiecks

Der Flächeninhalt eines Dreiecks F ist die Hälfte des Produkts der Länge einer Höhe h mit der Länge c der Seite, auf der diese Höhe senkrecht steht: $F = hc/2$.

Diese einfache Formel ist für jedes Dreieck – also insbesondere auch für *stumpfwinklige* Dreiecke – gültig. Mit diesem Fakt kann ich nun aber auch sofort sehen, dass die Fläche eines **Parallelogramms** durch hc gegeben ist, wobei c die Länge einer Grundseite und h die Länge der Höhe auf diese Grundseite ist. Dies zeigt Ihnen Abbildung 4.6 (Mitte), denn die graue Fläche des dortigen Parallelogramms ist das Doppelte der Fläche des Dreiecks mit den Eckpunkten A, B und C. Im nächsten Abschnitt werde ich Ihnen noch ein paar weitere Informationen über Parallelogramme und allgemeinere Vierecke geben.

Eine weitere prominente Berechnungsformel für den Flächeninhalt eines Dreiecks Δ hat den Vorteil, dass man keine Höhe bestimmen muss, sondern direkt die Längen der drei Seiten von Δ verwenden kann. Um zu sehen, wie man darauf kommt, schaue ich mir Abbildung 4.5 (rechts) an. Den Flächeninhalt des Teildreiecks mit den Eckpunkten A, B und M kann ich nach der obigen Formel einfach als $\varrho\ c/2$ berechnen, denn die Länge des Inkreisradius ϱ ist gerade die Länge der Höhe auf die Seite dieses Dreiecks mit den Eckpunkten A und B. Analog ergeben sich für den Flächeninhalt des Dreiecks mit den Eckpunkten B, C und M, beziehungsweise des Dreiecks mit den Eckpunkten

C, A und M, die Werte $\varrho\ a/2$ und $\varrho\ b/2$. Da der Flächeninhalt des Dreiecks mit den Eckpunkten A, B und C offenbar gerade die Summe der Flächeninhalte dieser drei Teildreiecke ist, erhalte ich das folgende Resultat:

Herons Formel

Der Flächeninhalt eines Dreiecks F mit Seitenlängen a, b und c ist die Hälfte des Produkts des Inkreisradius ϱ mit der Summe der Seitenlängen: $F = \varrho\ (a+b+c)/2$.

Oben hatte ich erwähnt, dass der Inkreisradius eines Dreiecks Δ durch die Formel $\varrho = \sqrt{(s-a)(s-b)(s-c)/s}$, wobei $s = (a+b+c)/2$, gegeben ist. Wenn ich ϱ in Herons Formel durch diesen Ausdruck ersetze, so erhalte ich für den Flächeninhalt F von Δ

$$\begin{aligned} F = \varrho s &= \sqrt{s(s-a)(s-b)(s-c)} \\ &= \tfrac{1}{4}\ \sqrt{(a+b+c)(b+c-a)(a+c-b)(a+b-c)}. \end{aligned}$$

Man benötigt also zur Berechnung von F lediglich die drei Seitenlängen a, b und c von Δ – durch Vorgabe von (sinnvollen) Seitenlängen eines Dreiecks ist dessen Flächeninhalt bestimmt.

Manchmal ist es nun so, dass ich für ein Dreieck Δ lediglich die drei Eckpunkte $A = (a_1|a_2), B = (b_1|b_2)$ und $C = (c_1|c_2)$ kenne und ich zur Bestimmung des Flächeninhalts mit den obigen Vorgehensweisen entweder die Länge einer Seite und die Länge der auf diese Seite senkrecht stehende Höhe berechnen oder aber alle drei Seitenlängen bestimmen müsste. Ich verrate Ihnen jetzt eine Formel zur Berechnung des Flächeninhalts F von Δ, die mir diesen eventuell etwas mühseligen Umweg erspart:

$$F = \tfrac{1}{2}\ \Big(a_1(b_2-c_2) + b_1(c_2-a_2) + c_1(a_2-b_2)\Big)\ .$$

Besonders einfach wird die Berechnung des Flächeninhalts eines Dreiecks darüberhinaus, wenn dieses *rechtwinklig* ist. In diesem Fall liegen nämlich zwei der drei Seiten auf Höhen des Dreiecks und ich muss die zur Flächenberechnung benötigte Höhe nicht berechnen. Es reicht dann, lediglich die Hälfte des Produkts der Längen a und b dieser beiden Seiten zu bestimmen (siehe Abbildung 4.6 (rechts)). Diese beiden Seiten schneiden sich im Eckpunkt mit dem rechten Innenwinkel und werden **Katheten** genannt, während die verbleibende Seite des rechtwinkligen Dreiecks (mit Länge c in Abbildung 4.6 (rechts)) **Hypotenuse** heißt.

Übungsaufgabe 4.1:
Es seien $A = (0|0)$ und $B = (4|0)$. Bestimmen Sie den Flächeninhalt des Dreiecks Δ durch die Punkte A, B und C, falls a) $C = (2|2)$, b) $C = (0|2)$, c) $C = (6|2)$ und d) C der Punkt mit Abstand $\sqrt{5}$ zu A und mit Abstand $\sqrt{13}$ zu B, ist. ■

Es bietet sich nun meiner Ansicht nach an, dass ich mich gemeinsam mit Ihnen ein wenig mit weiteren Aussagen über Flächen am *rechtwinkligen* Dreieck beschäftige.

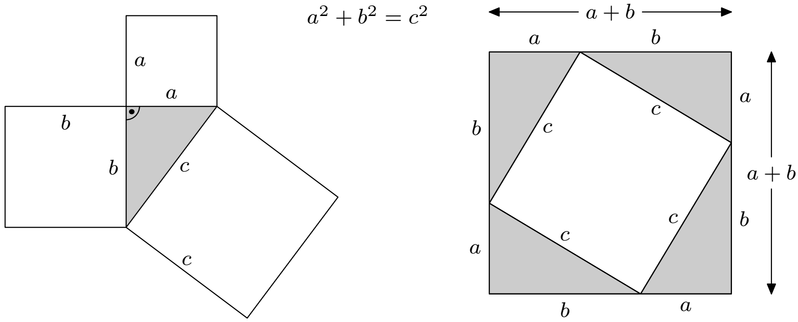

Abb. 4.7 Veranschaulichung des Satzes von Pythagoras (links) und eine der etwa 100 bekannten Herleitungen dieser Aussage (rechts).

Zu den absoluten Klassikern solcher Aussagen gehört der seit etwa 2 500 Jahren bekannte **Satz des Pythagoras** (nach Pythagoras von Samos, um 580 – 496 vor Christi Geburt), den ich Ihnen als Erstes wieder in Erinnerung bringen möche. Zur Herleitung betrachte ich Abbildung 4.7 (rechts). Der Flächeninhalt des dortigen (gesamten) Quadrats mit Seitenlänge $a + b$ ist gemäß der ersten **Binomischen Formel**

$$(a + b)^2 = a^2 + 2\,ab + b^2,$$

während c^2 der Flächeninhalt des (gekippten) Quadrats mit Seitenlänge c in dessen Inneren ist. Jedes der vier grauen Dreiecke ist rechtwinklig und gemäß der obigen Aussage ist deren Flächeninhalt somit jeweils $ab/2$. Addiere ich also zu c^2 die Summe der Flächeninhalte der vier grauen Dreiecke hinzu, so gelange ich für den Flächeninhalt des (gesamten) Quadrats auf die alternative Formel $c^2 + 2ab$. Da dieser Flächeninhalt aber einen eindeutigen Wert hat, erhalte ich

$$a^2 + 2\,ab + b^2 = c^2 + 2\,ab.$$

Jetzt kann ich den Term $2\,ab$ auf beiden Seiten der letzten Gleichung abziehen und gelange auf die Aussage, die ich in Abbildung 4.7 (links) veranschaulicht habe:

Satz des Pythagoras

Für rechtwinklige Dreiecke ist die Summe der Quadrate der Längen der Katheten a, b stets gleich dem Quadrat der Länge der Hypotenuse c: $a^2 + b^2 = c^2$

Übungsaufgabe 4.2:
In einem rechtwinkligen Dreieck habe die Hypotenuse die Länge 5 und die Fläche des Quadrats über einer der beiden Katheten sei 16. Wie lang ist die verbleibende Kathete? ■

Der Satz des Pythagoras zeigt insbesondere, dass die Länge c der Hypotenuse jeweils größer als die Längen der beiden Katheten ist

$$c > \max\{a, b\},$$

denn $c^2 = a^2 + b^2 > \max\{a^2, b^2\}$ (hier verwende ich max als Abkürzung für das Maximum zweier Zahlen). Diese Bemerkung wird weiter unten bei einer kleinen Diskussion im Rahmen des Sinussatzes von Bedeutung sein.

Übungsaufgabe 4.3:
In einem rechtwinkligen Dreieck Δ sei die Summe der Längen der Katheten 2 und die Hypotenuse sei doppelt so lang wie eine der beiden Katheten. Bestimmen Sie die Längen der Katheten und die Länge der Hypotenuse von Δ. ■

Beispiel 4.1:
In gleichseitigen Dreiecken Δ mit vorgegebener Seitenlänge a ist die Länge jeder Höhe h gleich. Da in diesem Fall die Mittelsenkrechten mit den Höhen übereinstimmen, gilt nach dem Satz des Pythagoras $h^2 + (a/2)^2 = a^2$. Es ergibt sich $h^2 = a^2 - a^2/4 = 3a^2/4$ und somit $h = \sqrt{3}\ a/2$. Hiermit kann ich jetzt den Flächeninhalt F eines solchen Dreiecks bequem ausrechnen: Es gilt die Formel $F = ha/2 = \sqrt{3}\ a^2/4$ für den **Flächeninhalt des gleichseitigen Dreiecks**. ■

Übungsaufgabe 4.4:
Zeigen Sie, dass der **Flächeninhalt *F* gleichschenkliger Dreiecke** durch $F = b\sqrt{4a^2 - b^2}/4$ gegeben ist, wobei a die Länge der beiden Schenkel und b die verbleibende Seitenlänge ist. ■

Eine oftmals auftretende Anwendung des Satzes von Pythagoras ist die Bestimmung des **Abstands *d* zweier Punkte** – nennen wir sie $(x_1|y_1)$ und $(x_2|y_2)$ – in der Ebene. Dieser ist gegeben durch die Länge der Strecke zwischen diesen beiden Punkten, die ich als Hypotenuse eines rechtwinkligen Dreiecks mit Kathetenlängen $x_1 - x_2$ und $y_1 - y_2$ auffassen kann. Es ergibt sich somit

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

Ein weiterer Klassiker, über den ich mit Ihnen zu sprechen habe, ist der **Satz des Euklids**. Hierzu können Sie sich Abbildung 4.8 anschauen. Betrachte ich das graue Dreieck Δ, so bildet die Kathete mit Länge a eine Seite dieses Dreiecks und die Länge der (außerhalb von Δ liegenden) Höhe $\tilde{h}$ von Δ über dieser Kathete hat dieselbe Länge: $\tilde{h} = a$. Es gilt also, nachdem was ich Ihnen oben über die Berechnung des Flächeninhalts von Dreiecken erzählt habe, dass $a^2/2$ die Fläche von Δ ist. Jetzt hat aber das Parallelogramm mit den Eckpunkten B, C, D und E den doppelten Flächeninhalt wie Δ – also a^2. Darüber hinaus stimmt der Flächeninhalt dieses Parallelogramms mit dem Flächeninhalt des Rechtecks mit Seitenlängen c (Hypotenusenlänge von Δ)

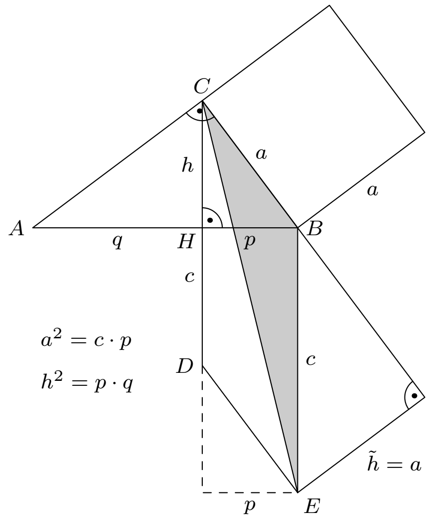

Abb. 4.8 Bezeichnungen zur Herleitung des Satzes des Euklids und des Höhensatz.

und p (Länge des Hypotenusenabschnitts) überein. Es ergibt sich also die folgende Aussage:

> **Satz des Euklid**
>
> Für rechtwinklige Dreiecke ist das Quadrat der Länge einer Kathete a gleich dem Produkt der Länge der Hypotenuse c mit der Länge des **Hypotenusenabschnitts** p, der einen gemeinsamen Punkt mit dieser Kathete hat: $a^2 = cp$.

Eine weitere interessante Beziehung von Flächen am rechtwinkligen Dreieck wird durch den **Höhensatz** ausgedrückt. Betrachte ich nochmals Abbildung 4.8 und beachte ich, dass das Dreieck mit den Eckpunkten H, B und C rechtwinklig ist, so stelle ich zunächst fest, dass nach dem Satz des Pythagoras $h^2 + p^2 = a^2$ gilt. Nach dem Satz von Euklid kann ich hier nun a^2 durch cp ersetzen und ich erhalte, indem ich in dieser Gleichung p^2 auf die rechte Seite schreibe,

$$h^2 = cp - p^2.$$

Da sich die Länge der Hypotenuse c als Summe der Längen der Hypotenusenabschnitte p und q schreiben lässt, $c = p + q$, folgt schließlich

$$h^2 = (p + q)p - p^2 = p^2 + qp - p^2 = qp.$$

> **Höhensatz**
>
> Für rechtwinklige Dreiecke ist das Quadrat der Höhe h auf der Hypotenuse gleich dem Produkt der beiden Hypotenusenabschnitte p und q: $h^2 = pq$.

Mit den drei obigen Sätzen lassen sich in einfacher Weise aus gewissen vorgegebenen Größen die Längen der verbleibenden Größen am rechtwinkligen Dreieck bestimmen.

Beispiel 4.2:
Weiß ich beispielsweise, dass für ein rechtwinkliges Dreieck Δ die Hypotenuse die Länge $c = 6$ hat und die beiden Hypotenusenabschnitte die Länge $p = 2$ und $q = 4$ haben, so berechne ich mit dem Höhensatz die Länge h der Höhe auf die Hypotenuse aus $h^2 = 2 \cdot 4 = 8$ als $h = 2\sqrt{2}$. Weiter bestimme ich dann nach dem Satz des Euklids die Länge der Kathete, die mit dem Hypotenusenabschnitt der Länge p einen gemeinsamen Punkt hat, als $a = \sqrt{p \cdot c} = \sqrt{2 \cdot 6} = \sqrt{12} = 2\sqrt{3}$. Schließlich kann ich die Länge b der verbleibenden Hypotenuse mit dem Satz des Pythagoras berechnen: $b = \sqrt{c^2 - a^2} = \sqrt{36 - 12} = \sqrt{24} = 2\sqrt{6}$. ■

Beispiel 4.3:
Ich betrachte ein rechtwinkliges Dreieck, für das ich die Länge eines Hypotenusenabschnitts $p = 1$ sowie die Länge $b = \sqrt{2}$ der Kathete, die *keinen* gemeinsamen Punkt mit diesem Hypotenusenabschnitt hat, kenne. Jetzt wird die Bestimmung der verbleibenden Längen etwas schwieriger, weil jede der durch die letzten drei Sätze gegebenen Formeln noch zwei unbekannte Faktoren enthält. Ein kleiner Trick hilft mir hier weiter: Ich erinnere mich zunächst, dass $c = p + q$ gilt, das heißt, den Hypotenusenabschnitt der Länge q, der einen gemeinsamen Punkt mit der Kathete der Länge b hat, kann ich als $q = c - p = c - 1$ schreiben. Der Satz des Euklids liefert mir nun $b^2 = qc = (c-1)c = c^2 - c$ und wegen $b^2 = 2$ gelange ich auf die **quadratische Gleichung** $c^2 - c - 2 = 0$. Diese kann ich mit der **p-q-Formel** auflösen und ich erhalte die Lösungen $c_1 = 2$ und $c_2 = -1$. Da nur positive Längen erlaubt sind, ist also $c = c_1 = 2$. Jetzt könnte ich die verbleibenden Größen q, h und a leicht bestimmen. Das überlasse ich aber Ihnen: ■

Übungsaufgabe 4.5:
Führen Sie Beispiel 4.2 fort, indem Sie dort die noch fehlenden Größen q, h und a bestimmen. ■

Eine weitere Standardaussage über rechtwinklige Dreiecke habe ich noch auf Lager. Hierzu betrachte ich die drei Mittelsenkrechten eines *rechtwinkligen* Dreiecks. Oben hatte ich Ihnen erklärt, dass diese sich für jedes Dreieck im Mittelpunkt S des Umkreises schneiden. Die Besonderheit rechtwinkliger Dreiecke Δ ist es, dass dieser Punkt S nun der Mittelpunkt der Hypotenuse von Δ ist (siehe Abbildung 4.9 (links)). Die Länge des Radius des Umkreis solcher Dreiecke ist also die halbe Länge der Hypotenuse: $c/2$. Entscheide ich mich lediglich, die zwei Eckpunkte A und B der Hypotenuse eines rechtwinkligen Dreiecks festzulegen, so liegt also der verbleibende Eckpunkt C stets auf einem Kreis mit Radius $c/2$ um den Mittelpunkt $S = (A + B)/2$ (siehe Abbildung 4.9 (rechts)).

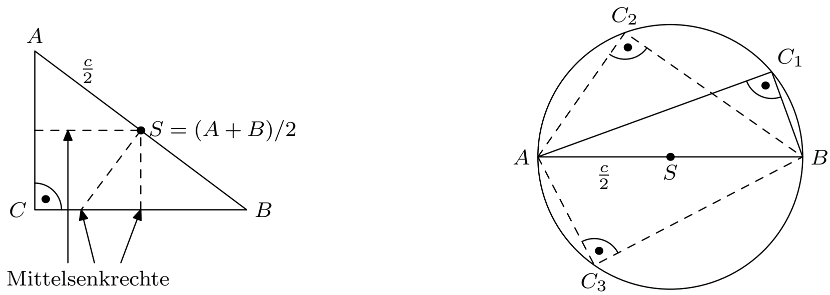

Abb. 4.9 Für rechtwinklige Dreiecke ist der Mittelpunkt des Umkreises (Schnitt der Mittelsenkrechten) der Mittelpunkt $S = (A + B)/2$ von deren Hypotenuse (links). Der Satz von Thales (rechts) besagt, dass der verbleibende Eckpunkt C stets auf dem gemeinsamen Umkreis solcher Dreiecke liegt.

> **Satz von Thales**
>
> Für rechtwinklige Dreiecke mit gemeinsamer Hypotenuse (der Länge c gebildet von den Eckpunkten A und B) liegt der verbleibende Eckpunkt C stets auf dem gemeinsamen Umkreis dieser Dreiecke. Dieser Umkreis hat den Mittelpunkt $S = (A + B)/2$ und den Radius $c/2$.

Eine Aussage, die mit dem Satz von Thales verwandt ist, erkläre ich Ihnen jetzt anhand von Abbildung 4.10. Diese besagt: Die **Peripheriewinkel** über einem festen Kreisbogen sind stets gleich. Ich werde auf diese später noch einmal eingehen, wenn ich über den **Sinussatz** sprechen werde, der die Innenwinkel in Relation zu den Seitenlängen des allgemeinen Dreiecks setzt.

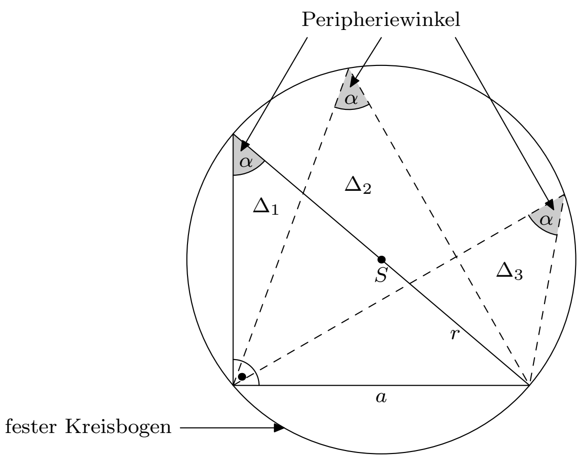

Abb. 4.10 Die Peripheriewinkel über einem festen Kreisbogen sind stets gleich. Die Dreiecke Δ_1, Δ_2 und Δ_3 haben die gemeinsame Seite mit Länge a über dem festen Kreisbogen – deshalb ist der dieser Seite gegenüberliegende Winkel stets dasselbe α.

Apropos allgemeines Dreieck. Ist es Ihnen auch aufgefallen? Stimmt, in den bisherigen Aussagen habe ich mich noch so gut wie gar nicht mit den (Innen-)Winkel am Dreieck beschäftigt. Ich bin bis jetzt oftmals von rechtwinkligen Dreiecken ausgegangen und die verbleibenden beiden Winkel $\neq 90°$ in diesen Dreiecken haben mich weniger interessiert – es reichte mir, aus gegebenen Größen, zum Beispiel die Längen der Katheten a und b, die wesentlichen verbleibenden Größen am Dreieck, beispielsweise die Länge der Hypotenuse $c = \sqrt{a^2 + b^2}$, zu bestimmen. Damit möchte ich mich jetzt aber nicht mehr zufrieden geben, denn erstens könnte es auch für rechtwinklige Dreiecke sinnvoll sein, die verbleibenden beiden Winkel $\neq 90°$ auszurechen, und zweitens möchte ich natürlich auch für allgemeine Dreiecke ähnliche Berechnungen wie für die rechtwinkligen Dreiecke durchführen können. Es wird sich herausstellen, dass ich dies in bequemer Weise tun kann, wenn ich die Winkel am Dreieck mit einbeziehe. Hierfür benötige ich allerdings **trigonometrische Funktionen**. Interessanterweise gelange ich auf diese, indem ich zunächst wieder lediglich *rechtwinklige* Dreiecke wie in Abbildung 4.6 (rechts) betrachte.

Sinus, Cosinus und Tangens

Der Quotient aus der Länge der **Gegenkathete** a und der Länge der Hypotenuse c wird als **Sinus** des Winkels α am Eckpunkt A bezeichnet: $\sin(\alpha) = a/c$. Der Quotient aus der Länge der **Ankathete** b und der Länge der Hypotenuse c wird als **Cosinus** des Winkels α am Eckpunkt A bezeichnet: $\cos(\alpha) = b/c$. Der Quotient aus der Länge der **Gegenkathete** a und der Länge der **Ankathete** b wird als **Tangens** des Winkels α am Eckpunkt A bezeichnet: $\tan(\alpha) = a/b$.

Beispiel 4.4:
Ist Δ ein rechtwinkliges und gleichschenkliges Dreieck mit Kathetenlängen $a = b = 1$ sowie Hypotenusenlänge $c = \sqrt{2}$, so gelten $\sin(\alpha) = \cos(\alpha) = 1/\sqrt{2} = \sqrt{2}/2$ sowie $\tan(\alpha) = 1$ und $\alpha = 45°$. Diesen Winkel bekomme ich, indem ich auf dem Taschenrechner $\sqrt{2}/2$ eingebe und dann (im Gradmaß) beispielsweise INV und COS in dieser Reihenfolge drücke. ■

Beispiel 4.5:
Betrachte ich ein rechtwinkliges Dreieck Δ mit den verbleibenden Innenwinkeln $\alpha = 30°$ und $\beta = 60°$ sowie der Hypotenusenlänge $c = 2$, so berechne ich die beiden verbleibenden Kathetenlängen beispielsweise wie folgt: $a = c \cdot \sin(\alpha) = 2 \cdot \sin(30°) = 2 \cdot 1/2 = 1$ und $b = c \cdot \sin(\beta) = 2 \cdot \sin(60°) = 2 \cdot \sqrt{3}/2 = \sqrt{3}$. ■

Übungsaufgabe 4.6:
In einem rechtwinkligen Dreieck seien die Längen der Katheten bekannt: $a = 3\sqrt{3}$ und $b = 3$. Bestimmen Sie die beiden verbleibenden Innenwinkel α und β sowie die Hypotenusenlänge c und den Flächeninhalt F dieses Dreiecks. ■

Eine Standardanwendung des Sinus für allgemeine Dreiecke liegt vor, wenn die Länge einer Höhe h auf einer Seite mit Länge c bestimmt werden soll. Kennt man nämlich noch die Länge einer anderen Seite a des Dreiecks und den von diesen beiden Seiten eingeschlossenen Winkel, so gilt $h = a\ \sin(\alpha)$. Möchte ich dann den Flächeninhalt F des Dreiecks bestimmen, so kann ich diesen dann gemäß

$$F = \frac{hc}{2} = \frac{ac\ \sin(\alpha)}{2}$$

berechnen.

Übungsaufgabe 4.7:
Für ein Dreieck seien die Längen zweier Seiten $a = 4$ und $c = 5$ sowie der von diesen Seiten eingeschlossene Winkel $\alpha = 30°$ bekannt. Bestimmen Sie den Flächeninhalt F dieses Dreiecks. ■

Aus den Festlegungen der trigonometrischen Funktionen ist sofort ersichtlich, dass

$$\frac{\sin(\alpha)}{\cos(\alpha)} = \frac{\frac{a}{c}}{\frac{b}{c}} = \frac{a}{b} = \tan(\alpha)$$

gilt: Der Tangens eines Winkels ist also gerade der Quotient aus dem Sinus und dem Cosinus dieses Winkels. Außerdem erhalte ich aus dem Satz des Pythagoras $(a/c)^2 + (b/c)^2 = 1$ und damit die nächste, wichtige Beziehung:

Summe der Quadrate des Sinus und Cosinus

Für jeden Winkel α gilt:

$$\sin^2(\alpha) + \cos^2(\alpha) = 1.$$

Diese Gleichheit begegnet Ihnen zuweilen in der Form $\sin(\alpha) = \sqrt{1 - \cos^2(\alpha)}$ oder $\cos(\alpha) = \sqrt{1 - \sin^2(\alpha)}$. Keine Angst – hier steht nicht plötzlich etwas Negatives unter der Wurzel, denn aus der letzten Aussage erkenne ich auch, dass die Werte des Sinus und Cosinus stets zwischen -1 und 1 liegen:

$$-1 \leq \sin(\alpha) \leq 1 \text{ (kurz: } |\sin(\alpha)| \leq 1) \text{ und } -1 \leq \cos(\alpha) \leq 1 \text{ (kurz: } |\cos(\alpha)| \leq 1).$$

Denn wäre für einen Winkel α beispielsweise $|\sin(\alpha)| > 1$, so wäre natürlich $\sin^2(\alpha) > 1$, und wegen $\cos^2(\alpha) \geq 0$ könnte damit die Summe aus $\sin^2(\alpha)$ und $\cos^2(\alpha)$ nicht gleich 1 sein. Dies kann aber nicht sein, denn sonst würde mit dem Satz des Pythagoras etwas nicht stimmen – und dies würde die Welt der Mathematiker erdbebengleich erschüttern.

Ich sagte es bereits: Mithilfe der trigonometrischen Funktionen plane ich, Relationen für die Größen am allgemeinen Dreieck zu finden. Allerdings steht mir derzeit noch ein kleineres Hindernis im Weg, um dies sofort in erfolgreicher Weise zu tun. Oben habe ich Ihnen nämlich gesagt, dass die beiden verbleibenden Winkel im rechtwinkligen Dreieck nur noch Werte (echt) zwischen 0° und 90° annehmen können. Wenn ich aber plane, zum Beispiel den Sinus für allgemeine, also beispielsweise auch stumpfwinklige, Dreiecke zu verwenden, so wäre es dazu nützlich zu wissen, wie sich der

Sinus auf beliebige Winkel erweitern lässt. Um dies zu erfahren, schaue ich mir einen **Einheitskreis** wie in Abbildung 4.11 an. Für Winkel α zwischen 0° und 90° ist der Sinus, wie die Abbildung zeigt, gerade die Länge der dem Winkel α *gegen*überliegenden *Kathete* (deshalb: *Gegenkathete*), denn hier ist die Länge der Hypotenuse 1. Analog ist für diese Winkel der Cosinus gerade die Länge der dem Winkel α *an*liegenden Kathete (deshalb: *Ankathete*).

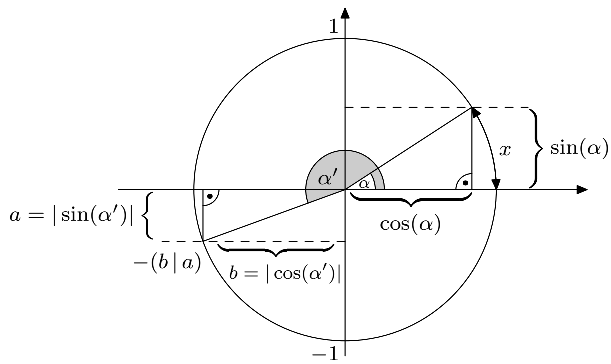

Abb. 4.11 Veranschaulichung des Sinus und Cosinus am Einheitskreis.

Möchte ich nun den Wert des Sinus beziehungsweise Cosinus für einen beliebigen anderen Winkel ermitteln (ich habe ein Beispiel solcher Winkel in Abbildung 4.11 mit α' bezeichnet), so kann ich dies tun, indem ich den dem Winkel entsprechenden Punkt $(b|a)$ auf dem Einheitskreis betrachte und $\sin(\alpha') = a$ sowie $\cos(\alpha') = b$ setze. Für solche allgemeineren Winkel α' lege ich den Tangens dann durch Beziehung $\tan(\alpha') = \sin(\alpha')/\cos(\alpha') = a/b$ ebenfalls fest. Der Wert des Tangens ist die Steigung der Geraden durch die beiden Punkte $(0|0)$ und $(b|a)$. Ein bisschen aufpassen muss ich hierbei allerdings: Da der Cosinus von 90° und 270° den Wert 0 hat, darf ich diese Formel des Tangens für diese speziellen Winkel nicht anwenden, denn Sie wissen ja, dass ich *nie nie nie* durch 0 teilen darf. Zu erwähnen ist hier noch, dass die obigen Beziehungen $\sin^2(\alpha') + \cos^2(\alpha') = 1$ und $|\sin(\alpha')| \leq 1$ sowie $|\cos(\alpha')| \leq 1$ für solche allgemeineren Winkel α' weiterhin uneingeschränkt erfüllt sind.

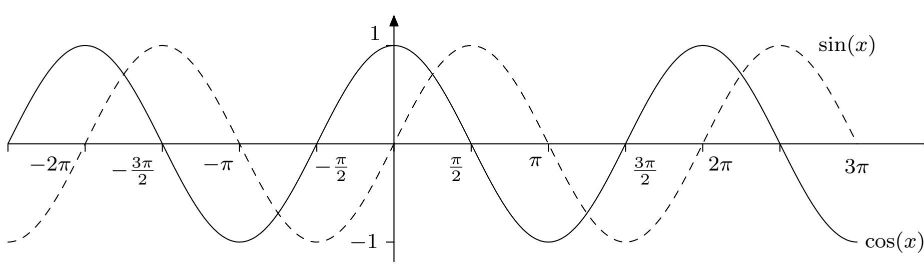

Abb. 4.12 Sinus und Cosinus als Funktion des (reellen) Bogenmaßes x.

Ein weiterer Punkt mag dem aufmerksamen Leser beim Vergleich mit den Inhalten von Kapitel 2 in den Sinn kommen. Ich spreche hier von trigonometrischen *Funktionen* und habe bisher – entgegen den üblichen Gepflogenheiten – als **Argumente** Winkel anstatt **reeller Zahlen** verwendet. Das ist natürlich zulässig – aber schön wäre es schon, wenn ich die trigonometrischen Funktionen, wie ich es von anderen Funktionen gewohnt bin, als Funktionen einer **reellen Variablen** – sagen wir x – auffassen könnte. Dies gelingt mir durch Verwendung des **Bogenmaßes**, zu dessen nachfolgender Erklärung es erlaubt ist, nochmals einen flüchtigen Blick auf Abbildung 4.11 zu werfen. Außerdem weise ich auf den nächsten Abschnitt hin, in dem das Bogenmaß nochmals auftreten wird.

Bogenmaß

Das Bogenmaß x eines Winkels α ist die vorzeichenbelegte Länge des zu α gehörigen Bogens am Einheitskreis, wobei das Vorzeichen von x genau dann positiv ist, wenn der Bogen gegen den Uhrzeigersinn abgetragen wird. Die Umrechnungsformel zwischen Grad- und Bogenmaß ist $x = \pi\ \alpha/180°$.

Beispiel 4.6:
Möchte ich den Wert des Bogenmaßes von $\alpha = 360°$ wissen, so berechne ich auf diese Art den Umfang des Einheitskreises $x = \pi\ 360°/180° = 2\pi$ (siehe auch nächster Abschnitt – dort gebe ich Hinweise darauf, warum dies so ist und was die Zahl $\pi \approx 3{,}14$ genau ist). Will ich das Bogenmaß von $\alpha = 45°$ wissen, so ermittle ich $x = \pi\ 45°/180° = \pi/4$ – dies ist ein Achtel des Umfangs des Einheitskreises. Ist $\alpha = -45°$, so erwarte ich, dass das Vorzeichen von x ebenfalls negativ ist, denn um diesen Winkel abzutragen, muss ich mit dem Uhrzeigersinn voranschreiten. Dies ist auch der Fall, denn ich berechne $x = \pi\ (-45°)/180° = -\pi/4$. Gibt mir umgekehrt jemand das Bogenmaß x vor, so bestimme ich den zugehörigen Winkel α gemäß $\alpha = 180°\ x/\pi$. Für $x = \pi/3$ erhalte ich also beispielsweise $\alpha = 180°\ \pi/(3\pi) = 60°$. Das ist – wie ich meine – recht übersichtlich. ■

Trigonometrische Funktionen

Mithilfe des Bogenmaßes kann ich die trigonometrischen Funktionen nun als Funktionen der reellen Variablen x im üblichen Sinne auffassen, indem ich

$$\sin(x) = \sin(\alpha),\ \ \cos(x) = \cos(\alpha) \text{ und } \tan(x) = \tan(\alpha),$$

setze, wobei x das zum Winkel α gehörige Bogenmaß ist.

Die Graphen der **Sinus-** und **Cosinusfunktion** finden Sie in Abbildung 4.12. In Tabelle 4.1 habe ich Ihnen einige Werte der trigonometrischen Funktionen für charakteristische Winkel zusammen mit deren Bogenmaß aufgelistet.

Tab. 4.1 Werte der trigonometrischen Funktionen an charakteristischen Stellen x.

α	$-90°$	$-60°$	$-45°$	$-30°$	$0°$	$30°$	$45°$	$60°$	$90°$
x	$-\pi/2$	$-\pi/3$	$-\pi/4$	$-\pi/6$	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$
$\sin(x)$	-1	$-\sqrt{3}/2$	$-\sqrt{2}/2$	$-1/2$	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1
$\cos(x)$	0	$1/2$	$\sqrt{2}/2$	$\sqrt{3}/2$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	$1/2$	0
$\tan(x)$	–	$-\sqrt{3}$	-1	$-\sqrt{3}/3$	0	$\sqrt{3}/3$	1	$\sqrt{3}$	–

Sie erkennen an den Graphen in Abbildung 4.12, dass die Sinus- und Cosinusfunktion jeweils 2π**-periodisch** sind, das heißt

$$\sin(x) = \sin(x + 2\pi) \text{ und } \cos(x) = \cos(x + 2\pi)$$

für alle x – wenig verwunderlich, denn nach Durchlauf (zum Beispiel gegen den Uhrzeigersinn) des kompletten Rands des Einheitskreises (der, wie Sie im nächsten Abschnitt erfahren werden, die Länge 2π hat) wiederholen sich die Werte des Sinus und Cosinus. Damit sind dies Funktionen, die unendlich viele **Nullstellen** besitzen. Betrachte ich beispielsweise den Sinus, so sehe ich zunächst, dass $\sin(0) = 0$ und $\sin(\pi) = 0$ gelten. Der Sinus ist aber 2π-periodisch und somit gelten auch beispielsweise $\sin(2\pi) = 0$ und $\sin(-\pi) = 0$. Möchte ich *alle* Nullstellen der Sinusfunktion beschreiben, so kann ich dies wie folgt tun:

$$\sin(x) = 0 \text{ genau dann, wenn } x = k\pi, \text{ wobei } k \text{ eine beliebige ganze Zahl ist.}$$

Beispielsweise gelten somit auch $\sin(311\ \pi) = 0$ und $\sin(-1022\ \pi) = 0$. Ähnlich kann ich die Nullstellen des Cosinus beschreiben:

$$\cos(x) = 0 \text{ genau dann, wenn } x = \tfrac{\pi}{2} + k\pi, \text{ wobei } k \text{ eine beliebige ganze Zahl ist.}$$

Übungsaufgabe 4.8:
Beschreiben Sie alle Stellen x, an denen a) die Sinusfunktion den Wert 1 hat (das heißt $\sin(x) = 1$ gilt), b) die Cosinusfunktion den Wert -1 hat (das heißt $\cos(x) = -1$ gilt). ■

Übungsaufgabe 4.9:
Beschreiben Sie alle reellen Zahlen x an denen die Tangensfunktion $\tan(x) = \sin(x)/\cos(x)$ sinnvoll definiert ist. ■

Weiterhin ist erwähnenswert, dass der **Graph** der Cosinusfunktion **achsensymmetrisch** zur y-Achse ist,

$$\cos(x) = \cos(-x),$$

für alle x, während der Graph der Sinusfunktion **punktsymmetrisch** zum **Ursprung** $(0|0)$ ist:

$$\sin(x) = -\sin(-x)$$

für alle x. Damit ist die Cosinusfunktion eine **gerade Funktion** und die Sinusfunktion eine **ungerade Funktion**. Zudem kann ich beobachten, dass durch Verschiebung des Graphen des Cosinus um $\pi/2$ entlang der x-Achse der Graph der Sinusfunktion entsteht. In Formeln ausgedrückt kann ich das folgendermaßen hinschreiben:

$$\cos\Big(x - \frac{\pi}{2}\Big) = \sin(x) \text{ beziehungsweise } \cos(x) = \sin\Big(x + \frac{\pi}{2}\Big)$$

für alle x. Die trigonometrischen Funktionen erfüllen eine ganze Reihe weiterer Beziehungen, die man als Normalbürger nicht sofort schnell erkennen kann. Beispielsweise erfordert der Nachweis der folgenden Beziehungen etwas aufwendigere Argumente, die ich Ihnen jetzt aber ersparen werde.

Additionstheoreme des Sinus und Cosinus

Für alle x und y gelten die folgenden Beziehungen:

$$\begin{aligned} \sin(x+y) &= \sin(x)\cos(y) + \cos(x)\sin(y) \quad \text{(1. Additionstheorem)} \\ \cos(x+y) &= \cos(x)\cos(y) - \sin(x)\sin(y) \quad \text{(2. Additionstheorem)} \end{aligned}$$

Beispiel 4.7:
Ich setze $x = -\pi/3$ und $y = \pi/3$ und verwende Tabelle 4.1. Offenbar gelten $\sin(x + y) = \sin(-\pi/3 + \pi/3) = \sin(0) = 0$ und $\cos(x + y) = \cos(-\pi/3 + \pi/3) = \cos(0) = 1$. Andererseits erhalte ich für den komplizierteren Term auf der rechten Seite des ersten Additionstheorems

$$\sin(x)\cos(y) + \cos(x)\sin(y) = \overbrace{\sin\Big(-\frac{\pi}{3}\Big)}^{-\frac{\sqrt{3}}{2}}\overbrace{\cos\Big(\frac{\pi}{3}\Big)}^{\frac{1}{2}} + \overbrace{\cos\Big(-\frac{\pi}{3}\Big)}^{\frac{1}{2}}\overbrace{\sin\Big(\frac{\pi}{3}\Big)}^{\frac{\sqrt{3}}{2}} = 0$$

beziehungsweise des zweiten Additionstheorems

$$\cos(x)\cos(y) - \sin(x)\sin(y) = \overbrace{\cos\Big(-\frac{\pi}{3}\Big)}^{\frac{1}{2}}\overbrace{\cos\Big(\frac{\pi}{3}\Big)}^{\frac{1}{2}} - \overbrace{\sin\Big(-\frac{\pi}{3}\Big)}^{-\frac{\sqrt{3}}{2}}\overbrace{\sin\Big(\frac{\pi}{3}\Big)}^{\frac{\sqrt{3}}{2}} = \frac{1}{4} - \Big(-\frac{3}{4}\Big) = 1.$$

■

Mit den Additionstheoremen kann ich viele andere Beziehungen der Sinusfunktion mit der Cosinusfunktion ableiten. Setze ich beispielsweise den Term $-y$ für y in das erste Additionstheorem ein, so erhalte ich $\sin(x-y) = \sin(x)\cos(-y)+\cos(x)\sin(-y)$, und weil der Cosinus eine gerade Funktion und der Sinus eine ungerade Funktion ist, folgt

$$\sin(x-y) = \sin(x)\cos(y) - \cos(x)\sin(y)$$

für alle x und y. Genauso erhalte ich mit dem zweiten Additionstheorem

$$\cos(x-y) = \cos(x)\cos(y) + \sin(x)\sin(y)$$

für alle x und y. Interessant ist, dass, wenn ich in der letzten Gleichung nun $y = x$ setze, $\cos(x-x) = \cos(x)\cos(x) + \sin(x)\sin(x)$, also $1 = \cos^2(x) + \sin^2(x)$, folgt – eine Beziehung, die ich Ihnen weiter oben schon einmal auf andere Art für lediglich gewisse Winkel α gezeigt habe. Mache ich Ähnliches in den Additionstheoremen – setze ich dort also $y = x$ – so sehe ich, dass die Beziehungen

$$\sin(2x) = 2\ \sin(x)\cos(x) \text{ und } \cos(2x) = \cos^2(x) - \sin^2(x)$$

beziehungsweise

$$\frac{\sin(2x)}{2} = \sin(x)\cos(x) \text{ und } \sin^2(x) = \cos^2(x) - \cos(2x)$$

erfüllt sind. Ersetze ich in der letzten Gleichung $\sin^2(x)$ durch $1-\cos^2(x)$, so gelange ich auf $1 - \cos^2(x) = \cos^2(x) - \cos(2x)$ und hieraus erhalte ich

$$\cos(2x) = 2\ \cos^2(x) - 1.$$

Wenn ich möchte, so kann ich mit diesen Gleichungen den Tangens (an den erlaubten Stellen x, siehe Übungsaufgabe 4.9) anders ausdrücken:

$$\tan(x) = \tan\left(2\ \tfrac{x}{2}\right) = \frac{\sin\left(2\ \frac{x}{2}\right)}{\cos\left(2\ \frac{x}{2}\right)} = \frac{2\ \sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right)}{2\ \cos^2\left(\frac{x}{2}\right) - 1}.$$

Schließlich werden möglicherweise die Formeln

$$\sin(x) + \sin(y) = 2\ \sin\left(\tfrac{x+y}{2}\right)\cos\left(\tfrac{x-y}{2}\right)$$

und

$$\cos(x) + \cos(y) = 2\ \cos\left(\tfrac{x+y}{2}\right)\cos\left(\tfrac{x-y}{2}\right)$$

Ihren zukünftigen (Studiums-)Weg kreuzen. Keine Angst – auch diese Gleichungen lassen sich ganz einfach aus den Additionstheoremen ableiten. Ich zeige Ihnen das anhand der ersten Formel – Sie dürfen dann die zweite Formel herleiten. In beiden Herleitungen macht man denselben kleinen Trick. Ich schreibe nämlich

$$x = \frac{x+y}{2} + \frac{x-y}{2}$$

und wende dann das erste Additionstheorem an

$$\sin(x) = \sin\left(\tfrac{x+y}{2} + \tfrac{x-y}{2}\right) = \sin\left(\tfrac{x+y}{2}\right)\cos\left(\tfrac{x-y}{2}\right) + \cos\left(\tfrac{x+y}{2}\right)\sin\left(\tfrac{x-y}{2}\right).$$

Außerdem benutze ich $y = (x+y)/2 + (y-x)/2$, um in derselben Weise

$$\sin(y) = \sin\left(\tfrac{x+y}{2} + \tfrac{y-x}{2}\right) = \sin\left(\tfrac{x+y}{2}\right)\cos\left(\tfrac{y-x}{2}\right) + \cos\left(\tfrac{x+y}{2}\right)\sin\left(\tfrac{y-x}{2}\right)$$

zu erkennen. Beachte ich, dass wegen der Symmetrie-Eigenschaften des Graphen der Sinus- und Cosinusfunktion $\sin((y-x)/2) = -\sin((x-y)/2)$ und $\cos((y-x)/2) = \cos((x-y)/2)$ gelten, so sehe ich, dass bei der Addition von $\sin(x)$ und $\sin(y)$ in diesen Darstellungen tatsächlich die erste der obigen Formeln herauskommt.

Übungsaufgabe 4.10:
Leiten Sie die zweite der obigen Formeln her. ■

Natürlich macht es viel Spaß, mit Ihnen über trigonometrische Funktionen zu plaudern, und wir könnten unsere Konversation – wenn es nur nach mir ginge – gerne noch ein bisschen fortsetzen. So ließe sich beispielsweise über die interessante Beziehung

$$e^{\sqrt{-1}\,\alpha} = \cos(\alpha) + \sqrt{-1}\ \sin(\alpha)$$

mit der **Euler'schen Zahl** e sicherlich einiges sagen und ich könnte Ihnen damit beispielsweise elegant zeigen, wie man die Additionstheoreme herleitet. **Komplexe Zahlen** (das sind grob gesprochen Zahlen, bei denen das Ziehen der Wurzeln aus negativen Zahlen, also auch aus der -1, erlaubt ist) sind jedoch in diesem Kapitel noch nicht unser Thema (siehe Kapitel 8), und wir sollten unsere gemeinsamen Ziele nicht aus dem Auge verlieren. Oben hatte ich Ihnen nämlich vor Längerem versprochen, dass ich die trigonometrischen Funktionen verwenden werde, um die Größen am allgemeinen Dreieck in übersichtliche Relationen zu stellen, wobei ich insbesondere die (Innen-)Winkel am Dreieck mit einbeziehen möchte. In diesem Sinne werde ich jetzt fortfahren.

Ich betrachte also im Folgenden ein beliebiges Dreieck mit Eckpunkten A, B und C und Seitenlängen a, b und c wie in Abbildung 4.13, und mein Ziel ist es, nun zunächst eine Beziehung zwischen diesen Größen und dem Innenwinkel α bei A herzustellen. Hierzu betrachte ich die Höhe auf die Seite, die A und B verbindet. Diese Höhe hat die Länge h, schneidet diese Seite mit Länge c in einem Punkt D und ich erhalte zwei rechtwinklige (Teil-)Dreiecke mit Eckpunkten A, D und C, beziehungsweise B, D und C. Verwende ich die Bezeichnungen aus Abbildung 4.13, so besagt der **Satz des Pythagoras**, dass die Gleichungen $p^2 + h^2 = b^2$ und $q^2 + h^2 = a^2$ erfüllt sein

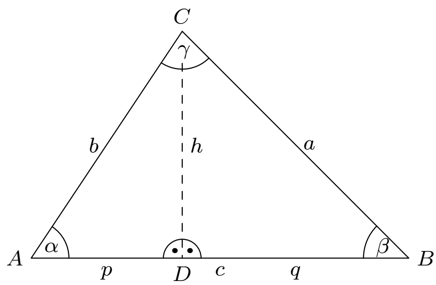

Abb. 4.13 Bezeichnungen zur Herleitung des Cosinussatzes.

müssen. Da $q = c-p$ gilt, berechne ich mit der **Binomischen Formel** $q^2 = (c-p)^2 = c^2-2\,pc+p^2$ und ich erhalte durch Einsetzen in die zweite Gleichung $c^2-2\,pc+p^2+h^2 = a^2$. Dies kann ich durch Verwendung der ersten Gleichung $(p^2 + h^2 = b^2)$ weiter vereinfachen:

$$b^2 + c^2 - 2\,pc = a^2.$$

Diese Gleichung stellt zwar die Längen der Seiten des allgemeinen Dreiecks in Beziehung – unglückseligerweise taucht hier allerdings noch die Länge des Abschnitts p auf: Eine Größe, deren Kenntnis ich hier nicht vorausgesetzt habe. Dementgegen gehe ich hier davon aus, dass ich den Winkel α, beziehungsweise $\cos(\alpha)$, kenne. Deshalb ersetze ich die mir unbekannte Größe p durch

$$p = b\cos(\alpha).$$

Das darf ich, denn die Strecke von A nach D ist die Ankathete im rechtwinkligen Dreieck mit dem zusätzlichen Eckpunkt C und der Cosinus von α ist bekanntlich durch die Formel $\cos(\alpha) = p/b$ festgelegt. Insgesamt erhalte ich jetzt

$$b^2 + c^2 - 2\,bc\,\cos(\alpha) = a^2.$$

Das ist die Gleichung, nach der ich gesucht habe, denn die drei Längen a, b und c der Seiten des allgemeinen Dreiecks stehen hier in Beziehung zu einem Winkel α – genauer: dem Winkel des Dreiecks, den die Seiten mit Länge b und c einschließen.

Cosinussatz

Für beliebige Dreiecke ist das Quadrat der Länge a einer Seite die Differenz der Summe der Quadrate der Längen b, c der verbleibenden Seiten mit dem Term $2\,bc\cos(\alpha)$, wobei α der von den Seiten mit Längen b und c eingeschlossene Winkel ist: $a^2 = b^2 + c^2 - 2\,bc\cos(\alpha)$.

Genauso wie mit dem Winkel α kann ich natürlich auch für die Winkel β und γ in Abbildung 4.13 argumentieren, und somit sehe ich ein, dass auch die Formeln

$$b^2 = a^2 + c^2 - 2\,ac\,\cos(\beta) \text{ und } c^2 = a^2 + b^2 - 2\,ab\,\cos(\gamma)$$

gültig sind. Ich sollte jetzt nicht unerwähnt lassen, dass es sich bei dem Cosinussatz um eine *Verallgemeinerung* des **Satzes von Pythagoras** für allgemeine Dreiecke handelt. Liegt der Spezialfall eines rechtwinkligen Dreiecks mit Eckpunkten A, B und C vor – ist also beispielsweise $\gamma = 90°$, so ist $\cos(\gamma) = \cos(90°) = 0$, und der Term $2\,ab\,\cos(\gamma)$ wird ebenfalls 0 sein. In diesem Fall ergibt der Cosinussatz also $c^2 = a^2+b^2$, was nichts anderes als die Aussage des Satzes von Pythagoras für rechtwinklige Dreiecke ist.

Beispiel 4.8:
Ein klassisches Beispiel zur Anwendung des Cosinussatzes liegt vor, wenn die Länge zweier Seiten eines Dreiecks, sagen wir $b = 2$ und $c = 3$, und der von diesen Seiten eingeschlossene Winkel, sagen wir $\alpha = 30°$, vorgegeben sind, und man gerne die Länge der verbleibenden Seite a ausrechnen möchte. Die Formel ergibt in diesem Fall

$$a^2 = 2^2 + 3^2 - 2 \cdot 2 \cdot 3 \cdot \cos(30°) = 4 + 9 - 12 \cdot \sqrt{3}/2 = 13 - 6 \cdot \sqrt{3}$$

und somit $a = \sqrt{13 - 6 \cdot \sqrt{3}} \approx 1{,}6148$. Umgekehrt kann ich mich jetzt fragen, wie wohl die beiden verbleibenden Winkel β und γ in diesem Dreieck sind. Hierzu stelle ich eine der beiden verbleibenden obigen Formeln um – $\cos(\beta) = (a^2 + c^2 - b^2)/(2\,ac)$ – und setze die Längen der Seiten des Dreiecks ein:

$$\cos(\beta) = \frac{13 - 6 \cdot \sqrt{3} + 9 - 4}{2\,(\sqrt{13 - 6 \cdot \sqrt{3}})3} = \frac{3 - \sqrt{3}}{\sqrt{13 - 6\sqrt{3}}} \approx 0{,}7852\ .$$

Drücken von INV und COS auf dem Taschenrechner ergibt $\beta \approx 38{,}26°$. Den verbleibenden Winkel γ könnte ich jetzt genauso berechnen. Ich mache mir dies aber leichter, indem ich mich daran erinnere, dass die Summe der Innenwinkel im Dreieck stets 180° ist. Somit ist hier der verbleibende Winkel γ in etwa $180° - 30° - 38{,}26° = 111{,}74°$ ist. Es handelt sich also in diesem Beispiel um ein stumpfwinkliges Dreieck. ■

Der Cosinussatz lehrt uns also, dass durch Vorgabe der Längen zweier Seiten a und b eines Dreiecks und dem von diesen eingeschlossenen Winkel α die verbleibende Seitenlänge sowie die beiden verbleibenden Winkel dieses Dreiecks festgelegt sind. Anders ausgedrückt bedeutet dies, dass durch diese drei Größen a, b und α bereits die Gestalt des Dreiecks Δ vollkommen festgelegt ist. Ich kann zwar weitere Dreiecke dieser Gestalt und Größe erzeugen, indem ich Δ in der Ebene verschiebe oder um einen Punkt drehe – so entstehende Dreiecke (man nennt sie zuweilen **kongruente Dreiecke**, siehe Abbildung 4.14) werden aber stets dieselbe Gestalt wie Δ aufweisen.

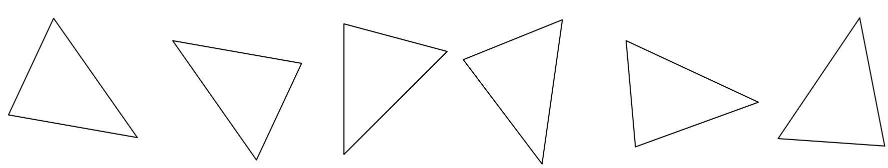

Abb. 4.14 Kongruenz von Dreiecken: Die drei Seitenlängen und die drei (Innen-)Winkel stimmen überein.

Ich stelle somit zunächst fest, dass zwei Dreiecke kongruent sind, falls die Längen zweier ihrer Seiten sowie der zwischen diesen Seiten liegende Innenwinkel übereinstimmen.

An der Stelle hat man vielleicht den Verdacht, dass es noch andere Situationen gibt, in denen Dreiecke kongruent sind. In der Tat: Zwei Dreiecke sind beispielsweise ebenfalls kongruent, falls die Längen ihrer drei Seiten übereinstimmen. Wie kann ich das sehen? Nun, die Antwort liegt ebenfalls im Cosinussatz begründet. Ich mache dazu ein weiteres Beispiel:

Beispiel 4.9:
In diesem Beispiel gebe ich mir die drei Seitenlängen eines Dreiecks vor – sagen wir $a = 5, b = 3$ und $c = 4$. Indem ich die Formel aus dem Cosinussatz umstelle, kann ich nun die drei (Innen-)Winkel des Dreiecks berechnen. Beispielsweise ist $\cos(\alpha) = (b^2 + c^2 - a^2)/(2\,bc)$. Ich berechne also $\cos(\alpha) = (9 + 16 - 25)/24 = 0$ und damit ist $\alpha = 90°$. Es handelt sich also zufällig um ein rechtwinkliges Dreieck. Den zweiten Winkel β berechne ich gemäß $\cos(\beta) = (a^2 + c^2 - b^2)/(2\,ac) = (25 + 16 - 9)/40 = 0{,}8$. Es ergibt sich $\beta \approx 36{,}87°$ und damit ist der verbleibende Winkel $\gamma \approx 53{,}13°$. ■

Übungsaufgabe 4.11:
Ein vorgegebenes Dreieck habe die Seitenlängen $2y, 2y$ und $3y$ für eine reelle Zahl $y \neq 0$. Zeigen Sie, dass die drei (Innen-)Winkel des Dreiecks unabhängig von y sind und bestimmen Sie diese. ■

Schnell einzusehen ist, dass bei der Vorgabe von drei Winkeln α, β und γ (mit $\alpha + \beta + \gamma = 180°$) ein Dreieck im Sinne der Kongruenz noch nicht festgelegt ist, denn ein solches Dreieck kann ich ja nach Belieben vergrößern und verkleinern. Eine andere Frage ist für Sie ähnlich leicht beantwortbar:

Übungsaufgabe 4.12:
Kann man zu vorgegebenen Längen a, b und c eigentlich immer ein Dreieck mit diesen Seitenlängen konstruieren? Begründen Sie Ihre Antwort. ■

Ziemlich am Anfang dieses Abschnitts habe ich Ihnen versprochen, dass ich Ihnen zwei Aussagen hinsichtlich **gleichschenkliger** und **gleichseitiger** Dreiecke zu einem späteren Zeitpunkt ganz einfach zeigen werden. Dieses Versprechen löse ich nun gerne ein. Liegt ein gleichschenkliges Dreieck vor, so sind zwei Seiten des Dreiecks gleich lang – es gilt also beispielsweise $b = a$. Der Cosinussatz ergibt dann

$$\cos(\alpha) = \frac{b^2 + c^2 - a^2}{2\,bc} = \frac{a^2 + c^2 - a^2}{2\,ac} = \frac{c}{2a}$$

und

$$\cos(\beta) = \frac{a^2 + c^2 - b^2}{2\,ac} = \frac{a^2 + c^2 - a^2}{2\,ac} = \frac{c}{2a}$$

und ich erkenne, dass die anliegenden Innenwinkel α, β der verbleibenden Seite übereinstimmen: $\alpha = \beta$. Liegt ein gleichseitiges Dreieck vor, so gilt $a = b = c$ und ich kann die letzten beiden Gleichungen noch weiter vereinfachen: $\cos(\alpha) = c/2a = a/2a = 1/2$

und $\cos(\beta) = 1/2$. Damit gilt $\alpha = \beta = 60°$ und der verbleibende Winkel ist ebenfalls $\gamma = 60°$.

Eine weitere nützliche Beziehung der Längen der Seiten mit den (Innen-)Winkeln eines beliebigen, vorgegebenen Dreiecks Δ lässt sich durch nochmaliges Betrachten von Abbildung 4.13 herleiten. Dort sind die beiden Teildreiecke mit den Eckpunkten A, D und C beziehungsweise B, D und C rechtwinklig. Der Sinus des Winkels bei A beziehungsweise bei B ist daher der Quotient der Längen von Gegenkathete und Hypotenuse dieser Dreiecke. Ich erhalte also

$$\sin(\alpha) = \frac{h}{b} \quad \text{und} \quad \sin(\beta) = \frac{h}{a} \,.$$

Die Länge der Höhe h auf die Dreiecksseite mit Länge c kann ich also auf zwei verschiedene Arten hinschreiben:

$$h = b \cdot \sin(\alpha) \quad \text{und} \quad h = a \cdot \sin(\beta) \ .$$

Da dieses h aber einen eindeutigen Wert haben muss, erhalte ich die Beziehung

$$b \cdot \sin(\alpha) = a \cdot \sin(\beta)$$

für das vorgegebene Dreieck Δ. Teile ich hier durch $sin(\alpha)$ und $\sin(\beta)$, so kann ich das auch wie folgt hinschreiben:

$$\frac{a}{\sin(\alpha)} = \frac{b}{\sin(\beta)} \ .$$

Durch genau dieselbe Vorgehensweise kann ich durch Verwendung der Höhe auf die Seite der Länge b genauso einsehen, dass auch

$$\frac{a}{\sin(\alpha)} = \frac{c}{\sin(\gamma)}$$

gilt.

Sinussatz

Für beliebige Dreiecke mit Seitenlängen a, b und c und den entsprechenden, gegenüberliegenden Innenwinkeln α, β und γ gilt: $a/\sin(\alpha) = b/\sin(\beta) = c/\sin(\gamma)$.

Anders ausgedrückt besagt der Sinussatz, dass der Quotient zweier Seitenlängen eines Dreiecks denselben Wert wie der Quotient der Sinusse der gegenüberliegenden Innenwinkel (auch: **Gegenwinkel**) ergibt:

$$\frac{a}{b} = \frac{\sin(\alpha)}{\sin(\beta)}, \ \frac{b}{c} = \frac{\sin(\beta)}{\sin(\gamma)} \ \text{und} \ \frac{c}{a} = \frac{\sin(\gamma)}{\sin(\alpha)}.$$

Bemerken möchte ich hier, dass der Sinussatz auch aus der Tatsache folgt, dass der Peripheriewinkel α wie in Abbildung 4.10 über einem festen Bogen des Umkreises eines Dreiecks Δ stets gleich ist. Die Festlegung des Sinus zeigt mir nämlich, dass für das rechtwinklige Dreieck Δ_1 in Abbildung 4.10 (und damit für jedes der dortigen Dreiecke) die Beziehung

$$\sin(\alpha) = \frac{a}{2r}$$

gilt, wobei r der Radius des Umkreises von $\Delta = \Delta_2$ beziehungsweise $\Delta = \Delta_1$ ist. Somit sind auch

$$\sin(\beta) = \frac{b}{2r} \text{ und } \sin(\gamma) = \frac{c}{2r}$$

für jedes Dreieck Δ wie in Abbildung 4.13 erfüllt. Die Auflösung dieser drei Formeln nach r und Gleichsetzen ergibt den Sinussatz.

Der Sinussatz ist besonders brauchbar, wenn zwei (Innen-)Winkel – sagen wir α und β – eines Dreiecks und eine Seitenlänge – sagen wir a – vorgegeben sind. In diesem Fall kann ich den verbleibenden Winkel zunächst leicht durch $\gamma = 180° - \alpha - \beta$ bestimmen und dann ergeben sich die noch fehlenden Seitenlängen b und c durch Umstellen der Formeln des Sinussatzes:

$$b = a\,\frac{\sin(\beta)}{\sin(\alpha)} \text{ und } c = a\,\frac{\sin(\gamma)}{\sin(\alpha)}.$$

Durch Anwendung des Sinussatzes sehe ich also, dass durch Vorgabe der Winkel und einer Seitenlänge eines Dreiecks die Längen der verbleibenden beiden Seiten festgelegt sind. Insbesondere sind zwei Dreiecke kongruent, falls eine ihrer Seitenlängen und alle ihre Innenwinkel übereinstimmen.

Beispiel 4.10:
Ich betrachte die Situation, in der die Länge einer Seite – sagen wir $a = 3$ – und die beiden dieser Seite anliegenden Innenwinkel – sagen wir $\beta = 120°$ und $\gamma = 20°$ – vorgegeben sind. Natürlich ist dann der verbleibende Winkel $\alpha = 180° - 120° - 20° = 40°$. Ich berechne jetzt die beiden verbleibenden Seitenlängen durch $b = a\ (\sin(\beta)/\sin(\alpha)) = 3\ (\sin(120°)/\sin(40°)) \approx 4{,}0419$ und $c = 3\ (\sin(20°)/\sin(40°)) \approx 1{,}5963$. ■

Oben hatte ich Ihnen gezeigt, dass der Cosinussatz direkt anwendbar ist, wenn die Längen zweier Seiten eines Dreiecks und der von diesen eingeschlossene Winkel vorgegeben ist. Sind nun die Längen zweier Seiten eines Dreiecks aber einer der beiden anderen Winkel vorgegeben, so stellt sich der Sinussatz *oftmals* als nützliches Werkzeug heraus.

Beispiel 4.11:
Für ein Dreieck wie in Abbildung 4.13 seien die Längen zweier Seiten – sagen wir $a = 3$ und $b = 1$ – und einer der Winkel, die *nicht* von diesen Seiten eingeschlossen sind, – sagen wir $\alpha = 50°$ – vorgegeben. Dann berechne ich mit dem Sinussatz zunächst den verbleibenden Winkel dieser Art: $\sin(\beta) = \sin(\alpha)(b/a) = \sin(50°)/3 \approx 0{,}2553$. Der Taschenrechner liefert mir nun nach Drücken von INV und SIN den Wert $\beta \approx 14{,}79°$.

Es ergibt sich nun $\gamma \approx 115{,}21°$. Für die verbleibende Länge habe ich die freie Auswahl, denn ich kann jetzt sowohl den Sinus- als auch den Cosinussatz anwenden. Ich berechne für diese $c \approx 3{,}5401$. ■

Ich habe Ihnen eben gezeigt, dass in diesem Beispiel durch Vorgabe der Länge zweier Seiten und eines Winkels die dritte Seitenlänge und die beiden verbleibenden Winkel durch Anwendung des Sinussatzes berechnet werden können. Ein bisschen muss ich hier im Allgemeinen aber aufpassen, denn sind die Seitenlängen a und b sowie α, der der Seite mit Länge a gegenüberliegende Winkel, vorgegeben, so muss *nicht* unbedingt

$$\left|\sin(\alpha)\ \tfrac{b}{a}\right| \leq 1$$

gelten. Da dies aber wegen $|\sin(\beta)| \leq 1$ immer erfüllt sein sollte, kann ich erkennen, dass für gewisse Vorgaben von a, b und α etwas schief geht, das heißt ein Dreieck für diese Größen nicht konstruierbar ist.

Beispiel 4.12:
Betrachte ich beispielsweise ein rechtwinkliges Dreieck $\alpha = 90°$, das heißt $\sin(\alpha) = 1$, so darf b nicht größer als a sein. Anderenfalls wäre nämlich die Länge der Hypotenuse a kleiner als die Länge b einer der Katheten, was nicht möglich ist. Also ist für $\alpha = 90°$ beispielsweise $b = 2$ und $a = 1$ keine gültige Wahl für Seitenlängen eines Dreiecks. ■

Gelingt die Berechnung andererseits, wie dies in Beispiel 4.11 der Fall ist, so ist das Dreieck durch die Vorgabe der Größen a, b und α bereits in seiner Gestalt vollkommen festgelegt. Zwei solche Dreiecke sind dann kongruent.

Übungsaufgabe 4.13:
Im Dreieck wie in Abbildung 4.13 seien $a = 2$ und $b = 8$ sowie $\alpha = 10°$. Bestimmen Sie die verbleibende Seitenlänge c und die beiden verbleibenden Innenwinkel β und γ. Betrachten Sie jetzt dieselben vorgegebenen Seitenlängen, aber $\alpha = 40°$. Was beobachten Sie? ■

Übungsaufgabe 4.14:
Leiten Sie die Formel $F = (abc)/(4r)$ für den Flächeninhalt eines Dreiecks mit Seitenlängen a, b und c und Umkreisradius r her. ■

Zum Abschluss dieses Abschnitts formuliere ich Ihnen aus Vollständigkeitsgründen den Tangenssatz, mit dem sich zuweilen einige Rechnungen verkürzen lassen. Ich verwende zu dessen Formulierung die Bezeichnungen aus Abbildung 4.13. Herleiten könnte ich mir diese Aussage aus dem Sinussatz unter Verwendung der Additionstheoreme und der Beziehung $\tan = \sin / \cos$.

Tangenssatz

Für beliebige Dreiecke mit Seitenlängen a und b und den diesen Seiten gegenüberliegenden Innenwinkeln α und β gilt: $\tan((\alpha-\beta)/2) = \tan((\alpha+\beta)/2)(a-b)/(a+b)$.

4.2 Ebene geometrische Figuren

In diesem Abschnitt möchte ich mich ein wenig mit geometrischen Figuren in der Ebene beschäftigen, die nicht ganz so einfach wie Dreiecke sind. Einige dieser Figuren erhalte ich, indem ich Dreiecke in geeigneter Weise zusammensetze – dies sind **Vierecke** und allgemeinere Figuren, die man **Vielecke** nennt. Vielecke sind durch Strecken berandet. Dies unterscheidet sie von ebenen Figuren, die einen krummen Rand haben. Solche ebenen Figuren – ich behandle hier den Kreis und die Ellipse – können nicht perfekt durch Zusammensetzung von Dreiecken konstruiert werden. Ich kann diese aber immerhin durch die Verwendung von genügend vielen Dreiecken beliebig genau annähern.

Die einfachsten ebenen Figuren, die sich vom Dreieck unterscheiden, sind Vierecke. Die Namensgebung lässt den wenig überraschenden Schluss zu, dass diese durch vier Eckpunkte A, B, C und D festgelegt sind. Anders als beim Dreieck muss ich bei der Beschreibung eines Vierecks aber ein bisschen aufpassen: Die Eckpunkte A, B, C und D sollten nämlich so angeordnet werden, dass ich die Strecken von A nach B, B nach C, C nach D und schließlich wieder von D nach A so abtragen kann, dass diese sich nicht überschneiden und damit tatsächlich auf die Umrandung einer zweidimensionalen Figur – dem Viereck – führen.

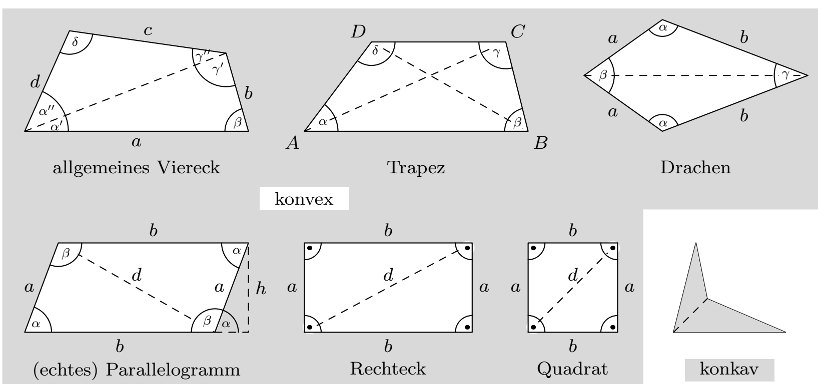

Abb. 4.15 Klassifizierung von Vierecken.

In Abbildung 4.15 zeige ich Ihnen verschiedene Typen von Vierecken $\diamondsuit$ zusammen mit ein paar Bezeichnungen, die zum Teil weiter unten eine Rolle spielen werden. Die dort vorgenommene Klassifizierung unterscheidet zwischen **allgemeinem Viereck**, **Trapez**, **Drachen** sowie **(echtem) Parallelogramm**, (auch: **Rhomboid**), **Rechteck** und **Quadrat**.

Einfache Vierecke sind sicherlich Rechtecke – hier stehen alle Seiten senkrecht aufeinander und ich habe genau zwei unterschiedliche Seitenlängen. Bei Quadraten sind diese Längen sogar gleich. Parallelogramme sind, salopp ausgedrückt, *schiefe* Rechtecke. Diese entstehen aus zwei Paaren paralleler Seiten und die Längen gegenüberliegender Seiten sind somit jeweils gleich. Sie bilden eine allgemeinere Klasse von

Vierecken als Rechtecke, da die Seiten nicht unbedingt senkrecht aufeinander stehen müssen. Sind alle Seiten eines Parallelogramms gleich lang, so nennt man es einen **Rhombus**. Bei Drachen(vierecken) haben jeweils zwei aneinander liegende Seiten gleiche Länge. Für Trapeze sind lediglich zwei Seiten parallel. Sind die Schenkel (im Bild von Abbildung 4.15 die Strecken von A nach D und B nach C) gleich lang, so heißt ein solches Trapez auch **gleichschenkliges Trapez**. Allgemeine Vierecke unterscheidet man noch in **konvexe Vierecke** und **konkave Vierecke**. Bei konvexen Vierecken ist jeder Innenwinkel kleiner gleich 180° und ich kann beide **Diagonalen** (das heißt Strecken, die gegenüberliegende Punkte im Viereck verbinden) einzeichnen, während in den konkaven Vierecken nur eine Diagonale im Innern des Vierecks verläuft.

Ein Viereck $\diamondsuit$ kann ich in zwei Dreiecke unterteilen, indem ich eine Diagonale in dessen Innern hinzuaddiere. Das habe ich beispielsweise am konkaven Vierecken in Abbildung 4.15 illustriert. Die Summe der Innenwinkel der beiden entstehenden Dreiecke ist 180° – dies ist ja bekanntlich für jedes Dreieck so der Fall. Mit den Bezeichnungen am allgemeinen Viereck in Abbildung 4.15 erhalte ich also

$$\alpha' + \beta + \gamma' = 180^\circ \text{ und } \alpha'' + \gamma'' + \delta = 180^\circ.$$

Sind $\alpha = \alpha' + \alpha''$ und $\gamma = \gamma' + \gamma''$ die neben β und δ auftretenden Winkel in $\diamondsuit$, so berechne ich damit die Summe der Winkel in $\diamondsuit$ als

$$\alpha+\beta+\gamma+\delta = (\alpha'+\alpha'')+\beta+(\gamma'+\gamma'')+\delta = (\alpha'+\beta+\gamma')+(\alpha''+\gamma''+\delta) = 2\cdot 180^\circ = 360^\circ.$$

Summe der Innenwinkel eines Vierecks

Die Summe $\alpha + \beta + \gamma + \delta$ der Innenwinkel $\alpha, \beta, \gamma, \delta$ eines Vierecks ist stets 360°.

Für manche Klassen von Vierecken gelten noch gewisse Eigenschaften und Beziehungen der Winkel. Oben habe ich Ihnen ja bereits gesagt, dass in *Recht*ecken alle Winkel **rechtwinklig**, also 90°, sind. Ich bemerke darüber hinaus, dass in Drachen stets zwei Winkel mit derselben Größe vorliegen – für den Drachen in Abbildung 4.15 habe ich diesen Winkel α genannt. Weiterhin treten in einem Parallelogramm lediglich zwei verschiedene Winkel α und β auf und diese summieren sich zu $\alpha + \beta = 180^\circ$.

Oftmals ist es sinnvoll, Kenntnisse über die **Länge der Diagonalen** d in einem Viereck $\diamondsuit$ zu besitzen. Für Rechtecke erhalte ich diese leicht durch Anwendung des **Satzes des Pythagoras**

$$d = \sqrt{a^2 + b^2},$$

denn das Hinzufügen einer Diagonale führt auf zwei **rechtwinklige Dreiecke** mit Kathetenlängen a und b. Ich beobachte, dass in diesem Fall die beiden Diagonalen gleich lang sind. Dasselbe gilt natürlich auch für Quadrate – hier haben die Diagonalen die Länge

$$d = \sqrt{a^2 + a^2} = \sqrt{2a^2} = \sqrt{2}a.$$

Etwas schwieriger wird die Bestimmung von d für Parallelogramme und die anderen allgemeineren Vierecke. Hier muss ich mich oftmals an die Verallgemeinerung des Satzes von Pythagoras aus dem letzten Abschnitt erinnern: Mit dem **Cosinussatz** rechne ich beispielsweise für das Parallelogramm in Abbildung 4.15

$$d = \sqrt{a^2 + b^2 - 2\,ab\,\cos(\alpha)}$$

aus. Die beiden Diagonalen in Parallelogrammen schneiden sich in ihren jeweiligen Mittelpunkten. Für Drachen $\diamondsuit$ ist diese Eigenschaft immerhin noch für die Diagonale der Fall, die $\diamondsuit$ in zwei gleichschenklige Dreieck zerlegt.

Länge der Diagonalen im Viereck

Für Quadrate mit Seitenlänge a ist $d = \sqrt{2}a$ die Länge der Diagonalen. Für Rechtecke mit Seitenlängen a und b ist $d = \sqrt{a^2 + b^2}$ die Länge der Diagonalen. Für Parallelogramme mit Seitenlängen a und b und von diesen Seiten eingeschlossenem Winkel α ist $d = \sqrt{a^2 + b^2 - 2\,ab\,\cos(\alpha)}$ die Länge der Diagonalen. Für allgemeine Vierecke $\diamondsuit$ sind die Längen der Diagonalen d ebenfalls mit dem Cosinussatz bestimmbar, falls neben den Seitenlängen auch die entsprechenden eingeschlossenen Winkel in $\diamondsuit$ bekannt sind.

Haben Sie schon einmal einen Teppichboden verlegt? Falls ja, dann haben Sie sich sicherlich schon einmal Gedanken über die Fläche von Vierecken und allgemeineren Figuren gemacht. Sollten Sie dies eventuell versäumt haben, so kennen Sie möglicherweise die fatalen Folgen der Unkenntnis dessen, was ich Ihnen nun beginne zu beschreiben. Es geht mir nämlich um den Flächeninhalt. Ich beginne mit Vierecken und später betrachte ich diesen noch für allgemeinere, geometrische Figuren.

Den Flächeninhalt F von Quadraten und Rechtecken kann man sehr einfach berechnen: Dieser ist $F = a^2$ beziehungsweise $F = ab$. Wie man den Flächeninhalt F eines Parallelogramms prinzipiell ausrechnet, habe ich Ihnen bereits im letzten Abschnitt gezeigt: Man benötigt hierzu die Länge einer Höhe h. Falls der Winkel α (siehe Abbildung 4.15) im Parallelogramm bekannt ist, so kann ich hierzu die Formel $h = \sin(\alpha)a$ benutzen. Mithilfe der Länge der verbleibenden Seite erhalte ich dann $F = \sin(\alpha)\,ab$. Unten werde ich Ihnen noch zeigen, wie man auf der Basis dieser Formel den Flächeninhalt eines Trapezes berechnen kann.

Flächeninhalt von Parallelogrammen

Der Flächeninhalt F eines Parallelogramms mit Seitenlängen a und b und von diesen Seiten eingeschlossenem Winkel α ist $F = \sin(\alpha)\, ab$. Spezieller ist der Flächeninhalt F eines Rechtecks ($\alpha = 90°$) $F = ab$ und der Flächeninhalt F eines Quadrats ($\alpha = 90°$ und $a = b$): $F = a^2$.

Übungsaufgabe 4.15:
In einem Parallelogramm $\diamondsuit$ wie in Abbildung 4.15 seien die Seitenlänge $a = 5$ und $b = 3$ sowie der Winkel $\beta = 120°$ vorgegeben. Berechnen Sie die Länge der Diagonale und den Flächeninhalt von $\diamondsuit$. ■

Ist das vorgegebene Viereck $\diamondsuit$ aber nun von allgemeinerer oder anderer (zum Beispiel Drachen-)Bauart, so muss ich mich zuweilen etwas mehr bemühen, um den Flächeninhalts von $\diamondsuit$ zu berechnen. Der *Haupttrick* hierzu ist es, dann wiederum das Viereck in zwei (oder eventuell sogar vier) Dreiecke zerlegen, indem ich eine (respektive beide) Diagonale(n) von $\diamondsuit$ hinzufüge. Abhängig von den gegebenen Größen kann ich mir dann die bequemste Formel zur Berechnung des Flächeninhalts der Teildreiecke aus dem vorherigen Abschnitt wählen.

Beispiel 4.13:
In einem Drachen $\diamondsuit$ seien die Längen a und b vorgegeben und der Winkel α wie in Abbildung 4.15 bekannt. Dann kann ich mit dem **Cosinussatz** aus dem letzten Abschnitt zunächst die Länge d der Diagonale in $\diamondsuit$ berechnen, die die beiden Eckpunkte mit den verbleibenden Innenwinkel verbindet $d = \sqrt{a^2 + b^2 - 2\,ab\,\cos(\alpha)}$. Die beiden durch Hinzufügen dieser Diagonale aus $\diamondsuit$ entstehenden Dreiecke Δ_1 und Δ_2 sind **kongruent** – haben also insbesondere denselben Flächeninhalt. Da ich nun die drei Seitenlängen von Δ_1 kenne, kann ich **Herons Formel** anwenden, um den Flächeninhalt von Δ_1 zu bestimmen. Das Doppelte davon ist dann der Wert des Flächeninhalts von $\diamondsuit$. ■

Beispiel 4.14:
In einem Drachen $\diamondsuit$ seien die Längen a und b und der Winkel β wie in Abbildung 4.15 vorgegeben. Ich füge dann die Diagonale zu $\diamondsuit$ hinzu, die die beiden Eckpunkte mit dem gleichen Innenwinkel α verbindet, und bezeichne deren Länge mit d. Es entstehen dann zwei **gleichschenklige Dreiecke**, deren Grundseite diese Diagonale ist. Um deren Länge d zu bestimmen, kann ich wieder den **Cosinussatz** anwenden: $d = \sqrt{2a^2 - 2\,a^2\,\cos(\beta)} = a\sqrt{2\,(1 - \cos(\beta))}$. Danach könnte ich wiederum Herons Formel anwenden. Eine alternative Möglichkeit ergibt sich aber aus Übungsaufgabe 4.4, in der ich Sie gebeten hatte, eine allgemeine Formel für den Flächeninhalt gleichschenkliger

Dreiecke herzuleiten. Die Summe der Flächeninhalte der beiden beteiligten Dreiecke ergibt dann den Flächeninhalt

$$F = \frac{d}{4}\left(\sqrt{4a^2 - d^2} + \sqrt{4b^2 - d^2}\right)$$

von $\diamondsuit$. ■

Wie so oft im Leben hängt die Einfachheit oder Schwere eines Vorhabens von den gegebenen Grundvoraussetzungen ab. Nicht anders ist es mit der Berechnung des Flächeninhalts eines Drachens – schauen Sie sich hierzu bitte die nächste Aufgabe an:

Übungsaufgabe 4.16:
Begründen Sie, weshalb der Flächeninhalt F eines Drachens $\diamondsuit$ die Hälfte des Produkts der Längen der Diagonalen in $\diamondsuit$ ist. ■

Sind für ein allgemeines Viereck die Seitenlängen a, b, c und d sowie die gegenüberliegenden Winkel β und δ, wie in Abbildung 4.15, bekannt, so kann ich denselben Trick wie in den obigen beiden Beispielen anwenden. Hinzufügen der Diagonale wie in Abbildung 4.15 unterteilt das Viereck in zwei Dreiecke und durch Anwendung des **Cosinussatzes** gelangt man nach etwas mühevoller Rechnung, die ich uns ersparen möchte, auf die folgende Formel für den Flächeninhalt F des allgemeinen Vierecks, die ihrer Bauart nach mit **Herons Formel** verwandt ist:

$$F = \sqrt{(s-a)(s-b)(s-c)(s-d) - abcd\cos(\sigma)},$$

wobei $s = (a+b+c+d)/2$ und $\sigma = (\beta+\delta)/2$. Ist das Viereck konkav, so sollten die hier vorgegebenen Winkel β und δ nicht von der Diagonale geteilt werden.

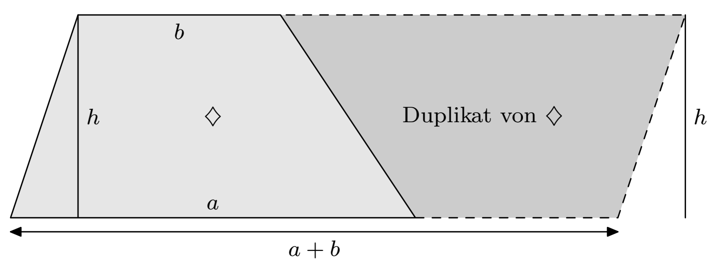

Abb. 4.16 Der Flächeninhalt eines Trapezes $\diamondsuit$ ist die Hälfte des Flächeninhalts eines geeigneten Parallelogramms, welches durch Duplizierung von $\diamondsuit$ entsteht.

Sie sehen also, dass es zumeist von Erfolg gekrönt ist, Größen von etwas komplizierten Figuren mithilfe von bekannten Beziehungen für einfachere Figuren auszurechnen. Als abschließendes Beispiel im Rahmen der Flächeninhaltsberechnungen von Vierecken betrachte ich noch das **Trapez**. Hier kann ich wie in Abbildung 4.16 vorgehen. Dort nehme ich an, dass die Längen a und b der beiden parallelen Seiten sowie die Länge der Höhe h eines Trapezes $\diamondsuit$ vorgegeben sind. Eine Idee ist es nun, ein umgedrehtes Duplikat des Trapezes an $\diamondsuit$ anzuhängen, sodass insgesamt eine einfachere Figur – ein Parallelogramm – entsteht. Dieses Parallelogramm hat eine Grundseite der Länge $a + b$ und die entsprechende Höhe der Länge h. Außerdem hat dieses Parallelogramm den doppelten Flächeninhalt wie $\diamondsuit$, sodass ich zu der folgenden Aussage gelange:

Flächeninhalt des Trapez

Der Flächeninhalt F eines Trapezes mit Seitenlängen a und b der parallelen Seiten und Höhenlänge h ist $F = (a + b)h/2$.

Übungsaufgabe 4.17:
Für ein Trapez $\diamondsuit$ seien die Längen $a = 6$ und $b = 4$ der beiden parallelen Seiten vorgegeben. Weiter schließe die Seite mit Länge a einen Winkel von $\alpha = 45°$ mit einer Seite von $\diamondsuit$ der Länge $c = \sqrt{2}$ ein. Bestimmen Sie den Flächeninhalt von $\diamondsuit$. ■

Übungsaufgabe 4.18:
Bestimmen Sie *alle* Seitenlängen und die Winkel eines gleichschenkligen Trapezes mit Flächeninhalt 1, bei dem eine der beiden parallelen Seiten dieselbe Länge wie die Höhe auf diese Seite besitzt und die verbleibende, parallele Seite doppelt so lang ist. ■

Möchte ich nun allgemeinere ebene Figuren als Dreiecke und Vierecke betrachten, die ich aber dennoch in einfacher Weise in die Ebene zeichnen kann, so gelange ich auf **Vielecke**, die man auch ***N*-Ecke** nennt, denn deren Anzahl von Ecken ist N – eine natürliche Zahl, die größer als 2 ist. Zur Konstruktion eines N-Ecks gebe ich mir N Punkte in der Ebene vor und verbinde diese mit N Strecken so, dass diese sich nicht überschneiden und der Anfangspunkt der ersten Strecke mit dem Endpunkt der letzten Strecke übereinstimmt (siehe Abbildung 4.17). Diese beiden Eigenschaften garantieren, dass ein solches **geschlossenes Polygon** (alle Strecken) auch tatsächlich die Umrandung einer zweidimensionalen Figur – dem N-Eck – ist.

Abb. 4.17 Ein konvexes Fünfeck (5-Eck) und ein konkaves Sechseck (6-Eck).

Ich erkenne nun, indem ich kurz überlege, wie N-Ecke für $N = 3$ und $N = 4$ aussehen, dass Dreiecke und Vierecke lediglich Spezialfälle von Vielecken sind. Ich habe Ihnen erklärt, dass die Winkel im Dreieck stets kleiner als 180° sind und dass dies für Vierecke nicht gelten muss, denn diese können konkav sein. Allgemeiner nennt man nun für $N \geq 4$ ein N-Eck Λ **konvex**, falls alle Winkel in Λ kleiner gleich 180° sind und anderenfalls heißt es **konkaves** N-Eck. Anders ausgedrückt bedeutet dies, dass in konvexen N-Ecken Λ mit je zwei Punkten X, Y auch jeder Punkt Z der Verbindungsstrecke zwischen diesen Punkten, also alle Punkte Z der Form

$$Z = t\ X + (1 - t)\ Y \text{ für ein } t \in [0, 1],$$

in Λ liegt (siehe Abbildung 4.17). In konkaven Vielecken Λ existieren andererseits Punkte X und Y in Λ, sodass ein Punkt Z auf deren Verbindungsstrecke nicht in Λ liegt. Die Addition von Punkten in der Ebene und die Multiplikation mit einem reellen Skalar ist hier so zu verstehen, wie ich Ihnen dies im letzten Abschnitt vor der Berechnung des Schnittpunkts der Seitenhalbierenden im Dreieck erklärt habe.

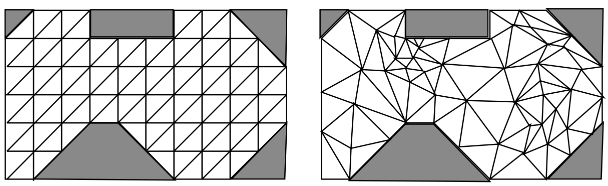

Abb. 4.18 Gleichmäßige und ungleichmäßige Triangulierungen des Bodens Ihres Wohnzimmers. Graue Bereiche: Schränke, Vitrine, Fernsehschrank etc.

Manchmal ist es notwendig, dass man ähnliche Berechnungen, wie ich Ihnen oben für das Viereck gezeigt habe, auch für vorgegebene N-Ecke durchführt. Das wissen Sie eventuell noch vom Verlegen Ihres Teppichbodens: In den seltensten Fällen ist es nämlich so, dass der Teil des Bodens, den Sie mit Teppich belegen wollen, auch tatsächlich mit einer rechteckigen Grundfläche übereinstimmt. Zumeist steht nämlich irgendein Schrank im Weg, der Raum geht um eine Ecke herum, die Wände in Ihrer Altbauwohnung sind schief oder Ähnliches. Mit anderen Worten: Die zu belegende Fläche bildet oftmals ein allgemeineres N-Eck, wie die umrandete Fläche in Abbildung 4.18. Wenn Sie nun im Teppichgeschäft stehen, stellt sich Ihnen zumeist die unangenehme Frage, wie groß ein dort angebotenes, rechteckiges Teppichstück sein sollte. Zum einen streben Sie ja an, den gesamten Boden zu belegen – zum anderen möchten Sie aber auch nicht gar zu viel Teppichreste nach der Verlegung haben, denn so dicke haben Sie es ja auch nicht – sonst würden Sie wohl kaum den Teppich selbst verlegen.

Glücklicherweise ist es nun so, dass ich N-Ecke immer in Dreiecke zerlegen kann – und über Dreiecke weiß ich nach dem Schmökern im letzten Abschnitt relativ gut Bescheid. Dies ist also nach wie vor mein *Haupttrick*, der nun also auch für N-Ecke anwendbar ist. Prinzipiell funktioniert das also genauso, wie ich Ihnen das oben für

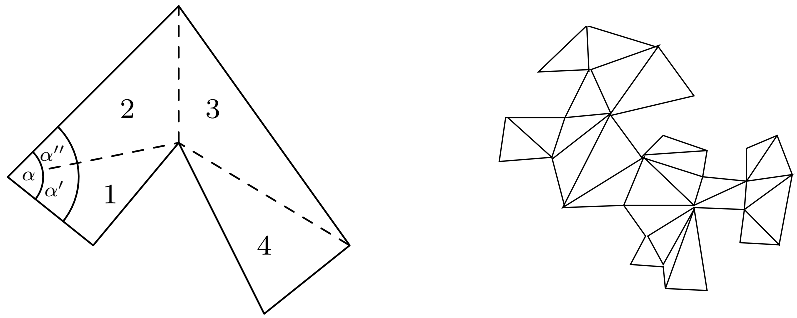

Abb. 4.19 N-Ecke lassen sich immer in Dreiecke zerlegen, sodass eine Triangulierung entsteht.

Vierecke gezeigt habe, das heißt, ich füge gewisse Strecken zu einem gegebenen N-Eck hinzu. Abbildung 4.19 (links) illustriert die Vorgehensweise. Für ein gegebenes N-Eck $\varLambda$ kann ich Schritt für Schritt die Dreiecke mit der dortigen Nummerierung abschälen beziehungsweise die entsprechenden gestrichelten Strecken einzeichnen, bis schließlich das ganze N-Eck in Dreiecke zerlegt ist. Die resultierenden Dreiecke bilden eine **Triangulierung** $\mathcal{T}$ von $\varLambda$ und man kennt sogar die Anzahl der Dreiecke: $N-2$. Diese Vorgehensweise funktioniert für jedes N-Eck – ein etwas größeres Beispiel zeige ich Ihnen in Abbildung 4.19 (rechts). Manchmal ist es hierbei durchaus sinnvoll, noch andere Triangulierungen $\mathcal{T}$ zu betrachten – zum Beispiel können Sie den Boden Ihres Wohnzimmers (siehe Abbildung 4.18 (links)) mit gleichmäßigen Dreiecken überdecken, indem Sie zusätzlich ein paar innere Punkte einfügen – das hat den Vorteil, dass Sie beispielsweise nur die Anzahl der Dreiecke und deren Flächeninhalt kennen müssen, um zu erkennen, wie viel Teppichboden Sie benötigen. Ist also beispielsweise 0,5 m^2 die Fläche jedes der 93 Dreiecke in Abbildung 4.18, so hat die mit Teppich zu belegende Fläche den Inhalt 46,5 m^2. Wird in Ihrem Teppichfachgeschäft der Teppich auf Rollen der Breite 2 Meter verkauft, so sollten Sie 24 Meter Teppich kaufen. Ihnen bleibt in diesem Fall 1,5 m^2 Rest – was kein Fehler ist, denn das nächste Rotweinglas wartet schon darauf, umzukippen. Sie können natürlich auch allgemeinere Triangulierungen wie in Abbildung 4.18 (rechts) betrachten. Mit solchen Objekten beschäftigen sich gewisse Mathematiker – aus praktischem Blickwinkel des Teppichbodenverlegens ist dies jedoch nach meinem Kenntnisstand weniger empfehlenswert.

Ein typisches Beispiel, bei dem ich die Zerlegung eines N-Ecks in Dreiecke anwenden kann, ist die Bestimmung der Summe aller N Innenwinkel α eines beliebigen N-Ecks $\varLambda$. Indem ich diese Winkel in Winkel wie beispielsweise α' und α'' in Abbildung 4.19 (links) gemäß aller in einem Eckpunkt endenden gestrichelten Stecken zerlege und verwende, dass die Summe der Winkel von jedem der $N-2$ Dreiecke in $\varLambda$ stets $180°$ ist, gelange ich auf das folgende Resultat:

Summe der Innenwinkel eines N-Ecks

Die Summe der N Innenwinkel eines N-Ecks ist stets $(N-2)\ 180°$.

Von besonderem Interesse sind regelmäßige N-Ecke. Dies sind N-Ecke wie in Abbildung 4.20, bei denen die N Seiten alle die gleiche Länge b haben und die N Innenwinkel α stets übereinstimmen, also

$$\alpha = \frac{(N-2)\ 180°}{N} = 180° - \frac{1}{N}\ 360°$$

gilt. Um den Flächeninhalt F_N eines solchen N-Ecks Λ zu berechnen, kann ich mir das Leben einfach machen, indem ich einen weiteren Punkt M im Zentrum von Λ wähle und die gleichmäßige Triangulierung $\mathcal{T}$ von Λ betrachte, die entsteht, wenn ich M mit allen Eckpunkten von Λ verbinde. Die N Eckpunkte des N-Ecks liegen dann auf einem Kreis mit Mittelpunkt M und Radius a (siehe Abbildung 4.20 (links)). Die Triangulierung $\mathcal{T}$ besteht offenbar aus N **gleichschenkligen** Dreiecken, die **kongruent** sind. Für jedes dieser Dreiecke Δ kann ich sofort erkennen, dass der der Seite mit Länge b gegenüberliegende Winkel in Δ den Wert $\beta = 360°/N$ hat. Die Verwendung des Sinus ergibt somit

$$\sin\left(\tfrac{180°}{N}\right) = \frac{\frac{b}{2}}{a}, \text{ beziehungsweise } b = 2a\sin\left(\tfrac{180°}{N}\right),$$

wobei a als Länge der Schenkel von Δ auftritt. Wenn ich mich nun an die Flächeninhaltsformel für gleichschenklige Dreiecke aus Übungsaufgabe 4.4 erinnere, so kann

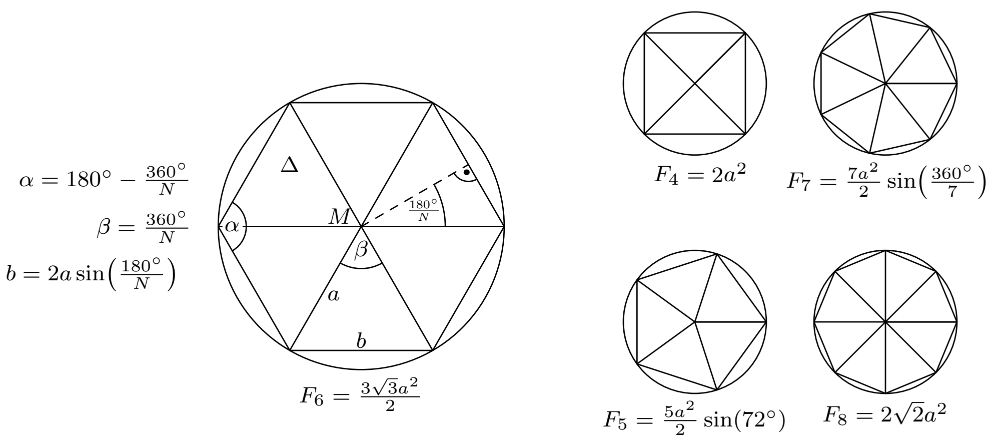

Abb. 4.20 Der Flächeninhalt F_N regelmäßiger N-Ecke bestimmt sich gemäß der Formel $F_N = (Na^2/2)\ \sin(360°/N)$. Links sieht man ein regelmäßiges Sechseck zusammen mit den in der Herleitung der allgemeinen Formel verwendeten Symbolen und Zusammenhängen. Hier gilt ausnahmsweise $a = b$. Rechts sieht man ein regelmäßiges Vier-, Fünf, Sieben- und Achteck zusammen mit den Formeln für deren Flächeninhalt für diese Wahlen von N, das heißt $N = 4, 5, 7, 8$.

ich den Flächeninhalt F_Δ der Dreiecke Δ ermitteln, indem ich b ersetze

$$F_\Delta = \frac{b}{4}\sqrt{4a^2 - b^2} = \frac{2a\ \sin(\frac{180°}{N})}{4}\sqrt{4a^2 - (2a\ \sin(\frac{180°}{N}))^2}\ .$$

Durch Kürzen, Ausquadrieren und Ziehen der Wurzel aus $4a^2$ erhalte ich

$$F_\Delta = \frac{a\ \sin(\frac{180°}{N})}{2}\sqrt{4a^2 - 4a^2\ \sin^2(\frac{180°}{N})} = \frac{a\ \sin(\frac{180°}{N})}{2}\ 2a\ \sqrt{1 - \sin^2(\frac{180°}{N})}\ .$$

Schließlich komme ich auf

$$\begin{aligned} F_\Delta &= a^2\ \sin(\tfrac{180°}{N})\sqrt{1 - \sin^2(\tfrac{180°}{N})} = a^2\ \sin(\tfrac{180°}{N})\ \cos(\tfrac{180°}{N}) \\ &= \tfrac{a^2}{2}\ \sin(\tfrac{180°}{N} + \tfrac{180°}{N}) = \tfrac{a^2}{2}\ \sin(\tfrac{360°}{N}), \end{aligned}$$

wobei ich bei der drittletzten und vorletzten Gleichheit die aus dem letzten Abschnitt bekannten Beziehungen $\sin(x) = \sqrt{1 - \cos^2(x)}$ beziehungsweise $\sin(2x)/2 = \sin(x)\cos(x)$ ausgenutzt habe. Das regelmäßige N-Eck besteht nun aus N solchen Dreiecken Δ mit dem Flächeninhalt F_Δ.

Flächeninhalt des regelmäßigen N-Ecks

Ein regelmäßiges N-Eck hat den Flächeninhalt $F_N = N\ (a^2/2\ \sin(360°/N))$, wobei a der Radius des kleinsten Kreises ist, der dieses enthält.

Übungsaufgabe 4.19:
a) Bestimmen Sie den Flächeninhalt des regelmäßigen 3600-Ecks mit Seitenlängen b. b) Zeigen Sie, wie man unter Verwendung von gleichmäßigen N-Ecken auf die Näherungsformel $U_N = 2N\sin(180°/N)$ für den Umfang eines Kreises mit Radius $a = 1$ gelangt und berechnen Sie U_{3600}. ■

In Abbildung 4.20 (rechts) zeige ich Ihnen jeweils ein regelmäßiges Vier-, Fünf-, Sieben- und Achteck zusammen mit den für diese Wahlen speziellen Formeln für deren Flächeninhalt. Sie können an diesen Spezialfällen – aber auch an der allgemeinen Formel – erkennen, dass der Flächeninhalt dieser N-Ecke neben dem Radius a des Kreises mit Mittelpunkt M, den ich diesen Figuren **umschrieben** habe, lediglich noch von dem durch N festgelegten Winkel $360°/N$ abhängt.

Apropos Kreise. Sie wissen natürlich längst, dass ich einen allgemeinen **Kreis** in der Ebene bilde, indem ich wie in Abbildung 4.21 die Menge aller Punkte $(x|y)$ betrachte, die zu einem festen **Mittelpunkt** $M = (x_1|y_1)$ denselben festen Abstand r – **Radius** genannt – besitzen. Die doppelte Länge des Radius nennt man **Durchmesser** des Kreises. Der Abstand zweier Punkte $(x|y)$ und $(x_1|y_1)$ in der Ebene kann ich nach dem, was ich Ihnen im letzten Abschnitt erzählt habe, unter Verwendung des **Satzes des Pythagoras** durch die Formel $r = \sqrt{(x - x_1)^2 + (y - y_1)^2}$ berechnen. Somit wird die Menge aller Punkte $(x|y)$ auf dem Kreis mit Mittelpunkt $M = (x_1|y_1)$ und

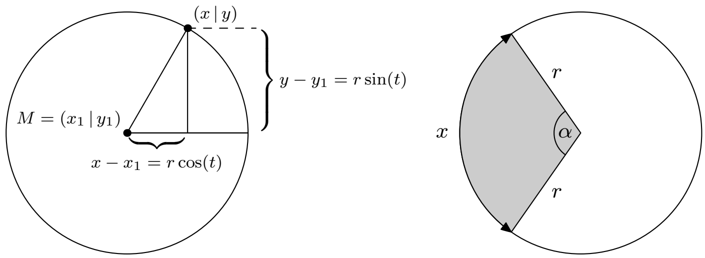

Abb. 4.21 Die Punkte $(x|y)$ auf einem Kreis mit Mittelpunkt $M = (x_1|y_1)$ und Radius r können mithilfe des Sinus und Cosinus parametrisch ausgedrückt werden. Die Mittelpunktwinkel α sind proportional zur Länge x des entsprechenden Kreisbogens.

Radius r durch die **Kreisgleichung**

$$(x - x_1)^2 + (y - y_1)^2 = r^2$$

beschrieben. Manche Leute schreiben das auch folgendermaßen hin:

$$\left(\frac{x - x_1}{r}\right)^2 + \left(\frac{y - y_1}{r}\right)^2 = 1.$$

Wenn ich mich an die Einführung des Sinus und Cosinus aus dem letzten Abschnitt erinnere und einen kurzen Blick auf Abbildung 4.21 werfe, so kann ich die Gültigkeit der alternativen **Parameterdarstellung des Kreises**

$$x = x_1 + r\cos(t) \text{ und } y = y_1 + r\sin(t),\ \ t \in [0, 2\pi)$$

sicherlich nachvollziehen. Zur Orientierung kann ich mir hierzu noch überlegen, dass die N Eckpunkte $(x_k|y_k),\ k = 0, \ldots, N-1$, eines gleichmäßigen N-Ecks wie in Abbildung 4.20 beispielsweise durch

$$(x_k|y_k) = \left(x_1 + r\cos\left(\frac{k\ 360°}{N}\right)\middle|\, y_1 + r\sin\left(\frac{k\ 360°}{N}\right)\right),\ \ k = 0, \ldots, N-1,$$

(wobei $r = a$) gegeben sind. Schon im letzten Abschnitt habe ich Ihnen erklärt, dass drei Punkte, die die Eckpunkte eines Dreiecks Δ bilden, einen Kreis – den **Umkreis von** Δ – schon vollständig festlegen. Jetzt können wir solche Kreise explizit ausrechnen. Ich betrachte hier ein – wegen der simplen Wahl von Punkten – leicht vereinfachendes Beispiel.

Beispiel 4.15:
Ich bestimme den Kreis durch die Punkte $(0|0), (1|0)$ und $(1|1)$, indem ich diese in die allgemeine Kreisgleichung einsetze:

$$x_1^2 + y_1^2 = r^2, (1 - x_1)^2 + y_1^2 = r^2 \text{ und } (1 - x_1)^2 + (1 - y_1)^2 = r^2.$$

Ziehe ich die erste Gleichung von der zweiten Gleichung ab, so erhalte ich $(1 - x_1)^2 - x_1^2 = 0$, das heißt nach Anwenden der **Binomischen Formel** $1 - 2x_1 = 0$, also $x_1 = 1/2$. Abziehen der dritten Gleichung von der zweiten Gleichung führt auf $y_1^2 - (1 - y_1)^2 = 0$ und ich berechne genauso wie eben $y_1 = 1/2$. Der Mittelpunkt ist somit $M = (1/2|1/2)$. Darüber hinaus erhalte ich den Radius r, indem ich $x_1 = y_1 = 1/2$

in eine dieser drei Gleichungen – sagen wir die erste – einsetze: $1/4 + 1/4 = r^2$ – also $r = 1/\sqrt{2} = \sqrt{2}/2$. Die gefundene Kreisgleichung ist $(x - 1/2)^2 + (y - 1/2)^2 = 1/2$. Alternativ könnte ich natürlich den Mittelpunkt als den Schnittpunkt zweier Mittelsenkrechten berechnen und dann den Radius wie eben ermitteln. ■

Festlegung von Kreisen

Drei Punkte, die nicht auf einer gemeinsamen Geraden liegen (also ein Dreieck bilden), legen einen ebenen Kreis eindeutig fest.

Tab. 4.2 Mit einer Computeralgebra (in wenigen Sekunden) berechnete Näherungen des Flächeninhalts π des Einheitskreises durch Verwendung von regelmäßigen N-Ecken für verschiedene Werte von N.

N	F_N
16	3,061467458920
64	3,136548490545
256	3,141277250932
1 024	3,141572940367
4 096	3,141591421511
16 384	3,141592576584
65 536	3,141592648776
262 144	3,141592653288
1 048 576	3,141592653570
4 194 304	3,141592653588

Kreise unterscheiden sich von den zuvor behandelten ebenen Figuren vor allem dadurch, dass sie nicht durch Strecken berandet sind – man spricht auch von einem **krummen Rand**. Jetzt erscheint es auf den ersten flüchtigen Blick hin schwierig, beispielsweise den Flächeninhalt eines Kreises durch eine bequeme Formel auszurechnen. Schließlich kann es wohl nicht gelingen, einen Kreis mit endlich vielen Dreiecken oder anderen, nur durch Strecken begrenzten, ebenen Figuren *exakt* darzustellen. Hier hilft mir eine Beobachtung, die ich bei genauer Betrachtung von Abbildung 4.20 machen kann. Mit ansteigender Eckenanzahl N scheint gemäß dieser der Flächeninhalt des regelmäßigen N-Ecks eine immer bessere Annäherung an den Flächeninhalt des **Einheitskreises** (das heißt den Kreis mit Radius $r = a = 1$) mit Mittelpunkt – sagen wir $M = (0|0)$ – zu sein. Der Flächeninhalt des regelmäßigen N-Ecks wie in Abbildung 4.20 ist zwar für alle natürlichen Zahlen N stets kleiner als der Flächeninhalt des Kreises – in der Tat ist es aber so, dass sich dieser selbigen beliebig genau – man sagt *von unten* – annähert. Ich möchte jetzt an dieser Stelle keine den Flächeninhalt des Kreises *von oben* annähernde, streckenbegrenzte ebene Figur festlegen, um dann eventuelle **Grenzwert**betrachtungen zu vollziehen, sondern biete Ihnen stattdessen

Tabelle 4.2 an. Dort habe ich an einem langen Winterabend in mühevoller häuslicher Arbeit den Flächeninhalt von N-Ecken für große Werte N ausgerechnet:

Ich habe dann ein riesiges N gewählt und den Flächeninhalt des Einheitskreises – auch **die Zahl** π genannt – auf 40 Nachkommastellen genau ausgerechnet

$$\pi \approx 3{,}1415926535897932384626433832795028841971 6.$$

Sie sehen, dass die Nachkommastellen von π vollkommen unregelmäßig sind und keine Schemata in deren Struktur erkennbar sind. Dies liegt in der Natur der Sache, denn π ist – wie man heute weiß – keine **rationale Zahl**. Mehr noch: π unterscheidet sich auch entscheidend von Wurzeln – wie zum Beispiel $\sqrt{2}$ – aber diese Details über **transzendente Zahlen** und die **Unmöglichkeit der Quadratur des Kreises** aus der Algebra sparen wir uns. Wichtiger ist für Sie zu Beginn Ihres Studiums, dass ich Folgendes festhalte:

Flächeninhalt des Kreises

Der Einheitskreis hat den Flächeninhalt $F = \pi$. Ein beliebiger Kreis mit Radius r hat den Flächeninhalt $F = \pi r^2$.

Wie bin ich jetzt so schnell auf den Flächeninhalt eines beliebigen Kreis mit Radius r gekommen? Nun – hier liegt kein Geheimnis vor. Der Flächeninhalt des regelmäßigen N-Ecks im Einheitskreis ist $F_N = (N/2)\ \sin(360°/N)$, während der Flächeninhalt des regelmäßigen N-Ecks in einem beliebigen Kreis mit Radius $r = a$ durch $F_N \cdot r^2$ gegeben ist. Der unterschiedliche multiplikative Faktor r^2 ist aber gerade der Faktor, der auch in der letzten Aussage auftrat.

Zur Berechnung der Länge der Strecke, die ich ablaufen muss um einmal auf einem Kreis herumzulaufen, den so genannten **Umfang eines Kreises**, kann ich nun, ähnlich wie bei der Berechnung des Flächeninhalts, wie folgt vorgehen. Hier hilft Ihnen Ihre Erfahrung aus Übungsaufgabe 4.19b) weiter.

Umfang des Kreises

Der Einheitskreis hat den Umfang 2π. Ein beliebiger Kreis mit Radius r hat den Umfang $U = 2\pi r$.

Frage ich mich an dieser Stelle, wie lang der Bogen x eines Kreises mit Radius r zu einem vorgegebenen Mittelpunktwinkel α wie in Abbildung 4.21 ist, so kann ich schnell erkennen, dass für diesen $x = 2\pi\alpha/360° = \pi\alpha/180°$ gilt. Eine geringfügige Verallgemeinerung des Werts x habe ich im vorherigen Abschnitt (dort ließ ich auch negative Werte zu) – wie Sie sich sicherlich noch erinnern, denn Sie sind vermutlich jünger wie ich – **Bogenmaß** genannt.

Möchte ich nun andererseits den Flächeninhalt F_α eines **Kreisausschnitts** (auch: **Kreissektor**) mit Mittelpunktwinkel α wie in Abbildung 4.21 ausrechnen, so stellt dies für den routinierten Pizzaesser auch kein Problem mehr da – dieser ist

$$F_\alpha = \frac{\alpha}{360^\circ}\,\pi r^2 = xr^2/2.$$

Übungsaufgabe 4.20:
Freitag ist wie immer großer Pizzatag und Sie bestellen für sich und Ihre sieben Freunde eine Family-Pizza mit 80 cm Durchmesser. Wie viel cm^2 Pizza bekommt – gleichmäßige Verteilung vorausgesetzt – in etwa jeder von Ihnen? Wie viel cm unbelegter Rand steht jedem von Ihnen in etwa zur Verfügung? Diese Woche macht Ihnen Ihr Pizzabäcker ein ganz besonderes Angebot: Er verkauft Ihnen zum selben Preis eine Pizza in Form eines Kreisrings mit äußerem Durchmesser 1 Meter und innerem Durchmesser von 50 cm. Sie rechnen kurz aus, wie viel cm^2 Pizza und wie viel cm unbelegter Rand nun jeder Ihrer Freunde bekommen würde. Nehmen Sie das Angebot an? ■

Zum Abschluss möchte ich Ihnen noch ein paar Dinge über **Ellipsen** in Erinnerung bringen: Diese sind allgemeinere ebene Figuren als Kreise, die ebenfalls krummlinig berandet sind. Sie werden gebildet aus der Menge aller Punkte $(x|y)$, für die die Summe $r_1 + r_2$ der Abstände zu zwei festen Punkten $A = (a_1|a_2)$ und $B = (b_1|b_2)$ – **Brennpunkte** genannt – einen konstanten Wert $2r$ hat. Wenn ich mich also daran erinnere, wie ich den Abstand zweier Punkte in der Ebene mit dem **Satz des Pythagoras** ausrechne, so erhalte ich die **allgemeine Ellipsengleichung**

$$\overbrace{\sqrt{(x-a_1)^2+(y-a_2)^2}}^{r_1} + \overbrace{\sqrt{(x-b_1)^2+(y-b_2)^2}}^{r_2} = 2r.$$

Ich nehme im Folgenden der Einfachheit der Beschreibung halber stets an, dass die beiden Brennpunkte wie in Abbildung 4.22 auf der x-Achse liegen und der **Mittelpunkt** der Ellipse (Mitte der Brennpunkte) der **Ursprung** $(0|0)$ ist: $A = (-c|0)$ und $B = (c|0)$, wobei $c > 0$. Die Ellipsengleichung ist dann von der folgenden Form:

$$\sqrt{(x+c)^2+y^2} + \sqrt{(x-c)^2+y^2} = 2r.$$

In diesem Fall sind die beiden **Hauptscheitel** der Ellipse die Punkte $H_1 = (-r|0)$ und $H_2 = (r|0)$, die diese Punkte verbindende Strecke der Länge $2r$ heißt **Hauptachse** der Ellipse und r wird **Hauptachsenradius** genannt. Die beiden **Nebenscheitel** der Ellipse sind dann die Punkte der Form $N_1 = (0|b)$ und $N_2 = (0|-b)$ auf der Ellipse – also die Schnittpunkte einer solchen Ellipse mit der y-Achse. Diese Punkte sind die Schnittpunkte der Kreise mit Radius r um die Brennpunkte und ich erhalte aus dem **Satz des Pythagoras** $b^2 = r^2 - c^2$. Somit ist $N_1 = (0|\sqrt{r^2-c^2})$ und $N_2 = (0|-\sqrt{r^2-c^2})$. Die Nebenscheitel bilden die **Nebenachse** der Ellipse und $b = \sqrt{r^2-c^2}$ bezeichnet den **Nebenachsenradius**.

Aus der Dreiecksungleichung am Anfang des vorherigen Abschnitts folgt, dass der Abstand der Brennpunkte $2c$ kleiner sein muss als $r_1 + r_2 = 2r$. Dies ist das Kriterium, das mir schon zeigt, wie groß ich den Hauptachsenradius r mindestens wählen sollte,

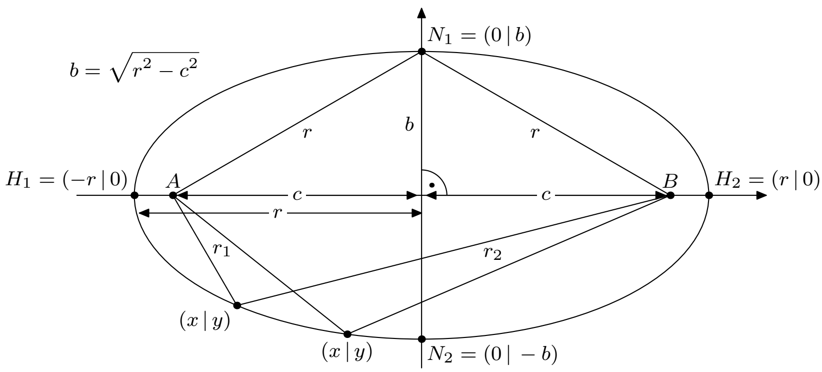

Abb. 4.22 Die Punkte $(x|y)$ auf einer Ellipse haben die Eigenschaft, dass die Summe r_1+r_2 der Abstände r_1 und r_2 zu den Brennpunkten A und B konstant $2r$ ist.

um mit vorgegebenen Brennpunkten eine Ellipse konstruieren zu können. Sind nun aber Punkte auf einer zu konstruierenden Ellipse vorgegeben, so ist die Situation schwieriger als bei Kreisen, denn drei Punkte reichen im Allgemeinen nicht aus, um eine Ellipse festzulegen. Zur Information verrate ich Ihnen hier, dass man dazu vier *geeignete* Punkte braucht. Weiter kann ich beobachten, dass Ellipsen symmetrische Figuren hinsichtlich der Haupt- und Nebenachse sind.

Betrachte ich den Spezialfall, dass die beiden Brennpunkte zusammenfallen $A = B = (0|0)$, beziehungsweise $c = 0$, so kann ich die Ellipsengleichung zunächst in die Form $\sqrt{x^2+y^2} + \sqrt{x^2+y^2} = 2r$ und dann als $2\,\sqrt{x^2+y^2} = 2r$ schreiben. Kürzen der 2 und nachfolgendes Quadrieren beider Seiten ergibt hier $x^2+y^2 = r^2$ – es handelt sich also bei diesem Spezialfall um einen Kreis mit Radius r und Mittelpunkt $(0|0)$. An dieser Stelle tritt die Frage auf, ob ich die obige Form der Ellipsengleichung nicht auch deutlich einfacher schreiben könnte. In der Tat: Durch relativ trickreiches Umformen gelangt man auf die alternative Darstellung

$$\frac{x^2}{r^2} + \frac{y^2}{b^2} = 1,$$

wobei nach wie vor $b^2 = r^2 - c^2$ gilt. Dies nennt man auch die **Mittelpunktsgleichung** der Ellipse. Mir fällt auf, dass die Quadrate der Länge der Haupt- und Nebenachse hier jeweils im Nenner auftreten und der Spezialfall $c = 0$ (also $b = r$) – wie schon oben gesehen – den Kreis mit Radius r und Mittelpunkt $(0|0)$ darstellt. Auch die Parameterdarstellung des Kreises lässt sich für Ellipsen verallgemeinern:

$$x = r\cos(t) \text{ und } y = b\sin(t),\ t \in [0, 2\pi).$$

Beispiel 4.16:
Die Ellipse $\mathcal{E}$ durch den Punkt $X = (6|4)$ mit Mittelpunkt $(0|0)$ und Hauptachsenradius $r = 10$ kann ich wie folgt berechnen: Da X auf $\mathcal{E}$ liegt, gilt $6^2/10^2 + 4^2/b^2 = 1$ beziehungsweise $16/b^2 = 1-36/100 = 16/25$. Also: $b^2 = 25$ und damit $c = \sqrt{r^2 - b^2} = \sqrt{75} = 5\,\sqrt{3}$. Die Ellipse hat also die Gleichung $x^2/100 + y^2/25 = 1$ und deren Parameterdarstellung sieht wie folgt aus: $x = 10\cos(t)$ und $y = 5\sin(t),\ t \in [0, 2\pi)$. Die

Brennpunkte von $\mathcal{E}$ sind $(5\sqrt{3}|0)$ und $(-5\sqrt{3}|0)$, während die Nebenscheitel von $\mathcal{E}$ die Punkte $(0|5)$ und $(0|-5)$ sind. ■

Abschließend verrate ich Ihnen noch, dass man den **Flächeninhalt** F einer **Ellipse** mit Hauptachsenradius r und Nebenachsenradius b gemäß der Formel $F = \pi rb$ berechnen kann und man für den **Umfang** U einer solchen **Ellipse** die Näherungsformel $U \approx \pi\ (3/2(r+b) - \sqrt{rb})$ kennt. Ähnliche Näherungsformeln kann man für die Länge eines **Ellipsenbogens** angeben. Sie sehen, dass dies komplizierter als für Kreise aussieht – tatsächlich ist die Herleitung dieser Aussagen gar nicht so einfach und benötigt schon recht fundierte Kenntnisse aus der **Integralrechnung** und über so genannte **elliptische Integrale**. Diese besitze ich zwar – und in Beispiel 6.43 werde ich Ihnen zumindest zeigen, wie man auf die Flächeninhaltsformel für Ellipsen kommt – ich für meinen Teil werde aber jetzt die Schreibarbeit unterbrechen und eine kleine Grillparty vorbereiten.

Übungsaufgabe 4.21:
Beim sommerlichen Grillen stellen Sie fest, dass die im gefrorenen Zustand runden Hamburger mit 20 cm Durchmesser auf eine Ellipse mit zwei Drittel Größe (Fläche) zusammenschmurgeln, deren Hauptachsenradius doppelt so groß wie der Nebenachsenradius ist. Das Problem lässt Sie nicht mehr los: Sie wollen unbedingt wissen, wie groß der Haupt- und Nebenachsenradius der so entstehenden Hamburger-Ellipsen ist. ■

5 Einführung in die Lineare Algebra

Inzwischen sind Sie doch schon ein ganz schönes Stück weitergekommen. Sie haben sich an die alten unangenehmen Dinge wie Vorklammern, Ausmultiplizieren oder auch Logarithmen erinnert, Sie haben gesehen, wie man mit bestimmten Funktionen umgeht, Sie wurden mit verschiedenen Arten von Gleichungen traktiert und sogar ein wenig Geometrie wurde Ihnen geboten. Das war gar nicht mal so wenig, und wenn Sie einmal genauer hinsehen, dann werden Sie feststellen, dass die ersten drei der bisherigen Kapitel eine deutliche Gemeinsamkeit haben: Überall, selbst bei der Besprechung der elementaren Rechenmethoden, mussten wir den Bereich des reinen Zahlenrechnens verlassen und stattdessen auf das Rechnen mit Buchstaben zurückgreifen. Sicher mag das Rechnen mit konkreten Zahlen angenehmer sein, aber wenn es beispielsweise darum geht, allgemeine Regeln zu formulieren oder irgendwelche Formeln zu vereinfachen, kommt man um das Rechnen mit den so genannten Variablen einfach nicht herum. Im Allgemeinen bezeichnet man dieses Buchstabenrechnen mit dem alten Begriff „Algebra“.

Nun konnten Sie aber feststellen, dass diese Algebra Sie auf die verschiedensten Wege geführt hat. Immerhin versammeln wir unter diesem Begriff so unterschiedliche Dinge wie das Rechnen mit Logarithmen und das simple Addieren, so unterschiedliche Gleichungstypen wie die einfache lineare Gleichung und die ungleich kompliziertere Wurzel- oder gar Exponentialgleichung. Wir haben es also mit einem ziemlich weit gefassten Gebiet zu tun, und in diesem Kapitel will ich das Gebiet etwas einschränken und mich nur noch mit der „Linearen Algebra“ befassen. Wenn Sie sich einmal kurz die linearen Gleichungen ins Gedächtnis rufen, dann sind das gute Aussichten, denn so lästige Operationen wie Quadrieren oder Logarithmieren waren dort beim besten Willen nicht zu finden: Lineare Gleichungen sehen im Wesentlichen so aus wie $5x + 3 = 17$, und kein x^2 oder $\log_2 x$ der Welt hat das Recht, sich dorthin zu verirren. Alles, was ich in diesem Kapitel betrachte, wird **linear** sein, und das bedeutet, dass die vorkommenden Rechenoperationen nicht über die Grundrechenarten hinausgehen.

Klingt gut, oder? Im Leben ist aber nichts umsonst (wenn auch manches vergeblich). Während ich mich bei den zulässigen Rechenoperationen sehr starken Einschränkungen unterwerfen werde, darf ich auf der anderen Seite, sozusagen durch die Hintertür, etwas mehr Spannung in die Sache bringen, sonst könnte ich mir dieses Kapitel nämlich schenken. Bei der Lösung linearer und auch anderer Gleichungen gab es immer *eine einzige* Unbekannte, deren Wert Sie herausfinden mussten. Gelegentlich gab es mehrere Lösungen, aber das waren immer Lösungen für die eine gesuchte

Unbekannte. Es kommt leider in der Realität häufiger vor, dass nicht nur eine Größe unbekannt ist, sondern gleich zwei, drei oder siebzehn auf einmal. Über solche Situationen haben wir hier noch gar nicht gesprochen, und inzwischen haben Sie genug Algebra gelernt, um auch solche Probleme anzugehen. Wir machen also in diesem Kapitel über Lineare Algebra ein kleines Geschäft: Einerseits verzichte ich darauf, unangenehme Rechenoperationen ins Spiel zu bringen und beschränke mich strikt aufs Lineare, andererseits darf ich die Anzahl der verwendeten Größen, die bisher meistens bei eins lag, ganz nach Belieben erhöhen – sonst wäre die Lineare Algebra auch zu langweilig und Sie könnten das Kapitel getrost überschlagen.

Nach dieser langen Vorrede sehen wir uns am Beispiel der Vektoren an, was ich eigentlich meine.

5.1 Vektoren

In der Linearen Algebra hat man es immer wieder mit Vektoren zu tun, und auch ich will sie Ihnen nicht vorenthalten. Sie werden gleich in Beispiel 5.1 sehen, dass man mit ihrer Hilfe verschiedene Teilinformationen zu einer Gesamtinformation zusammenfassen kann:

Beispiel 5.1:
Stellen Sie sich vor, Sie sind glücklicher Besitzer eines Unternehmens *MeineFirma*, das drei verschiedene Produkte A_1, A_2 und A_3 herstellt, wobei ich die Ausgestaltung der Produkte Ihrer Fantasie und Ihren Vorlieben überlasse. Nun wird *MeineFirma* sicher nicht von allen drei Produkten jeweils die gleiche Menge herstellen, denn man kann nicht erwarten, dass der Absatzmarkt beispielsweise genauso viele Toilettendeckel wie Schreibtischlampen verkraften kann. Ihr Unternehmen wird also pro Produkt eine bestimmte Absatzmenge aufweisen, und wir nehmen einmal an, dass die Absatzmenge für das erste Produkt bei 18 Einheiten liegt, für das zweite bei 34 und für das dritte bei 17. Dabei können mit Einheiten natürlich auch Tausenderpacks oder irgendetwas anderes gemeint sein, das braucht uns jetzt nicht zu interessieren. Die Frage ist: Wie kann ich diese Situation sinnvoll formelmäßig darstellen, am besten in einer Datenstruktur, die ich auch mit einem Computer abspeichern kann?

Es gibt aber nicht nur Absatzmengen; zur Berechnung des Umsatzes brauchen Sie natürlich noch die Stückpreise der drei Produkte. Wir nehmen an, die Preise lauten 12 Geldeinheiten für eine Einheit des ersten Produkts, 9 Geldeinheiten beim zweiten Produkt und 13 beim dritten. Wie kann man nun diese drei Stückpreise in einer einzigen Datenstruktur zusammenfassen? ■

Sie sehen vermutlich schon, worauf das Ganze hinausläuft. Um drei Zahlenwerte, die irgendwie eine Einheit bilden sollen, zusammenzufassen, ohne die einzelnen Werte zu verlieren, mache ich einfach einen so genannten dreidimensionalen Vektor aus ihnen. Wir schauen uns gleich an, wie so etwas aussieht:

Dreidimensionaler Vektor

Unter einem dreidimensionalen Vektor versteht man eine Größe der Form

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

wobei $x_1 \in \mathbb{R}, x_2 \in \mathbb{R}, x_3 \in \mathbb{R}$ gilt. Die Menge aller dreidimensionalen Vektoren wird mit $\mathbb{R}^3$ bezeichnet.

In einem dreidimensionalen Vektor fasst man also nur drei reelle Zahlen zusammen, nicht mehr und nicht weniger. Oft benennt man solche Vektoren mit fett gedruckten Kleinbuchstaben, und ich will von dieser Regel hier auch nicht abweichen, deshalb heißt der Vektor $\mathbf{x}$ und nicht einfach nur x. Die beiden Probleme aus dem Beispiel 5.1 kann ich jetzt ganz leicht mithilfe von zwei dreidimensionalen Vektoren lösen, nämlich dem „Absatzvektor“

$$\mathbf{a} = \begin{pmatrix} 18 \\ 34 \\ 17 \end{pmatrix}$$

sowie einem „Preisvektor“

$$\mathbf{p} = \begin{pmatrix} 12 \\ 9 \\ 13 \end{pmatrix}.$$

Es spielt also keine Rolle, ob man in den Vektor Absatzmengen, Preise oder sonst etwas einträgt; die Struktur bleibt immer die gleiche.

Bestimmt brennt Ihnen jetzt die Frage auf den Nägeln, was man denn nun mit einem Unternehmen anstellen soll, das nur zwei Produkte oder aber gleich vier oder fünf Produkte herstellt. Es wäre wenig sinnvoll, für jede Dimension eine neue Definition aufzuschreiben, zumal es unendlich viele natürliche Zahlen gibt und ich nicht unendlich viele Seiten zur Verfügung habe. Wir brauchen daher einen Vektorbegriff, der mit einem Schlag alle möglichen Fälle, also alle denkbaren Dimensionen erledigt, und so etwas macht man, indem man von n-dimensionalen Vektoren spricht. Wenn man sich nicht darauf festlegen möchte, wie viele Einträge der Vektor haben soll, dann nimmt man sich die Freiheit, die Anzahl der Einträge mit der Variablen n zu bezeichnen, und je nach Bedarf ist dann $n = 3$ oder $n = 2$ oder $n =$ irgendetwas anderes.

Vektor

Unter einem n-dimensionalen Vektor versteht man eine Größe der Form

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

wobei $x_1 \in \mathbb{R}, ..., x_n \in \mathbb{R}$ gilt. Die Menge aller n-dimensionalen Vektoren wird mit $\mathbb{R}^n$ bezeichnet.

Meistens spricht man einfach nur von einem Vektor und lässt den Zusatz n-dimensional weg, denn irgendeine Dimension muss der Vektor ja sowieso haben. Falls Sie die drei Punkte irritieren, die mitten im Vektor herumstehen: Die symbolisieren nur, dass hier insgesamt n Zahlen untereinander stehen sollen, und da man den Wert von n in diesem allgemeinen Fall nicht kennt, bequemt man sich zu dieser einfachen Dreipunktschreibweise. Zwei- und dreidimensionale Vektoren erlauben übrigens noch eine ganz andere Darstellung, die etwas mit Geometrie zu tun hat; darauf kommen wir aber erst im vierten Abschnitt dieses Kapitels zu sprechen.

Jetzt können wir also Vektoren aufschreiben, aber das ist auch schon alles. Natürlich könnten Sie einen Vektor auch durch die Gegend werfen oder grün anstreichen, aber wer interessiert sich schon für grün angestrichene Vektoren? Wir haben es hier mit Objekten aus Zahlen zu tun, und es sollte möglich sein, mit ihnen zu rechnen. Im nächsten Beispiel sehen Sie, wie man Vektoren sinnvoll addieren kann.

Beispiel 5.2:
Ich verwende wieder den Absatzvektor **a** aus Beispiel 5.1. Nimmt man an, dass er den Absatz im ersten Quartal des laufenden Jahres beschreibt, und hat man weiterhin einen Vektor **b**, in dem die Produktion des zweiten Quartals gespeichert ist, so kann man sich fragen, wie viel wohl im Laufe des ersten Halbjahres Ihre Firma *MeineFirma* zustande gebracht hat. Mit den Vektoren

$$\mathbf{a} = \begin{pmatrix} 18 \\ 34 \\ 17 \end{pmatrix} \text{ und } \mathbf{b} = \begin{pmatrix} 19 \\ 33 \\ 21 \end{pmatrix}$$

ergibt sich für das erste Halbjahr eine Ausbringungsmenge von

$$\begin{pmatrix} 18+19 \\ 34+33 \\ 17+21 \end{pmatrix} = \begin{pmatrix} 37 \\ 67 \\ 38 \end{pmatrix}.$$

Und schon wissen Sie, wie man Vektoren addiert. ■

Nach dem Schema aus Beispiel 5.2 kann man immer vorgehen, wenn man zwei Vektoren addieren will: Man addiert einfach ihre zueinander passenden **Komponenten**.

Das kann allerdings nur funktionieren, wenn man es mit Vektoren gleicher Dimension zu tun hat, sonst bleibt am Ende bei einem Vektor eine Komponente übrig, und keiner kann ihr sagen, wo sie ihr müdes Haupt betten soll. Wir addieren also nur Vektoren gleicher Dimension, von gemischten Kombinationen lassen Sie besser die Finger.

Addition und Subtraktion von Vektoren

Sind

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ und } \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

zwei n-dimensionale Vektoren, dann setzt man

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \pm \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 \pm y_1 \\ \vdots \\ x_n \pm y_n \end{pmatrix}.$$

Man addiert oder subtrahiert also zwei Vektoren gleicher Dimension, indem man die jeweiligen Komponenten der Vektoren, also ihre Einträge, addiert oder subtrahiert.

Wie Sie sehen, habe ich die Subtraktion gleich miterledigt, was wohl nicht sehr überraschend war. Ein Beispiel zur Addition haben Sie ja schon oben gesehen. Trotzdem zeige ich Ihnen gleich noch mehr Beispiele, aber zuerst will ich noch einen Schritt weitergehen und eine weitere Rechenmethode vorführen: die Multiplikation mit einer Zahl. Auch dahinter steckt aber rein gar nichts Kompliziertes, wie man sich leicht überlegen kann. Nehmen wir zum Beispiel den zweidimensionalen Vektor $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, so gilt natürlich $\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$ nach der Regel für die Addition, die Sie gerade gelernt haben. Aber ein Ding plus das gleiche Ding ist immer noch das zweifache Ding, und folglich habe ich den Vektor $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ mit der Zahl 2 multipliziert und dabei den Vektor $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ erhalten. Sie sehen wohl, worauf ich hinaus will, denn es gilt:

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 \\ 2 \cdot 2 \end{pmatrix}.$$

Man multipliziert also einen Vektor mit einer Zahl, indem man jede seiner Komponenten mit dieser Zahl multipliziert.

Multiplikation eines Vektors mit einer Zahl

Für eine reelle Zahl λ und einen Vektor $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ setzt man

$$\lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_1 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}.$$

Stören Sie sich nicht an dem ominösen λ, das ist nur die griechische Fassung eines l, und es hat sich ein wenig eingebürgert, im Zusammenmhang mit Vektoren die simplen Zahlen mit griechischen Buchstaben zu bezeichnen, damit man sie leicht von den Vektoren unterscheiden kann. Ausgesprochen wird es übrigens als „lambda“.

Jetzt wird es aber Zeit für das eine oder andere Beispiel:

Beispiel 5.3:
Es gilt:

$$\begin{pmatrix} -5 \\ 3 \\ 2 \\ 8 \end{pmatrix} + \begin{pmatrix} 9 \\ 7 \\ 1 \\ 17 \end{pmatrix} = \begin{pmatrix} -5+9 \\ 3+7 \\ 2+1 \\ 8+17 \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \\ 3 \\ 25 \end{pmatrix}.$$

Weiterhin ist

$$3 \cdot \begin{pmatrix} 17 \\ 8 \\ 12 \end{pmatrix} = \begin{pmatrix} 3 \cdot 17 \\ 3 \cdot 8 \\ 3 \cdot 12 \end{pmatrix} = \begin{pmatrix} 51 \\ 24 \\ 36 \end{pmatrix}.$$

Dagegen ist die Addition

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

nicht möglich, da die beiden Vektoren verschiedene Dimensionen haben. ■

Ein wenig Übung gefällig? Daran soll es nicht scheitern:

Übungsaufgabe 5.1:
Berechnen Sie

$$\begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} - 4 \cdot \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix} + 2 \cdot \begin{pmatrix} 12 \\ 0 \\ 1 \end{pmatrix} \quad \text{und} \quad -\begin{pmatrix} 3 \\ -4 \end{pmatrix} + 3 \cdot \left[\begin{pmatrix} 2 \\ 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} -1 \\ 4 \end{pmatrix} \right].$$

■

Die Grundrechenarten sollten damit schon erledigt sein, denn was soll man sonst noch machen außer addieren, subtrahieren und multiplizieren? Die Antwort liegt nahe:

Eine Division wäre auch nicht schlecht. Ich muss Ihnen aber leider mitteilen, dass es nicht möglich ist, einen Vektor durch einen anderen zu teilen; dabei kann nichts Vernünftiges herauskommen. Was allerdings funktioniert, ist das Teilen eines Vektors durch eine Zahl, und das ist genau genommen kaum der Rede wert, weil wir es eigentlich schon besprochen haben. Ein kleines Beispiel zeigt, was ich meine.

Beispiel 5.4:
Was kommt heraus, wenn man $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ durch 2 dividiert? Wie Sie wissen, entspricht das Dividieren durch 2 dem Multiplizieren mit $\frac{1}{2}$, und deshalb lautet das Ergebnis:

$$\frac{1}{2} \cdot \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \qquad \blacksquare$$

Man dividiert also einen Vektor durch eine Zahl, indem man ihn mit dem Kehrwert der Zahl multipliziert – und das Multiplizieren mit einer Zahl ist für Sie inzwischen ein alter Hut. Ich sagte ja, dass Sie die Division durch eine Zahl im Grunde schon kennen.

Jetzt aber zu etwas wirklich Neuem. Multipliziert man einen Vektor mit einer Zahl, so kommt dabei wieder ein Vektor heraus. Es gibt aber eine Art von Multiplikation zwischen Vektoren, die als Ergebnis zu einer Zahl führt, die also das bisher Gewohnte sozusagen gerade auf den Kopf stellt. Sehen wir uns an, wie so etwas passieren kann.

Beispiel 5.5:
Ich gehe wieder davon aus, dass Ihr Unternehmen *MeineFirma* drei Artikel A_1, A_2 und A_3 verkauft. Von A_1 werden 18 Einheiten verkauft, von A_2 34 Einheiten und von A_3 17 Einheiten. Die entsprechenden Stückpreise betragen 12 Euro für eine Einheit A_1, 9 Euro für eine Einheit A_2 und 13 Euro für eine Einheit A_3. Der Umsatz Ihrer aufstrebenden Firma berechnet sich dann aus:

$$\text{Umsatz } = 18 \cdot 12 + 34 \cdot 9 + 17 \cdot 13 = 743.$$

Nun kann man aber die gegebenen Informationen wieder mithilfe von Vektoren aufschreiben, denn ich habe hier einen Absatzvektor $\mathbf{a}$ und einen Preisvektor $\mathbf{p}$ mit

$$\mathbf{a} = \begin{pmatrix} 18 \\ 34 \\ 17 \end{pmatrix} \text{ und } \mathbf{p} = \begin{pmatrix} 12 \\ 9 \\ 13 \end{pmatrix}.$$

Den Umsatz habe ich berechnet, indem ich die erste Komponente des einen Vektors multipliziert habe mit der ersten Komponente des anderen, dann die zweite Komponente des ersten mit der zweiten Komponente des anderen und schließlich die dritte Komponente des ersten mit der dritten Komponente des anderen – danach wurden die einzelnen Produkte addiert. Weil hier einiges multipliziert wird, spricht man von einem Produkt, und weil sich am Ende eine Zahl ergibt, könnte man das Ganze ein Zahlprodukt nennen. Macht man aber nicht. Der vornehmere Ausdruck für „Zahl“ lautet „Skalar“, und daher bezeichnet man die Operation, die ich gerade durchgeführt habe, als das Skalarprodukt der beiden Vektoren **a** und **b**. In Formeln geschrieben sieht das dann so aus:

$$\begin{pmatrix} 18 \\ 34 \\ 17 \end{pmatrix} \cdot \begin{pmatrix} 12 \\ 9 \\ 13 \end{pmatrix} = 18 \cdot 12 + 34 \cdot 9 + 17 \cdot 13 = 743.$$

■

Dieses Verfahren funktioniert natürlich nicht nur für zwei dreidimensionale Vektoren, sondern immer dann, wenn Sie es mit zwei Vektoren gleicher Dimension zu tun haben.

Skalarprodukt

Sind $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ zwei n-dimensionale Vektoren, dann setzt man

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 \cdot y_1 + \cdots + x_n \cdot y_n.$$

Man bezeichnet diese Operation als das Skalarprodukt der beiden Vektoren.

Das Skalarprodukt von zwei Vektoren dient also dazu, die Inhalte dieser Vektoren auf bestimmte Weise miteinander zu kombinieren. Dazu müssen allerdings beide Vektoren gleich lang sein, also die gleiche Dimension haben. Damit Sie sich an diese Konstruktion gewöhnen, sehen wir uns noch weitere Beispiele an.

Beispiel 5.6:
Es gilt:

$$\begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix} = 2 \cdot 1 + 5 \cdot (-3) + (-1) \cdot 2 = -15.$$

Dagegen ist die Berechnung von

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$$

nicht möglich, da die beiden Vektoren verschiedene Längen, also verschiedene Dimensionen, haben. ■

Sie wissen jetzt also, wie man zwei Vektoren miteinander multipliziert und dabei eine Zahl herausbekommt. Manchmal ist das Leben aber auch etwas komplizierter und es reicht nicht, sich nur mit zwei Vektoren zu begnügen. Stellen Sie sich beispielsweise vor, Ihnen liegt nicht nur ein Vektor vor, der die Absatzmenge beschreibt, und ein weiterer Vektor, in dem die Stückpreise stehen, sondern der Preisvektor ist aufgeteilt in einen Nettopreisvektor und einen Mehrwertsteuervektor; ein Blick auf Ihre letzte Rechnung aus der Autowerkstatt zeigt, dass das im richtigen Leben jederzeit vorkommen kann. In diesem Fall muss man den Absatzmengenvektor skalarmultiplizieren mit der Summe der beiden anderen Vektoren, denn beide zusammen ergeben erst den Stückpreis. Es kann also manchmal nötig sein, Skalarprodukte in komplizierteren Situationen auszurechnen, indem man nicht nur einen Vektor mit einem anderen multipliziert, sondern einen Vektor mit der Summe von zwei oder gar mehreren anderen. Sobald aber solche Verwicklungen auftreten, bei denen mehrere Rechenoperationen beteiligt sind, schreibt man gerne Rechenregeln auf, damit man jederzeit weiß, was erlaubt ist und was nicht. Die Grundsituation sehen wir uns zunächst an einem Beispiel an:

Beispiel 5.7:
Einerseits ist

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \left[\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \right] = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} = 2 + 0 - 3 = -1,$$

wobei ich nach dem ersten Gleichheitszeichen die beiden Vektoren innerhalb der eckigen Klammern addiert habe. Andererseits könnte man auch versuchen, die Klammer auszumultiplizieren, wie Sie das schon im ersten Kapitel bei den üblichen Regeln für das Zahlenrechnen gelernt haben. Dabei ergibt sich:

$$\begin{aligned} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \left[\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \right] &= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix} \\ &= 0 + 2 - 3 + 2 - 2 + 0 = -1. \end{aligned}$$

Ich erhalte also auf beiden Wegen das gleiche Ergebnis, und das ist durchaus kein Zufall. ■

Das Ausmultiplizieren geht beim Skalarprodukt genauso gut wie beim vertrauten Multiplizieren von Zahlen. Und das ist noch nicht alles, denn in Wahrheit gelten *alle* Regeln, die Sie über das Multiplizieren von Zahlen kennen, auch für das Skalarprodukt.

Rechenregeln für das Skalarprodukt

Für $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gelten die folgenden Regeln:
a) $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$;
b) $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z}$;
c) $\lambda \cdot (\mathbf{x} \cdot \mathbf{y}) = (\lambda \cdot \mathbf{x}) \cdot \mathbf{y}$.

Beachten Sie dabei, dass in Regel b) die beiden auftretenden Pluszeichen zwei verschiedene Bedeutungen haben. Auf der linken Seite der Gleichung handelt es sich um die Addition von Vektoren, auf der rechten Seite um die schlichte Addition gewöhnlicher Zahlen. Trotzdem können Sie die Rechnung ganz nach Ihrem persönlichen Geschmack auf die eine oder die andere Weise ausführen und erhalten beide Male das gleiche Ergebnis. In Regel c) hat dagegen der Multiplikationspunkt verschiedene Bedeutungen: Erstens die Multiplikation einer Zahl mit einem Vektor ($\lambda \cdot x$), zweitens das Skalarprodukt zweier Vektoren ($x \cdot y$) und drittens die schlichte Multiplikation zweier Zahlen (Produkt aus λ und $x \cdot y$). Die Regel zeigt aber, dass es nichts schadet, so grundverschiedene Operationen mit dem gleichen Zeichen zu benennen, denn schließlich geht am Ende alles gut aus.

Jetzt wird es wieder Zeit für ein wenig Übung.

Übungsaufgabe 5.2:

a) Berechnen Sie

$$\begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -12 \end{pmatrix} \text{ und } \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 5 \\ 6 \end{pmatrix}.$$

b) Berechnen Sie das Skalarprodukt

$$\begin{pmatrix} -1 \\ 2 \end{pmatrix} \cdot \left(\begin{pmatrix} 3 \\ 4 \end{pmatrix} + \begin{pmatrix} -2 \\ 5 \end{pmatrix} \right)$$

auf zwei Arten. ■

Der bisherige Stand ist schon gar nicht so schlecht, denn die wesentlichen Rechenoperationen für Vektoren haben wir jetzt zusammen. Auf ihrer Basis können Sie auch etwas komplexere Aufgabenstellungen angehen, wie das folgende Beispiel zeigt.

Beispiel 5.8:

Ihr altvertrautes Unternehmen *MeineFirma* stellt in zwei Hallen H_1 und H_2 jeweils drei Produkte A_1, A_2 und A_3 her. Aufgrund von Qualitätsunterschieden haben die Produkte aus H_1 einen niedrigeren Preis als die Produkte aus H_2. So kosten Produkt A_1 aus Halle H_1 12 Euro, Produkt A_2 8 Euro und Produkt A_3 17 Euro. Dagegen kosten die drei Produkte, wenn sie in Halle H_2 hergestellt wurden, 17 Euro, 10 Euro und 20 Euro. Wie lautet der Gesamtumsatz, wenn in H_1 von jedem Produkt 10 Einheiten und in H_2 von jedem Produkt 15 Einheiten hergestellt werden?

Der Preisvektor von H_1 lautet $\begin{pmatrix} 12 \\ 8 \\ 17 \end{pmatrix}$ der Preisvektor von H_2 lautet $\begin{pmatrix} 17 \\ 10 \\ 20 \end{pmatrix}$. Die Ergebnisse pro Artikel erhalten wir als eine Kombination beider Vektoren, nämlich:

$$10 \cdot \begin{pmatrix} 12 \\ 8 \\ 17 \end{pmatrix} + 15 \cdot \begin{pmatrix} 17 \\ 10 \\ 20 \end{pmatrix} = \begin{pmatrix} 120 \\ 80 \\ 170 \end{pmatrix} + \begin{pmatrix} 255 \\ 150 \\ 300 \end{pmatrix} = \begin{pmatrix} 375 \\ 230 \\ 470 \end{pmatrix}.$$

Der Ergebnisvektor beschreibt in seiner ersten Komponente den Umsatz für Produkt A_1, dann den Umsatz für Produkt A_2 und schließlich den Umsatz für Produkt A_3. Den Gesamtumsatz erhält man dann aus

$$\text{Umsatz } = 375 + 230 + 470 = 1075.$$

■

Was habe ich hier mit den beiden auftretenden Vektoren angestellt? Ich habe jeden mit einer Zahl multipliziert und anschließend die resultierenden Vektoren addiert. Das ist zwar eine ganz einfache Sache, aber weil sie in der Vektorrechnung unglaublich häufig vorkommt, hat man ihr einen eigenen Namen gegeben und nennt das eine Linearkombination: Offenbar werden hier die Vektoren kombiniert, und die Art der Kombination beruht auf harmlosen linearen Operationen wie Addition und Multiplikation mit einer Zahl. Natürlich kann man solche Linearkombinationen nicht nur aus zwei Vektoren bilden, sondern aus beliebig vielen. Das Prinzip ist aber immer gleich.

Linearkombination

Sind $\mathbf{x}_1, ..., \mathbf{x}_m \in \mathbb{R}^n$ Vektoren und $c_1, ..., c_m \in \mathbb{R}$ reelle Zahlen, dann heißt

$$c_1 \cdot \mathbf{x}_1 + \cdots + c_m \cdot \mathbf{x}_m$$

eine Linearkombination der Vektoren $\mathbf{x}_1, ..., \mathbf{x}_m$.

Das ist nur ein hochtrabender Name für eine Selbstverständlichkeit, was auch das nächste Beispiel noch einmal zeigen wird:

Beispiel 5.9:

$$2 \cdot \begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix} - 3 \cdot \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 0 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} 10 \\ 6 \\ 2 \end{pmatrix} - \begin{pmatrix} -3 \\ 6 \\ 12 \end{pmatrix} + \begin{pmatrix} 0 \\ 8 \\ 9 \end{pmatrix} = \begin{pmatrix} 13 \\ 8 \\ -1 \end{pmatrix}.$$

Hier wird eine Linearkombination aus drei Vektoren berechnet, wobei der dritte Vektor keinen Vorfaktor besitzt, aber da man natürlich immer die Zahl 1 vor einen Vektor schreiben kann, passt die Rechnung trotzdem in das Schema der Linearkombination.

■

Der letzte Begriff, mit dem ich Sie in diesem Abschnitt plagen muss, ist der Begriff der linearen Abhängigkeit. Auch er hat einen recht praktischen Hintergrund. Stellen Sie sich vor, Sie wollen einen aus drei Zutaten bestehenden Kuchen backen, wobei die Zutaten natürlich in einem bestimmten Mischungsverhältnis stehen müssen. Hat man also beispielsweise zwei Einheiten der ersten Zutat, drei Einheiten der zweiten Zutat und eine Einheit der dritten Zutat zu mischen, so können Sie offenbar das Zutatenschema durch den dreidimensionalen Vektor $\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$ beschreiben. Wie sieht es nun aus bei einem Kuchen, für den Sie vier Einheiten der ersten Zutat, sechs Einheiten der zweiten Zutat und zwei Einheiten der dritten Zutat brauchen? Auf den ersten Blick sieht das aus wie ein ganz anderer Kuchen, aber eigentlich ist es der gleiche, nur die doppelte Menge, denn Sie haben nur die jeweiligen Einheiten verdoppelt, das Mischungsverhältnis aber gleich gelassen. Für den neuen Zutatenvektor gilt daher auch die Gleichung $\begin{pmatrix} 4 \\ 6 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$. Da der neue Vektor sich auf diese Weise leicht aus dem alten berechnen lässt, also in Wahrheit keine neuen Informationen bietet, sagt man, dass die beiden Vektoren **linear abhängig** sind.

Das geht auch mit mehr als zwei Vektoren. Gehen wir beispielsweise davon aus, dass Ihr schon häufiger bemühtes Unternehmen *MeineFirma* Fruchtsäfte herstellt, dann könnte etwa ein Saft darauf beruhen, dass man zwei Teile Orangensaft mit einem Teil Zitronensaft mischt, und der zweite Saft im Sortiment macht es genau umgekehrt: für ihn mischt man einen Teil Orangensaft mit zwei Teilen Zitronensaft. Die entsprechenden Mischungsvektoren lauten dann $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. Für beide Saftmischungen brauchen Sie jeweils eine Maschine, die leider auch Geld kostet, und ganz schlimm wird es, wenn Sie feststellen, dass viele Kunden auch gerne eine Saftmischung hätten, bei der man Orangen- und Zitronensaft im Verhältnis eins zu eins mischt. Bevor Sie nun aber in düsteren Nächten die Kosten für eine neue Maschine kalkulieren, sollten Sie sich auf die Vektorrechnung besinnen und daran denken, dass

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{3} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \frac{1}{3} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

gilt. Sobald Sie also die schon im Sortiment vorhandenen Saftmischungen zu gleichen Teilen mischen, haben Sie schon den neuen Kundenwunsch erfüllt, und das lässt sich vermutlich leichter bewerkstelligen als der Bau einer neuen Fabrikhalle. Änlich sieht es aus mit einem gewünschten Mischungsverhältnis von fünf zu vier, denn es gilt:

$$\begin{pmatrix} 5 \\ 4 \end{pmatrix} = 2 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Sie sehen, worauf es ankommt. Man wird auch hier die Vektoren

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

als linear abhängig bezeichnen, denn Sie können den ersten Vektor als Linearkombination der beiden anderen darstellen. Und natürlich sind aus dem gleichen Grund auch

$$\begin{pmatrix} 5 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

linear abhängig.

Da Sie nun wissen, was lineare Abhängigkeit bedeutet, kann ich den Begriff auch offiziell formulieren.

Lineare Abhängigkeit und Unabhängigkeit

Die Vektoren $\mathbf{x}_1, ..., \mathbf{x}_m \in \mathbb{R}^n$ heißen *linear abhängig*, falls man einen von ihnen als Linearkombination der restlichen $m-1$ Vektoren darstellen kann. Vektoren, die nicht linear abhängig sind, heißen *linear unabhängig*.

Das ist noch mal das Gleiche wie eben, nur diesmal allgemein für eine beliebige Anzahl von Vektoren formuliert. Sehen wir uns dazu gleich einige Beispiele an:

Beispiel 5.10:

a) $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sind linear unabhängig, denn es ist nicht möglich, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ als $c \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ oder $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ als $c \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ mit einer reellen Zahl c darzustellen.

b) Die Vektoren
$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 7 \end{pmatrix}$$
sind linear abhängig, denn
$$\begin{pmatrix} 4 \\ 7 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$
Der Vektor $\begin{pmatrix} 4 \\ 7 \end{pmatrix}$ bietet also sozusagen keine neuen Informationen, die nicht schon in den beiden anderen Vektoren enthalten wären. Deshalb spricht man auch von Abhängigkeit. Allerdings entsteht hier die Frage, wie ich eigentlich auf die Vorfaktoren der beiden anderen Vektoren gekommen bin, die ich für die Linearkombination brauche. Beim derzeitigen Stand der Dinge scheinen sie noch vom Himmel zu fallen, und wir werden uns überlegen müssen, wie man an sie herankommt.

c) Die Vektoren
$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$
sind linear unabhängig, denn es ist nicht möglich, $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ als $c \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ oder $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ als $c \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ mit einer reellen Zahl c darzustellen.

d) Die Vektoren
$$\begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}$$

sind linear unabhängig. Zum Test auf lineare Abhängigkeit würde man z. B. ansetzen:

$$\begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix} = c_1 \cdot \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} + c_2 \cdot \begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix},$$

denn ein Vektor muss ja bei linearer Abhängigkeit als Linearkombination der anderen Vektoren darstellbar sein. Wenn man nun die Gleichungen für die einzelnen Komponenten aufschreibt, dann heißt das:

$$\begin{array}{rcrcl} c_1 & + & c_2 & = & 5 \\ 4c_1 & + & 2c_2 & = & 1 \\ 2c_1 & + & 6c_2 & = & 3. \end{array}$$

Ich habe also ein so genanntes lineares Gleichungssystem mit drei Gleichungen und zwei Unbekannten. Falls dieses System irgendeine Lösung hat, falls es also zwei reelle Zahlen c_1 und c_2 gibt, sodass beim Einsetzen in die drei Gleichungen alles aufgeht, dann sind die Vektoren linear abhängig, denn in diesem Fall kann ich einen Vektor tatsächlich als Linearkombination der anderen Vektoren darstellen. Falls es aber keine Lösung gibt, sind sie linear unabhängig, denn in diesem Fall kann man den einen Vektor eben nicht als Linearkombination der anderen Vektoren darstellen. Lineare Abhängigkeit hat also offenbar etwas mit der Lösbarkeit linearer Gleichungssysteme zu tun, und das ist einer der Gründe, warum ich mich solchen Gleichungssystemen im dritten Abschnitt widmen werde. Beim momentanen Stand der Dinge ist das vorliegende Problem noch nicht lösbar, Sie müssen sich noch bis zu den linearen Gleichungssystemen gedulden und mir die lineare Unabhängigkeit der Vektoren fürs Erste glauben. ■

Lineare Abhängigkeit oder Unabhängigkeit kann man auch noch etwas anders aufschreiben. Ich komme noch einmal zurück auf die Vektoren im Beispiel 5.10 b), wo ich mit der Gleichung

$$\begin{pmatrix} 4 \\ 7 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

die lineare Abhängigkeit der drei beteiligten Vektoren gezeigt habe. Bringt man in dieser Gleichung alles auf eine Seite, so ergibt sich

$$2 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 4 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Es gibt somit eine Möglichkeit, den so genannten Nullvektor, der in jeder Komponente den Wert Null hat, aus den gegebenen drei Vektoren zu kombinieren, ohne dass sich dabei jeder Vektor den Vorfaktor Null gefallen lassen muss. Das ist keineswegs selbstverständlich; versuchen Sie es mal, aus den beiden Vektoren $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$

den Nullvektor zusammen zu kombinieren. Dazu müssten Sie eine Linearkombination der beiden Vektoren finden, die den Nullvektor ergibt, also:

$$c_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Wenn Sie nun aber die Linearkombination auf der linken Seite ausrechnen, dann finden Sie als Resultat den Vektor $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$, und das heißt, dass $\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ und damit auch $c_1 = c_2 = 0$ gilt.

Manche Vektoren kann man also zum Nullvektor kombinieren, ohne dass man schlicht alle Vektoren mit der Zahl Null multiplizieren muss, bei anderen geht das nicht. Und Sie haben schon gesehen, welche Vektoren die gutwilligen sind: Das waren genau die linear abhängigen. Auf diese Weise kommt man zur folgenden Beschreibung von linearer Abhängigkeit und Unabhängigkeit:

Lineare Abhängigkeit und Unabhängigkeit

Die Vektoren $\mathbf{x}_1, ..., \mathbf{x}_m \in \mathbb{R}^n$ sind genau dann linear unabhängig, wenn aus

$$c_1 \cdot \mathbf{x}_1 + c_2 \cdot \mathbf{x}_2 + \cdots + c_m \cdot \mathbf{x}_m = \mathbf{0}$$

stets folgt:

$$c_1 = c_2 = \cdots = c_m = 0,$$

wobei unter $\mathbf{0}$ der Nullvektor verstanden wird. Kann man die Vektoren dagegen zum Nullvektor kombinieren, ohne dass jeder der Vorfaktoren $c_1, c_2, ..., c_m$ zu Null werden muss, so sind die Vektoren linear abhängig.

Sehen wir uns dazu wieder zwei Beispiele an:

Beispiel 5.11:

a) Die Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sind linear unabhängig. Aus

$$c_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

folgt nämlich sofort

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \text{ und damit } c_1 = 0, c_2 = 0, c_3 = 0.$$

b) Das entsprechende Gleichungssystem für die Vektoren

$$\begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 6 \end{pmatrix}$$

lautet

$$\begin{array}{ccccccc} 5c_1 & + & c_2 & + & c_3 & = & 0 \\ c_1 & + & 4c_2 & + & 2c_3 & = & 0 \\ 3c_1 & + & 2c_2 & + & 6c_3 & = & 0. \end{array}$$

Dieses Gleichungssystem hat natürlich die Lösung $c_1 = c_2 = c_3 = 0$, denn wenn Sie für c_1, c_2 und c_3 jeweils den Wert Null einsetzen, dann ergibt sich insgesamt als Ergebnis jeder linken Seite wieder Null. Die Frage ist nur, ob das Gleichungssystem auch noch andere Lösungen hat, weil Sie im Falle weiterer Lösungen, die nicht nur aus Nullen bestehen, die lineare Abhängigkeit nachgewiesen hätten. Wie man die verschiedenen Lösungen eines linearen Gleichungssystems berechnet, werde ich Ihnen im dritten Abschnitt zeigen, und dort komme ich auch auf dieses Beispiel zurück. ■

Zum Abschluss dieses Abschnitts verrate ich Ihnen noch einen kleinen Trick, mit dem man sich manchmal unerfreuliche Rechnungen ersparen kann. Hat man nämlich mindestens vier dreidimensionale Vektoren in der Hand, dann sind diese vier Vektoren automatisch linear abhängig, ohne dass man noch etwas rechnen müsste. Und so ist das immer: Wenn die Anzahl der infrage stehenden Vektoren die Dimension dieser Vektoren übersteigt, dann sind die Vektoren ganz von alleine linear abhängig, und keiner kann von Ihnen verlangen, dass Sie auch nur einen Finger zum Rechnen rühren.

Lineare Abhängigkeit

Hat man m n-dimensionale Vektoren und gilt $m > n$, so sind diese Vektoren linear abhängig.

Also sind zum Beispiel $\begin{pmatrix} 1 \\ 17 \end{pmatrix}, \begin{pmatrix} -12 \\ 38 \end{pmatrix}, \begin{pmatrix} 123 \\ -125 \end{pmatrix}$ auf jeden Fall linear abhängig, denn ich habe hier drei Vektoren der Dimension zwei.

Mit diesem praktischen Trick zur Rechenvermeidung beende ich den Abschnitt über Vektoren. Im nächsten Abschnitt zeige ich Ihnen, wie man mit Matrizen umgeht und was sie mit Vektoren zu tun haben – aber nicht, ohne Ihnen noch einmal Gelegenheit zum Üben zu geben.

Übungsaufgabe 5.3:
Testen Sie die folgenden Vektoren auf lineare Abhängigkeit bzw. Unabhängigkeit.

a) $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix}$

b) $\begin{pmatrix} -2 \\ 5 \end{pmatrix}, \begin{pmatrix} 12 \\ 17 \end{pmatrix}, \begin{pmatrix} 234 \\ 1 \end{pmatrix}$

c) $\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \\ -4 \end{pmatrix}$ ■

5.2 Matrizen

Matrizen sind nichts Geheimnisvolles, auch wenn es in den Filmen der „Matrix“-Reihe so aussieht. Hätten die Hauptfiguren dort sich etwas mit Linearer Algebra befasst, dann könnten sie ihre wiederholt auftretende Frage „Was ist die Matrix?“ problemlos mit dem Satz beantworten, dass jede Matrix eigentlich nur ein zu breit geratener Vektor ist. Ich gebe zu: Daraus hätte man keinen Film machen können, und vielleicht wurde das Problem deshalb in den Filmen ein wenig anders gelöst.

Nun sind Sie hier aber nicht im Kino und deshalb sehen wir uns wie üblich ein einführendes Beispiel an.

Beispiel 5.12:
Ihr mittlerweile wohlbekanntes Unternehmen *MeineFirma* hat eine sehr erfreuliche Entwicklung genommen und ist inzwischen so groß, dass Sie es in drei Abteilungen A, B und C untergliedern mussten. Diese Abteilungen sind selbstverständlich keine isolierten Inseln im Ozean der Firma, sondern geben untereinander Leistungen ab; was eine Abteilung abgibt, das nimmt die andere auf. Die Leistungsbeziehungen werden durch die folgende Vereinbarung beschrieben:

„Abteilung A gibt 20 Einheiten an Abteilung B und 30 Einheiten an Abteilung C. Abteilung B gibt 30 Einheiten an Abteilung A und 40 Einheiten an Abteilung C. Abteilung C gibt 10 Einheiten an Abteilung A und 20 Einheiten an Abteilung B.“

Was auch immer Ihre Firma herstellen mag und wie auch immer die Einheiten beschaffen sein mögen, in jedem Fall wird man die vorkommenden Leistungsströme in ein übersichtliches Schema packen müssen, nicht nur, damit man es leichter lesen kann, sondern auch und vor allem, damit die Leistungsverrechnung maschinell durchgeführt werden kann. Ich brauche also eine vernünftige Datenstruktur, um die gegebene Situation anständig darstellen zu können. Und hier kommen die Matrizen ins Spiel, unter denen man nichts anderes versteht als eine Art tabellarischer Auflistung von Zahlen. In unserem Beispiel kann ich die folgende Matrix verwenden:

$$\begin{array}{cc} & \begin{array}{ccc} \mathrm{A} & \mathrm{B} & \mathrm{C} \end{array} \\ \begin{array}{c} \mathrm{A} \\ \mathrm{B} \\ \mathrm{C} \end{array} & \begin{pmatrix} 0 & 20 & 30 \\ 30 & 0 & 40 \\ 10 & 20 & 0 \end{pmatrix} \end{array}$$

Ich habe also nur die vorkommenden Zahlen in ein rechteckiges Schema geschrieben und dieses Schema auch noch ein wenig beschriftet, damit der Leser weiß, worum es bei diesen Zahlen überhaupt geht. Die Beschriftung am Rand ist aber nur eine Bequemlichkeitshilfe für den Leser, die Matrix selbst besteht nur aus dem mit Zahlen gefüllten rechteckigen Schema. ■

Eine Matrix ist also nur ein rechteckiges Schema, in dem Zahlen stehen, und weiter ist sie gar nichts. Dieses Ergebnis werde ich jetzt erst einmal für den speziellen Typ, der in Beispiel 5.12 vorkam, sichern.

Matrix

Eine Matrix mit drei Zeilen und drei Spalten ist ein rechteckiges Schema der Form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

wobei die Einträge in diesem Schema reelle Zahlen sind.

Beachten Sie, wie ich die Einträge in der Matrix definiert habe. Bei drei Zeilen und drei Spalten brauche ich neun Einträge, und es hätte auf den ersten Blick nahe gelegen, sie genau wie bei den Vektoren als $a_1, a_2, ..., a_9$ zu bezeichnen. Dabei würde aber einiges an Information verloren gehen, denn auf Anhieb weiß ich so nicht, an welcher Stelle in der Matrix beispielsweise das Element a_5 steht. Besser ist es, eine doppelte Nummerierung zu verwenden, und zum Beispiel den dritten Eintrag der zweiten Zeile als a_{23} zu bezeichnen, was man „a zwei drei“ ausspricht und nicht etwa „a dreiundzwanzig“. Die erste Zahl gibt also die Nummer der aktuellen Zeile an und die zweite die Nummer der aktuellen Spalte, sodass ich jedem Element sofort ansehen kann, an welchem Platz der Matrix es sich zu Hause fühlt.

Natürlich gibt es nicht nur Matrizen mit drei Zeilen und drei Spalten:

Matrix

Sind m und n natürliche Zahlen, so heißt das rechteckige Zahlenschema aus m Zeilen und n Spalten

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

eine $m \times n$-Matrix. Die Menge aller $m \times n$-Matrizen wird mit $\mathbb{R}^{m \times n}$ bezeichnet.

Das ist nun genau das Gleiche wie im Falle der speziellen Matrix, nur dass hier die Zeilenzahl und die Spaltenzahl unbestimmt sind. Sie sehen übrigens, dass eine Matrix keineswegs immer quadratisch sein muss, die Anzahl der Zeilen und die Anzahl der Spalten sind voneinander unabhängig. Das sieht man ganz besonders deutlich, wenn man Matrizen mit nur einer Spalte oder nur einer Zeile betrachtet: Eine Matrix mit nur einer Spalte ähnelt verdächtig einem Vektor, den man genauer als Spaltenvektor bezeichnen kann, und eine Matrix mit nur einer Zeile ist eben auch ein Vektor, der keine Kraft mehr zum Stehen hatte; so etwas bezeichnet man als Zeilenvektor.

Genau wie im ersten Abschnitt haben wir jetzt also Objekte definiert und müssen sehen, was man damit anstellen kann. Die einfachste Operation ist die Addition, die genauso durchgeführt wird, wie Sie sich das wahrscheinlich vorstellen.

Beispiel 5.13:
Ihr Betrieb *MeineFirma* produziert wieder einmal drei Güter, die ich jetzt mit I, II, III bezeichne, und liefert sie an die Händler A, B, C, D aus. Im ersten bzw. im zweiten Quartal werden die jeweiligen Lieferungen durch die folgenden Matrizen beschrieben:

$$\begin{array}{c c} & \begin{array}{cccc} \text{A} & \text{B} & \text{C} & \text{D} \end{array} \\ \begin{array}{c} \text{I} \\ \text{II} \\ \text{III} \end{array} & \begin{pmatrix} 12 & 8 & 0 & 7 \\ 17 & 9 & 4 & 8 \\ 5 & 7 & 4 & 10 \end{pmatrix} \end{array}$$

für das erste Quartal sowie

$$\begin{array}{c c} & \begin{array}{cccc} \text{A} & \text{B} & \text{C} & \text{D} \end{array} \\ \begin{array}{c} \text{I} \\ \text{II} \\ \text{III} \end{array} & \begin{pmatrix} 13 & 7 & 2 & 9 \\ 15 & 10 & 6 & 10 \\ 6 & 6 & 5 & 9 \end{pmatrix} \end{array}$$

für das zweite Quartal. So hat also beispielsweise Händler A im ersten Quartal fünf Einheiten des Produkts III erhalten, die sich aber anscheinend so gut verkauft haben, dass er im zweiten Quartal immerhin eine Einheit mehr bestellt hat. Um nun festzustellen, welcher Händler wie viele Einheiten jedes Produkts im gesamten ersten Halbjahr erhalten hat, werden die beiden Matrizen komponentenweise addiert. Man erhält damit für das erste Halbjahr die Matrix:

$$\begin{array}{c c} & \begin{array}{cccc} \mathrm{A} & \mathrm{B} & \mathrm{C} & \mathrm{D} \end{array} \\ \begin{array}{c} \mathrm{I} \\ \mathrm{II} \\ \mathrm{III} \end{array} & \begin{pmatrix} 25 & 15 & 2 & 16 \\ 32 & 19 & 10 & 18 \\ 11 & 13 & 9 & 19 \end{pmatrix} \end{array}$$

und weiß damit beispielsweise, dass Händler D im ersten Halbjahr insgesamt 18 Einheiten des Produkts II erhalten hat. ■

Matrizen werden also ganz einfach komponentenweise addiert.

Addition und Subtraktion von Matrizen

Es seien

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

und

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

Matrizen mit m Zeilen und n Spalten. Dann setzt man:

$$A \pm B = \begin{pmatrix} a_{11} \pm b_{11} & a_{12} \pm b_{12} & \cdots & a_{1n} \pm b_{1n} \\ a_{21} \pm b_{21} & a_{22} \pm b_{22} & \cdots & a_{2n} \pm b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} \pm b_{m1} & a_{m2} \pm b_{m2} & \cdots & a_{mn} \pm b_{mn} \end{pmatrix}.$$

Die Summe und die Differenz der beiden Matrizen sind dann wieder Matrizen mit m Zeilen und n Spalten.

Matrizen addiert und subtrahiert man also genauso wie Vektoren, nämlich komponentenweise. Wenn Sie also zwei Matrizen mit je drei Zeilen und vier Spalten addieren müssen, dann hat offenbar jede der beiden Matrizen zwölf Einträge und Sie werden deshalb zwölf Additionen durchführen. Da man aber nur Komponenten zusammenzählen kann, die auch vorhanden sind, ist es klar, dass Sie niemals Matrizen addieren oder subtrahieren dürfen, die in ihrer Zeilen- oder Spaltenanzahl nicht zusammenpassen. Matrizen können nur addiert und subtrahiert werden, wenn ihre Zeilen- und Spaltenanzahlen zueinander passen.

Auch dazu wieder Beispiele:

Beispiel 5.14:
a)

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

b) Die Addition

$$\begin{pmatrix} 5 & 3 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

kann nicht durchgeführt werden, da die erste Matrix zwei Spalten hat, die zweite aber drei Spalten. ■

Die Addition von Matrizen ist also nichts Aufregendes, und auch die Subtraktion verspricht keinen gesteigerten Nervenkitzel. Für den Augenblick kann ich auch nicht viel mehr Aufregung versprechen, denn als Nächstes werde ich Ihnen zeigen, wie man eine Matrix mit einem Skalar, also mit einer Zahl, multipliziert. Da Sie sich sicher noch an die entsprechende Operation für Vektoren erinnern können, wird Sie hier nichts überraschen.

Beispiel 5.15:
Man muss auch mal über den Tellerrand hinausschauen und sich um andere Länder und andere Sitten bemühen. Ich verlasse also Ihr gutes altes Unternehmen *Meine-Firma* und werfe einen Blick auf die internationalen Handelsbeziehungen. Damit es nicht zu unübersichtlich wird, beschränke ich mich auf drei Länder A, B und C. Die Außenhandelsbeziehungen zwischen diesen Ländern sollen durch die folgende Matrix beschrieben werden:

$$\begin{array}{c} \\ \text{A} \\ \text{B} \\ \text{C} \end{array} \begin{array}{c} \begin{array}{ccc} \text{A} & \text{B} & \text{C} \end{array} \\ \begin{pmatrix} 0 & 12 & 8 \\ 6 & 0 & 4 \\ 10 & 2 & 0 \end{pmatrix} \end{array},$$

wobei die Zeilen den jeweiligen Export angeben und die Spalten den Import. Die Abrechnung erfolgt dabei in Milliarden Euro, sodass also das Land A Waren im Wert von acht Milliarden Euro nach C exportiert, aber in der Gegenrichtung Waren für zehn Milliarden Euro von C nach A exportiert werden. Soll nun die Abrechnung in US-Dollar vorgenommen werden, so muss man jede Komponente der Matrix mit 1,4 multiplizieren, weil nach dem heutigen Stand (das ist der Stand von Oktober 2010) ein Euro ziemlich genau einem Dollar und vierzig Cent entspricht. Damit ergibt sich die neue Matrix:

$$\begin{pmatrix} 0 & 16{,}8 & 11{,}2 \\ 8{,}4 & 0 & 5{,}6 \\ 14 & 2{,}8 & 0 \end{pmatrix}.$$

Ich habe also die alte Matrix mit der Zahl 1,3 multipliziert, indem ich jeden einzelnen Eintrag der Matrix mit 1,3 multipliziert habe. Damit man auch sieht, dass diese Operation nichts mit A, B oder C zu tun hat, habe ich die Spalten- und Zeilenbezeichnungen weggelassen; der Multiplikation ist das egal. ■

Damit kennen Sie auch schon das Prinzip der Multiplikation Zahl · Matrix, das ich jetzt allgemein formulieren werde.

Multiplikation einer Matrix mit einer Zahl

Für eine Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

mit m Zeilen und n Spalten und eine reelle Zahl λ setzt man:

$$\lambda \cdot A = \begin{pmatrix} \lambda \cdot a_{11} & \lambda \cdot a_{12} & \cdots & \lambda \cdot a_{1n} \\ \lambda \cdot a_{21} & \lambda \cdot a_{22} & \cdots & \lambda \cdot a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda \cdot a_{m1} & \lambda \cdot a_{m2} & \cdots & \lambda \cdot a_{mn} \end{pmatrix}.$$

Am besten rechnen Sie einmal kurz selbst:

Übungsaufgabe 5.4:
Berechnen Sie:

a) $2 \cdot \begin{pmatrix} 1 & 0 & -2 \\ -3 & 5 & 1 \\ 4 & 2 & -1 \end{pmatrix} + 3 \cdot \begin{pmatrix} -6 & 3 & 4 \\ 1 & 2 & -2 \\ -3 & -4 & 0 \end{pmatrix}.$

b) $-3 \cdot \begin{pmatrix} -1 & 5 & 2 \\ -2 & 1 & 0 \\ 3 & 1 & -2 \end{pmatrix} + \begin{pmatrix} 2 & -5 & -2 \\ 2 & -1 & 0 \\ -3 & -1 & 3 \end{pmatrix}.$ ■

Um es noch einmal in Worte zu fassen: Man multipliziert eine Matrix mit einer Zahl, indem man jeden einzelnen Eintrag der Matrix mit dieser Zahl multipliziert. Auch das ist nicht so schwer, dass es einem schlaflose Nächte bereiten muss, aber immerhin sind Matrizen doch recht komplexe Gebilde, und so kann es nicht schaden, einmal kurz die Rechenregeln für die Operationen aufzuschreiben, die ich Ihnen bisher gezeigt habe. Sie werden sie allerdings wenig überraschend finden, denn es stellt sich heraus, dass in Bezug auf die Addition von Matrizen und auf die Multiplikation von Matrizen mit Skalaren alles ganz genauso ist, wie Sie es schon beim üblichen Zahlenrechnen gelernt haben. Man kann also zum Beispiel die Reihenfolge bei der Addition vertauschen, man kann Klammern ausmultiplizieren und Zahlen, die als gemeinsame Faktoren vorkommen, vorklammern.

Rechenregeln für Matrizen

Sind A und B Matrizen mit m Zeilen und n Spalten und λ und μ reelle Zahlen, dann gelten die folgenden Rechenregeln:
a) $A + B = B + A$;
b) $\lambda \cdot (A + B) = \lambda \cdot A + \lambda \cdot B$;
c) $(\lambda + \mu) \cdot A = \lambda \cdot A + \mu \cdot A$.

Wie schon angekündigt, ist nichts davon überraschend, und die Regeln sagen im Grunde genommen nur aus, dass alles so funktioniert wie immer. Trotzdem ist das eine oder andere Beispiel nie verkehrt:

Beispiel 5.16:
Ich berechne den Ausdruck

$$3 \cdot \begin{pmatrix} 2 & 5 \\ -1 & 7 \end{pmatrix} - 2 \cdot \left[\begin{pmatrix} 1 & 3 \\ 4 & 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} \right].$$

Zunächst gehe ich die innere Klammer an und bestimme:

$$\begin{pmatrix} 1 & 3 \\ 4 & 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 4 & 0 \end{pmatrix} + \begin{pmatrix} 10 & 5 \\ 0 & -5 \end{pmatrix} = \begin{pmatrix} 11 & 8 \\ 4 & -5 \end{pmatrix}.$$

Also wird

$$\begin{aligned} & 3 \cdot \begin{pmatrix} 2 & 5 \\ -1 & 7 \end{pmatrix} - 2 \cdot \left[\begin{pmatrix} 1 & 3 \\ 4 & 0 \end{pmatrix} + 5 \cdot \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} \right] \\ = & \begin{pmatrix} 6 & 15 \\ -3 & 21 \end{pmatrix} - 2 \cdot \begin{pmatrix} 11 & 8 \\ 4 & -5 \end{pmatrix} \\ = & \begin{pmatrix} 6 & 15 \\ -3 & 21 \end{pmatrix} - \begin{pmatrix} 22 & 16 \\ 8 & -10 \end{pmatrix} = \begin{pmatrix} -16 & -1 \\ -11 & 31 \end{pmatrix}. \end{aligned}$$

Sie hätten aber mit dem gleichen Recht auch zuerst den Faktor 2 in die eckige Klammer hineinmultiplizieren, dann die einzelnen Vorfaktoren in die Matrizen multiplizieren und schließlich alles addieren können, was ich Ihnen als Übung sehr empfehlen möchte. Herauskommen muss in jedem Fall dasselbe. ■

Es scheint wieder einmal Zeit zu sein für eine Übungsaufgabe.

Übungsaufgabe 5.5:
Berechnen Sie

$$2 \cdot \begin{pmatrix} 1 & 4 \\ -2 & 6 \end{pmatrix} - 3 \cdot \left[\begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix} + 4 \cdot \begin{pmatrix} 1 & 0 \\ -1 & -2 \end{pmatrix} \right].$$

■

Vermutlich werden Sie zugeben, dass bisher alles recht harmlos war. Auch die Matrizenmultiplikation, die ich Ihnen als Nächstes vorstelle, zeichnet sich nicht durch allzu große Kompliziertheit aus, aber immerhin ist sie ein wenig komplizierter als die anderen Operationen – wenn auch nicht sehr. Das Beste wird sein, wir sehen uns wieder mal ein Beispiel an.

Beispiel 5.17:
Wir kehren zurück in Ihr gutes altes Unternehmen *MeineFirma* und nehmen an, dass Ihr Betrieb eine zweistufige Produktion betreibt. Aus den Rohstoffen R_1 und R_2 werden in der ersten Produktionsstufe die Zwischenprodukte Z_1, Z_2 und Z_3 hergestellt, während in der zweiten Produktionsstufe aus diesen Zwischenprodukten die Endprodukte E_1 und E_2 gefertigt werden. Den Materialverbrauch in jeder Produktionsstufe kann man dann mithilfe der beiden folgenden Matrizen beschreiben:

$$\begin{array}{c c} & \begin{array}{ccc} Z_1 & Z_2 & Z_3 \end{array} \\ \begin{array}{c} R_1 \\ R_2 \end{array} & \begin{pmatrix} 1 & 4 & 5 \\ 2 & 3 & 6 \end{pmatrix} \end{array}$$

sowie

$$\begin{array}{c c} & \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{array}{c} Z_1 \\ Z_2 \\ Z_3 \end{array} & \begin{pmatrix} 1 & 3 \\ 5 & 3 \\ 4 & 1 \end{pmatrix} \end{array}.$$

Bevor man mit solchen Matrizen rechnet, muss man sich natürlich erst einmal über ihre Interpretation einigen. Die erste Matrix beschreibt, wie viele Einheiten der beiden Rohstoffe bei der Produktion je einer Einheit der Zwischenprodukte benötigt werden. Um eine Einheit von Z_1 herzustellen, braucht man also genau eine Einheit des Rohstoffs R_1 und zwei Einheiten des Rohstoffs R_2. Dagegen werden zur Produktion von einer Einheit des Zwischenproduktes Z_2 vier Einheiten R_1 und drei Einheiten R_2 gebraucht. Genauso können Sie auch die zweite Matrix lesen: Aus den Zwischenprodukten werden jetzt die Endprodukte gefertigt, und für eine Einheit des Endproduktes E_1 werden Sie den Verbrauch von einer Einheit Z_1, fünf Einheiten Z_2 und vier Einheiten Z_3 verbuchen müssen.

So viel zur Ausgangslage. Da ich am Ende aber wissen muss, wie viele Rohstoffe ich für die weitere Produktion besorgen sollte, bin ich vor allem am jeweiligen Rohstoffverbrauch pro Einheit des Endprodukts interessiert und will deshalb eine Gesamtmaterialverbrauchsmatrix aufstellen, die mir genau diesen Rohstoffverbrauch angibt.

Sobald ich über diese Matrix verfüge, kann ich dann auch Fragen wie „Wie groß ist der Rohstoffbedarf, wenn der Betrieb 1000 Einheiten E_1 und 1200 Einheiten E_2 herstellen soll?“ beantworten.

Aus der zweiten Matrix können Sie ablesen, dass man für eine Einheit E_1 genau $1Z_1, 5Z_2$ und $4Z_3$ braucht. Für eine Einheit Z_1 braucht man $1R_1$ und $2R_2$, für eine Einheit Z_2 braucht man $4R_1$ und $3R_2$, für eine Einheit Z_3 braucht man $5R_1$ und $6R_2$. Folglich benötigt man für eine Einheit E_1 die folgende Menge R_1:

$$1 \cdot 1 + 4 \cdot 5 + 5 \cdot 4 = 41$$

und die folgende Menge R_2:

$$2 \cdot 1 + 3 \cdot 5 + 6 \cdot 4 = 41.$$

Auf diese einfache Weise haben Sie schon den Verbrauch an Rohstoffen bei der Herstellung einer Einheit des Endproduktes E_1 berechnet, und vielleicht erinnert Sie die Berechnungsweise an eine Rechenmethode, die ich Ihnen im Abschnitt über Vektoren gezeigt habe. Diese Kombination aus Multiplizieren und Addieren hatten wir schon beim Skalarprodukt gesehen, und tatsächlich liegen hier die Skalarprodukte

$$\begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 5 \\ 4 \end{pmatrix} \text{ und } \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 5 \\ 4 \end{pmatrix}$$

vor. Sie sehen, dass der Vektor $\begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix}$ genau der ersten Zeile der ersten Matrix entspricht, nur eben als Spalte geschrieben, während der Vektor $\begin{pmatrix} 1 \\ 5 \\ 4 \end{pmatrix}$ nichts anderes ist als die erste Spalte der zweiten Matrix. Bei der zweiten Rechnung haben Sie die gleiche Situation: Der Vektor $\begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}$ entspricht der zweiten Zeile der ersten Matrix, nur eben als Spalte geschrieben, während der Vektor $\begin{pmatrix} 1 \\ 5 \\ 4 \end{pmatrix}$ noch immer die erste Spalte der zweiten Matrix darstellt.

Man erhält also den Gesamtrohstoffverbrauch für eine Einheit E_1, indem man das Skalarprodukt aus dem ersten bzw. zweiten *Zeilenvektor* der ersten Matrix mit dem ersten *Spaltenvektor* der zweiten Matrix berechnet und dabei den Zeilenvektor für einen Moment als Spaltenvektor ansieht, damit man das Skalarprodukt überhaupt bilden kann. Analog findet man für eine Einheit E_2:

$$1 \cdot 3 + 4 \cdot 3 + 5 \cdot 1 = 20$$

und

$$2 \cdot 3 + 3 \cdot 3 + 6 \cdot 1 = 21,$$

woran Sie ablesen können, dass zur Herstellung einer Einheit Ihres Endproduktes E_2 sowohl 20 Einheiten R_1 als auch 21 Einheiten R_2 benötigt werden.

Die Gesamtmaterialverbrauchsmatrix lautet also:

$$\begin{array}{c c} & \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{array}{c} R_1 \\ R_2 \end{array} & \begin{pmatrix} 41 & 20 \\ 41 & 21 \end{pmatrix} \end{array}.$$

Und jetzt sind wir so weit, ein Matrizenprodukt aufschreiben zu können. Man bezeichnet nämlich die eben durchgeführte Verknüpfung der beiden Matrizen als ihr Produkt und schreibt dafür:

$$\begin{pmatrix} 1 & 4 & 5 \\ 2 & 3 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 \\ 5 & 3 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 41 & 20 \\ 41 & 21 \end{pmatrix}.$$

Wie man dieses Produkt ausrechnet, haben Sie eben gesehen: Man kippt die Spalten der zweiten Matrix der Reihe nach auf die Zeilen der ersten Matrix und bildet die Skalarprodukte. Das funktioniert nicht nur bei diesen beiden Matrizen, sondern auch bei allen möglichen anderen, unter einer kleinen Voraussetzung, auf die ich gleich noch zu sprechen komme. Vorher gehen wir noch kurz der Frage nach: „Wie groß ist der Rohstoffbedarf, wenn der Betrieb 1 000 Einheiten E_1 und 1 200 Einheiten E_2 herstellen soll?“ Mithilfe der berechneten Verbrauchsmatrix ist sie jetzt leicht zu beantworten, denn für 1 000 Einheiten E_1 und 1 200 Einheiten E_2 braucht man an R_1-Einheiten:

$$41 \cdot 1\,000 + 20 \cdot 1\,200 = 65\,000$$

und an R_2-Einheiten:

$$41 \cdot 1\,000 + 21 \cdot 1\,200 = 66\,200.$$

Man schreibt dafür:

$$\begin{pmatrix} 41 & 20 \\ 41 & 21 \end{pmatrix} \cdot \begin{pmatrix} 1\,000 \\ 1\,200 \end{pmatrix} = \begin{pmatrix} 65\,000 \\ 66\,200 \end{pmatrix},$$

und auch das funktioniert genauso wie eben, denn Sie müssen zuerst das Skalarprodukt aus der ersten Zeile der Matrix und dem gegebenen Vektor bilden und anschließend auf die gleiche Weise das Skalarprodukt aus der zweiten Zeile der Matrix und dem Vektor. ■

Allgemeine Matrizenprodukte definiert man auf die gleiche Weise, wie ich es Ihnen in Beispiel 5.17 gezeigt habe. Sobald man das in eine allgemeine Formel fasst, bekommt man etwas unhandliche Ausdrücke, die auf Anhieb schwer zu verstehen sind, aber manchmal lässt sich das nicht vermeiden.

Matrizenmultiplikation

Es seien A und B Matrizen, wobei die Zeilen von A genauso lang sein sollen wie die Spalten von B. Das Produkt $C = A \cdot B$ der beiden Matrizen hat dann so viele Zeilen wie A und so viele Spalten wie B und wird folgendermaßen berechnet. Ist

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

eine Matrix mit m Zeilen und n Spalten und

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nk} \end{pmatrix}$$

eine Matrix mit n Zeilen und k Spalten, so sind die Zeilen von A genauso lang wie die Spalten von B. Die Matrix $C = A \cdot B$ hat dann die Form

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1k} \\ c_{21} & c_{22} & \cdots & c_{2k} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mk} \end{pmatrix}.$$

Die einzelnen Einträge c_{ij} in C berechnen sich als Skalarprodukt der i-ten Zeile von A und der j-ten Spalte von B, also:

$$c_{ij} = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix} \cdot \begin{pmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{pmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

Ist $\mathbf{x}$ ein Vektor, der genauso lang ist wie die Zeilen von A, so ist das Produkt $A \cdot \mathbf{x}$ ein Vektor, der genauso lang ist wie die Spalten von A und berechnet wird, indem man der Reihe nach das Skalarprodukt der ersten Zeile von A mit $\mathbf{x}$, das Skalarprodukt der zweiten Zeile von A mit $\mathbf{x}$,..., das Skalarprodukt der letzten Zeile von A mit $\mathbf{x}$ in einen Vektor schreibt.

Vielleicht klingt das jetzt noch ein wenig abstrakt und ungewohnt, aber das wird sich nach den folgenden Beispielen gleich geben.

Beispiel 5.18:
a) Ich betrachte die beiden Matrizen

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, B = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix}.$$

Beide Matrizen haben zwei Zeilen und zwei Spalten, also sind die Zeilen von A sicher so lang wie die Spalten von B und ich kann mich ans Multiplizieren machen. Die Produktmatrix $C = A \cdot B$ muss ebenfalls aus zwei Zeilen und zwei Spalten bestehen, und mit $C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$ liefert die Rechenvorschrift:

$$c_{11} = 1 \cdot 2 + 1 \cdot 0 = 2, \quad c_{12} = 1 \cdot (-1) + 1 \cdot 1 = 0$$

und

$$c_{21} = 1 \cdot 2 + (-1) \cdot 0 = 2, \quad c_{22} = 1 \cdot (-1) + (-1) \cdot 1 = -2.$$

Folglich ist

$$A \cdot B = \begin{pmatrix} 2 & 0 \\ 2 & -2 \end{pmatrix}.$$

b) Nun geht es um die Matrizen

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

Die Zeilenlänge von A ist drei, und auch die Spalten von B haben die Länge drei, sodass einer Multiplikation nichts im Wege steht. Die Produktmatrix $C = A \cdot B$ hat dann so viele Zeilen wie A, also zwei, und so viele Spalten wie B, also drei. Jetzt muss ich nur noch streng nach Vorschrift die richtigen Skalarprodukte an die richtige Stelle schreiben und erhalte:

$$\begin{aligned} &A \cdot B \\ &= \begin{pmatrix} 1 \cdot 1 + 4 \cdot (-2) + 2 \cdot 0 & 1 \cdot 1 + 4 \cdot 3 + 2 \cdot 1 & 1 \cdot 0 + 4 \cdot 5 + 2 \cdot 4 \\ 4 \cdot 1 + 0 \cdot (-2) + (-3) \cdot 0 & 4 \cdot 1 + 0 \cdot 3 + (-3) \cdot 1 & 4 \cdot 0 + 0 \cdot 5 + (-3) \cdot 4 \end{pmatrix} \\ &= \begin{pmatrix} -7 & 15 & 28 \\ 4 & 1 & -12 \end{pmatrix}. \end{aligned}$$

Dagegen kann man $B \cdot A$ überhaupt nicht berechnen, denn jetzt hat die Matrix, die vorne steht (also B), eine Zeilenlänge von drei, während die hintere Matrix (und das ist A) mit der Spaltenlänge zwei geplagt ist – das passt nicht zusammen, und daher gibt es dieses Matrizenprodukt nicht. ■

Ich habe schon mehrfach erwähnt, dass man nicht alle Matrizen miteinander multiplizieren kann: Damit $A \cdot B$ ein sinnvoller Ausdruck ist, müssen die Zeilen von A so lang sein wie die Spalten von B. Nun ist aber die Zeilenlänge von A gerade die Anzahl der Spalten von A, und auf der anderen Seite ist die Spaltenlänge von B gleich der Anzahl seiner Zeilen. Man kann also genau dann $A \cdot B$ ausrechnen, wenn die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B ist. Ich wollte deshalb auf diese Formulierung hinaus, weil man sie leicht in eine Formel übersetzen kann. Die Menge aller Matrizen mit m Zeilen und n Spalten haben wir nämlich am Anfang des Abschnitts mit $\mathbb{R}^{m \times n}$ abgekürzt, und wenn A in der Menge $\mathbb{R}^{m \times n}$ ist, dann muss B in der Menge $\mathbb{R}^{n \times k}$ sein, also n Zeilen und wer weiß wie viele Spalten haben. In diesem Fall hat A genau n Spalten und B hat n Zeilen, weshalb alles zusammen passt. Und Sie wissen auch schon, wie viele Zeilen und Spalten $A \cdot B$ hat, denn das Produkt hat so viele Zeilen wie A und so viele Spalten wie B. Also hat $A \cdot B$ genau m Spalten und k Zeilen.

Bedingung für die Matrizenmultiplikation

Genau dann kann man das Matrizenprodukt $A \cdot B$ bilden, wenn $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{n \times k}$ gilt. In diesem Fall hat das Produkt $A \cdot B$ m Zeilen und k Spalten, also ist $A \cdot B \in \mathbb{R}^{m \times k}$.

Es könnte sein, dass Ihnen das Hantieren mit dem Matrizenprodukt noch etwas schwer fällt; das kann einerseits an der Neuheit der Rechenmethode liegen, andererseits aber auch daran, dass wir noch kein wirklich übersichtliches Mittel in der Hand haben, mit dem man Matrizenmultiplikationen durchführen kann. Das muss aber nicht so bleiben, denn es gibt ein einfaches und gut überschaubares Schema, das die Berechnung von Matrizenprodukten zur bloßen Routine werden lässt: das so genannte Falk'sche Schema, das ich Ihnen im nächsten Beispiel zeige.

Beispiel 5.19:
Ich will das Produkt aus

$$A = \begin{pmatrix} 1 & 0 \\ -2 & 1 \\ 0 & 3 \end{pmatrix} \text{ und } B = \begin{pmatrix} 2 & 1 & 9 \\ 0 & -1 & -2 \end{pmatrix}$$

berechnen. Das Falk'sche Schema sieht dann folgendermaßen aus:

$$\begin{array}{cc|ccc} & & 2 & 1 & 9 \\ & & 0 & -1 & -2 \\ \hline 1 & 0 & 2 & 1 & 9 \\ -2 & 1 & -4 & -3 & -20 \\ 0 & 3 & 0 & -3 & -6 \end{array}$$

Man schreibt die beiden Matrizen A und B in dieser Anordnung auf: A steht links unten, B steht rechts oben. In den Raum, den A und B einschließen, passt genau eine 3×3-Matrix, denn da A eine 3×2-Matrix und B eine 2×3-Matrix ist, muss das Produkt $A \cdot B$ eine 3×3-Matrix sein. Die Einträge der Produktmatrix kann man dann leicht bestimmen. Jede Position in der Ergebnismatrix bekommt man, indem man eine A-Zeile nach rechts und eine B-Spalte nach unten fortführt, denn diese beiden Linien müssen sich genau innerhalb des Platzes für die Ergebnismatrix schneiden. Bei der ersten Zeile von A und der ersten Spalte von B schneiden sich die fortgeführten Linien genau an der Stelle, an der die Zahl 2 eingetragen ist. Diese Zahl 2 entsteht, indem man den A-Zeilenvektor $(1, 0)$ mit dem B-Spaltenvektor $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ komponentenweise multipliziert und anschließend die Ergebnisse addiert, also das an dieser Stelle verlangte Skalarprodukt berechnet. Nach dem gleichen Prinzip geht es weiter. Der nächste Eintrag in der ersten Zeile der Produktmatrix entsteht, indem man auf die angegebene Weise die erste A-Zeile $(1, 0)$ mit der zweiten B-Spalte $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ kombiniert und so weiter, bis alle Einträge gefüllt sind. ■

Um das Falk'sche Schema zur Berechnung von $A \cdot B$ anzuwenden, müssen Sie also nur A nach links unten und B nach rechts oben schreiben und dann den entstandenen leeren Platz mit den passenden Skalarprodukten füllen. In den folgenden Übungen haben Sie dazu Gelegenheit.

Übungsaufgabe 5.6:
Berechnen Sie:

a) $\begin{pmatrix} 1 & 0 & -2 \\ -3 & 5 & 1 \\ 4 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} -6 & 3 & 4 \\ 1 & 2 & -2 \\ -3 & -4 & 0 \end{pmatrix}$.

b) $\begin{pmatrix} -1 & 5 & 2 \\ -2 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & -5 & -2 & 1 \\ 2 & -1 & 0 & 2 \\ -3 & -1 & 3 & 0 \end{pmatrix}$. ■

Wieder einmal sind wir so weit, dass eine neue Operation zur Verfügung steht und ich Ihnen berichten muss, wie sie sich mit den anderen Operationen verträgt. Kurz gesagt: Es ist Zeit für ein paar Rechenregeln. Sie bieten wenig Überraschendes, aber es gibt immerhin einen Punkt, der ziemlich aus dem Rahmen fällt. Zunächst einmal die Regeln:

Rechenregeln für das Matrizenprodukt

Gegeben seien Matrizen A, B und C, für die die folgenden Operationen durchführbar sein sollen. Dann gelten die folgenden Regeln:

a)
$$A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$

b)
$$A \cdot (B + C) = A \cdot B + A \cdot C.$$

c)
$$(A + B) \cdot C = A \cdot C + B \cdot C.$$

d) Im Allgemeinen ist $A \cdot B \neq B \cdot A$, sofern man überhaupt beide Produkte berechnen kann.

Gehen wir die Sache einmal kurz durch. Der Eingangssatz dieser Regeln ist wichtig, denn ich habe in den Regeln selbst darauf verzichtet, immer genau aufzuschreiben, welchen Typ die einzelnen Matrizen haben sollen, damit man sie auch addieren oder gar multiplizieren kann. Ich gehe also durchgängig davon aus, dass die Zeilen- und Spaltenanzahlen so gewählt sind, dass alle Operationen gut gehen. Die erste Regel sagt dann nur aus, dass man, genau wie beim Zahlenrechnen auch, auf Klammern verzichten kann, sofern man es nur mit Multiplikationen zu tun hat, denn es kommt bei jeder Art der Klammersetzung ohnehin das Gleiche heraus. Die zweite Regel formuliert das so genannte Distributivgesetz, was nur ein hochtrabender Name für die schlichte Tatsache ist, dass Sie auch in Bezug auf das Matrizenprodukt jederzeit ausmultiplizieren und vorklammern dürfen: Die Matrizenmultiplikation verträgt sich also mit der Matrizenaddition. Aber warum wird dann in der dritten Regel noch einmal ausmultipliziert bzw. vorgeklammert? In c) steht doch eigentlich das Gleiche wie in b)! Sieht so aus, ist aber nicht so. Ein genauerer Blick zeigt, dass in b) die einzelne Matrix links steht und die Summe rechts, während es in c) umgekehrt ist. Und das macht tatsächlich einen Unterschied, denn nach Regel d) kann man die Reihenfolge beim Multiplizieren nicht so einfach vertauschen, weshalb ich die Regel über das Ausmultiplizieren in zwei Variationen aufschreiben musste.

Das eigentlich Überraschende ist also Regel d): $A \cdot B$ und $B \cdot A$ sind – sofern überhaupt beide existieren – zwei verschiedene Matrizen. Dass man oft genug gar nicht beide Produkte bilden kann, ist leicht einzusehen. Hat A beispielsweise zwei Zeilen und drei Spalten, B dagegen drei Zeilen und vier Spalten, so passen A und B wunderbar zusammen, denn A hat so viele Spalten wie B Zeilen hat. Umgekehrt geht aber gar nichts, weil B vier Spalten hat, A dagegen nur zwei Zeilen, das Produkt $B \cdot A$ also nicht möglich ist. Aber auch wenn Sie beide Produkte ausrechnen können, kommt normalerweise nicht dasselbe heraus.

Beispiel 5.20:

a) Für die Matrizen

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \text{ und } B = \begin{pmatrix} 0 & 3 \\ 1 & -1 \end{pmatrix}$$

will ich die Produkte $A \cdot B$ und $B \cdot A$ ausrechnen. Mit dem Falk'schen Schema können Sie feststellen, dass

$$A \cdot B = \begin{pmatrix} 2 & 1 \\ 1 & -7 \end{pmatrix} \text{ und } B \cdot A = \begin{pmatrix} -6 & 3 \\ 3 & 1 \end{pmatrix}$$

gilt. Folglich kann man zwar beide Produkte bilden, aber man erhält zwei verschiedene Ergebnisse.

b) Sehen wir uns einmal an, wie sich das Distributivgesetz auswirkt. Zu diesem Zweck berechne ich das Produkt

$$\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \cdot \left[\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix} \right]$$

auf zwei Arten. Eine Möglichkeit besteht darin, erst die beiden Matrizen in den eckigen Klammern zu addieren und anschließend die Multiplikation durchzuführen. Die Addition ergibt:

$$\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 3 \\ 1 & -1 \end{pmatrix}$$

und somit reduziert sich das Ganze auf die Multiplikation

$$\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 3 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & -7 \end{pmatrix},$$

denn das hatte ich gerade in Beispiel a) ausgerechnet. Nun sagt mir das Distributivgesetz aus Regel b), dass ich auch anders vorgehen und die Klammer ausmultiplizieren kann. Das ergibt dann:

$$\begin{aligned} &\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 \\ -2 & -3 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & -7 \end{pmatrix}. \end{aligned}$$ ■

Auch ein wenig eigenes Rechnen kann nicht schaden:

Übungsaufgabe 5.7:

Berechnen Sie auf zwei Arten

$$\begin{pmatrix} 2 & 3 & -1 \\ -2 & 4 & 1 \\ 0 & 5 & -2 \end{pmatrix} \cdot \left[2 \cdot \begin{pmatrix} 1 & 2 & -2 \\ -1 & 3 & 1 \\ 4 & 0 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 4 & -1 \\ -2 & 3 & 1 \\ 1 & 5 & -2 \end{pmatrix} \right].$$ ■

Wie man Matrizen multipliziert, wissen Sie jetzt, aber das wirft sofort eine weitere Frage auf: Wenn man schon Matrizen miteinander multiplizieren kann und dabei wieder eine Matrix herauskommt, könnte es dann nicht auch so etwas wie eine Matrizendivision geben? Das klingt zunächst ein wenig theoretisch, hat aber einen sehr konkreten praktischen Hintergrund. Erinnern Sie sich für einen Moment an den zweistufigen Produktionsprozess, der uns zum Matrizenprodukt geführt hat. Ein ähnliches Beispiel zeigt auch, warum die Umkehrung sinnvoll sein kann. Nehmen Sie also an, in Ihrem vielzitierten Unternehmen *MeineFirma* wird ein zweistufiger Produktionsprozess mit zwei Rohstoffen R_1, R_2, zwei Zwischenprodukten Z_1, Z_2 und zwei Endprodukten E_1, E_2 durchgeführt. Durch Messungen oder eine brauchbare Buchhaltung kennt man die *Verbrauchsmatrix* B beim Übergang von Z nach E und die Gesamtverbrauchsmatrix C, die beschreibt, wie viele Rohstoffeinheiten R_1, R_2 für die Produktion einer Einheit E_1 bzw. E_2 verbraucht werden. Bei beiden Matrizen handelt es sich natürlich um Matrizen mit zwei Zeilen und zwei Spalten, da Sie sowohl mit zwei Rohstoffen als auch mit zwei Zwischenprodukten und zwei Endprodukten hantieren. Das Problem ist: Wie kann man daraus die Verbrauchsmatrix A beim Übergang von Rohstoffen zu Zwischenprodukten bestimmen?

In Beispiel 5.17 hatten wir festgestellt, dass $A \cdot B = C$ gilt. Jetzt kenne ich sowohl B als auch C und kann deshalb leider A nicht mehr durch eine simple Multiplikation berechnen. Würde es sich um Zahlen handeln, dann wäre die Sache einfach: Die Gleichung $a \cdot b = c$ würde ich, sofern $b \neq 0$ ist, locker durch b teilen und bekäme $a = \frac{c}{b}$ oder auch $a = c \cdot \frac{1}{b}$. So etwas bräuchte ich auch für Matrizen und schon wäre mein Problem gelöst. Ich muss mir also eine Art „Kehrmatrix“ überlegen (das hat nichts mit der Kehrwoche im Schwäbischen zu tun, sondern mit dem Kehrwert), die die Rolle von $\frac{1}{b}$ spielt. Da es sich um Matrizen handelt, schreibt man dann nicht $\frac{1}{B}$, sondern etwas vornehmer B^{-1}, meint damit im Grunde genommen aber genau dasselbe, denn am Ende soll die Gleichung $A = C \cdot B^{-1}$ gelten. Auf diese Weise können Sie dann die gesuchte Matrix A ausrechnen.

Alles schön und gut, aber wie kommt man jetzt an die Kehrmatrix, die man übrigens gar nicht Kehrmatrix, sondern in Wahrheit inverse Matrix nennt? Sie soll die Rolle von $\frac{1}{b}$ spielen, und deshalb muss ich erst einmal herausfinden, was im Reich der Matrizen eine Eins darstellt. Bei den Zahlen ist die Eins die einzige Zahl, mit der man ohne Angst vor den Folgen multiplizieren kann, denn $1 \cdot a = a \cdot 1 = a$, ganz egal, welches $a \in \mathbb{R}$ Sie auch auswählen mögen. Ich brauche also eine Matrix, die beim Multiplizieren keine Wirkung zeigt, und die ist leicht zu finden. So ist zum Beispiel

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

wie Sie unschwer nachrechnen können, und für Matrizen mit drei Zeilen und drei Spalten hat die Matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ die gleiche wundersame Wirkung. Das ist kein Zufall, sondern ein Hinweis darauf, wie die „Einsermatrix“ aussehen muss.

Einheitsmatrix

Ist A eine Matrix mit n Zeilen und n Spalten, dann gilt mit

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

immer:

$$A \cdot I_n = I_n \cdot A = A.$$

Daher heißt die Matrix I_n **Einheitsmatrix**.

Die Einheitsmatrix hat also n Zeilen und n Spalten und zeichnet sich dadurch aus, dass ihr Produkt mit jeder beliebigen Matrix aus $\mathbb{R}^{n \times n}$ wieder die Matrix selbst ergibt. Ihr Aufbau ist einfach genug: In der von links oben nach rechts unten verlaufenden Diagonale – der so genannten Hauptdiagonale – stehen Einsen, überall sonst stehen Nullen.

Jetzt habe ich die Matrixeins in der Hand, und somit steht einer Definition der Kehrmatrix, genauer gesagt der inversen Matrix, nichts mehr im Wege. Erinnern Sie sich an unseren Startpunkt: Es geht mir hier darum, zu einer gegebenen Matrix eine inverse Matrix zu finden, die die Rolle des Kehrwertes spielen soll. Nun ist aber für $a \neq 0$ der Kehrwert $\frac{1}{a}$ dadurch gekennzeichnet, dass $a \cdot \frac{1}{a} = 1$ gilt. Nichts übermäßig Neues, aber sehr hilfreich, denn das kann man nun direkt auf die Matrizen übertragen. Aus der Zahl a wird die Matrix A, aus dem Kehrwert $\frac{1}{a}$ wird die inverse Matrix A^{-1} und aus der Zahl 1 wird die Einheitsmatrix I_n.

Inverse Matrix

Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt *invertierbar* oder auch *regulär*, wenn es eine Matrix $A^{-1} \in \mathbb{R}^{n \times n}$ gibt mit

$$A \cdot A^{-1} = A^{-1} \cdot A = I_n.$$

In diesem Fall heißt A^{-1} die inverse Matrix von A.

Beachten Sie, dass ich hier nur noch so genannte quadratische Matrizen betrachte, also Matrizen, die genauso lang wie breit sind; nur in dieser Situation kann man wirklich von Invertierbarkeit sprechen. Die inverse Matrix von A ist also die Matrix, die beim Multiplizieren mit A die Einheitsmatrix ergibt, wobei es hier (und nur hier!) keine Rolle spielt, ob Sie $A \cdot A^{-1}$ oder $A^{-1} \cdot A$ berechnen – es muss immer die Einheitsmatrix herauskommen.

Kommen wir noch einmal zurück auf mein Eingangsbeispiel, bei dem ich die Matrix A aus B und C rekonstruieren wollte. Um an A heranzukommen, müsste ich also B^{-1} bestimmen. Aus $A \cdot B = C$ folgt dann

$$A \cdot B \cdot B^{-1} = C \cdot B^{-1},$$

denn ich kann natürlich auf beiden Seiten mit der gleichen Matrix multiplizieren. Also ist

$$A \cdot I_n = C \cdot B^{-1},$$

denn schließlich ist $B \cdot B^{-1} = I_n$. Und damit wird

$$A = C \cdot B^{-1},$$

denn das Multiplizieren mit der Einheitsmatrix ändert nichts an der Matrix A. Auf diese Weise kann man also tatsächlich die Matrix A ausrechnen, sofern die inverse Matrix von B zur Verfügung steht.

Es wird einem nichts geschenkt im Leben, auch nicht bei den inversen Matrizen. Erstens habe ich nämlich noch kein Wort darüber verloren, wie man so eine inverse Matrix konkret ausrechnet, und zweitens gibt es – ganz im Gegensatz zu den vertrauten Zahlen – leider Matrizen, die nicht nur aus Nullen bestehen und trotzdem keine inverse Matrix haben, zu denen man also keine „Kehrmatrix" bilden kann.

Beispiel 5.21:

a) Ich versuche, die inverse Matrix von $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$ zu berechnen. Zu diesem Zweck nehme ich einfach mal an, dass A tatsächlich invertierbar ist, also eine inverse Matrix besitzt. Dann gibt es eine inverse Matrix $A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, mit der Eigenschaft $A \cdot A^{-1} = I_2$, denn genauso war die inverse Matrix definiert. Das heißt aber:

$$\begin{aligned} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= I_2 \\ &= A \cdot A^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &= \begin{pmatrix} a+c & b+d \\ 2a+2c & 2b+2d \end{pmatrix}, \end{aligned}$$

wobei ich am Ende nichts weiter getan habe als die beiden Matrizen auf dem üblichen Weg miteinander zu multiplizieren. Insgesamt hat sich also die Gleichung

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ 2a+2c & 2b+2d \end{pmatrix}$$

ergeben, und somit müssen diese beiden Matrizen die gleichen Einträge haben. Daraus folgt insbesondere $1 = a+c$ und $0 = 2a+2c$. Das ist aber völlig unmöglich, da $2a + 2c$ genau das Doppelte von $a + c$ ist. Wenn nun $a + c = 1$ ist, dann bleibt $2a+2c$ nichts anderes übrig, als den Wert 2 anzunehmen, und ganz sicher ist $2 \neq 0$. Wenn die Matrix A also eine inverse Matrix hätte, dann würde das dazu führen, dass eine Zahl gleichzeitig 0 und 2 sein kann, was doch eher selten vorkommt. Also kann meine Matrix A keine inverse Matrix besessen haben, und es folgt, dass A nicht invertierbar ist.

b) Jetzt betrachte ich die etwas größere Matrix

$$A = \begin{pmatrix} 1 & 0 & -1 \\ -8 & 4 & 1 \\ -2 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Wie können Sie nachrechnen, dass

$$A^{-1} = \begin{pmatrix} 1 & 1 & -4 \\ 2 & 2 & -7 \\ 0 & 1 & -4 \end{pmatrix}$$

gilt? Ganz einfach, indem Sie $A \cdot A^{-1}$ ausrechnen und feststellen, dass die Einheitsmatrix I_3 heraus kommt. Das ist leicht, aber wie bin ich auf diese Matrix A^{-1} gekommen? Das ist schon wesentlich schwieriger und wird im nächsten Abschnitt behandelt. ■

Noch ein kleines Beispiel zum Üben:

Übungsaufgabe 5.8:

Weisen Sie nach, dass die Matrix $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ keine Inverse hat. ■

Sie haben in Beispiel 5.21a) gesehen, dass das Invertieren von Matrizen wohl etwas mit linearen Gleichungssystemen zu tun hat, denn ich bin beim Rechnen auf zwei Gleichungen mit zwei Unbekannten gestoßen, die zusammen ein lineares Gleichungssystem bildeten. Schon im Zusammenhang mit der linearen Abhängigkeit von Vektoren haben sich diese Gleichungssysteme in den Vordergrund gedrängt, und jetzt schon wieder! Es wird offenbar langsam Zeit, dass ich sie Ihnen etwas näher bringe.

5.3 Lineare Gleichungssysteme

Dass lineare Gleichungssysteme zur Lösung des einen oder anderen der Probleme gebraucht werden, mit denen ich Sie in den ersten beiden Abschnitten traktiert habe, ist inzwischen klar geworden. Aber nicht nur für die selbst gemachten Probleme muss man sie in Anspruch nehmen, sie treten auch in den verschiedensten Anwendungszusammenhängen auf, seien sie nun eher technisch oder stärker ökonomisch orientiert. Worum es eigentlich geht, zeige ich Ihnen an einem ganz einfachen Anwendungsfall: der Berechnung von Frühstückskosten.

Beispiel 5.22:
Ein Student holt sich zum Frühstück zwei Brötchen und einen Liter Milch. Dafür zahlt er einen Euro und zehn Cent. Bezeichnet man mit B den Brötchenpreis und mit M den Milchpreis, so kann man diesen Sachverhalt in die Gleichung $2B + M = 1{,}1$ übersetzen. Sie hat leider keine eindeutige Lösung; zum Beispiel sind $B = 0{,}5$ und $M = 0{,}1$ Lösungen, aber auch $B = 0{,}25$ und $M = 0{,}6$. Natürlich ist es mein Ziel, die Lebensmittelpreise festzustellen, und daher brauche ich noch mehr Informationen über die Essgewohnheiten meines Studenten. Die bekomme ich am nächsten Tag, denn jetzt hat er Besuch zum Frühstück und kauft fünf Brötchen, zwei Liter Milch und sechs Eier zum Preis von drei Euro und fünfunddreißig Cent. Ich will mich nicht darüber auslassen, warum die beiden Frühstücker eine so heftige Stärkung schon am frühen Morgen brauchen, sondern lieber die daraus entstehende neue Gleichung aufstellen. Sie lautet:

$$5B + 2M + 6E = 3{,}35,$$

wobei E für den Preis eines Eis steht. Beide Gleichungen zusammen sind immer noch nicht eindeutig lösbar; zum Beispiel sind $B = 0{,}25, M = 0{,}6, E = 0{,}15$, aber auch $B = 0{,}3, M = 0{,}625, E = 0{,}1$ Lösungen. Ich bin also immer noch nicht in der Lage, die einzelnen Lebensmittel mit eindeutigen Preisen zu versehen. Tags darauf wird sich das aber ändern, denn der unter Beobachtung stehende Student verändert wieder seine Frühstückseinkäufe, die diesmal zu der Gleichung $4B + 2M + 4E = 2{,}8$ führen.

Nach drei Tagen habe ich jetzt drei Gleichungen mit drei Unbekannten zusammengetragen, und so etwas nennt man ein lineares Gleichungssystem. Es lautet also:

$$\begin{array}{rcrcrcl} 2B & + & M & & & = & 1{,}1 \\ 5B & + & 2M & + & 6E & = & 3{,}35 \\ 4B & + & 2M & + & 4E & = & 2{,}8. \end{array}$$

Gesucht sind natürlich B, M und E. Um die erste und die dritte Gleichung ein wenig ähnlicher zu machen, multipliziere ich die erste Gleichung mit zwei und erhalte $4B + 2M = 2,2$. Das kann ich nun gut mit der dritten Gleichung kombinieren, denn die fängt genauso an wie die verdoppelte zweite. Zu diesem Zweck schreibe ich beide untereinander

$$\begin{array}{rcrcrcl} 4B & + & 2M & & & = & 2{,}2 \\ 4B & + & 2M & + & 4E & = & 2{,}8 \end{array}$$

und stelle fest, dass es hilfreich sein könnte, die erste Gleichung von der zweiten abzuziehen, das heißt: die linke Seite der ersten Gleichung von der linken Seite der zweiten und die rechte Seite der ersten Gleichung von der rechten Seite der zweiten. Da es sich um Gleichungen handelt, sind diese Operationen erlaubt, und glücklicherweise fallen dabei die Unbekannten B und M weg. Übrig bleibt:

$$4E = 0{,}6, \text{ also } E = 0{,}15.$$

Das ist gut, denn jetzt weiß ich schon mal, dass ein Ei fünfzehn Cent kostet. Diesen Wert kann ich nun in die ersten beiden Gleichungen meines ursprünglichen Gleichungssystems einsetzen und erhalte:

$$\begin{array}{rcrcr} 2B & + & M & = & 1{,}1 \\ 5B & + & 2M & = & 2{,}45, \end{array}$$

denn die alte zweite Gleichung hieß $5B + 2M + 6E = 3{,}35$, also $5B + 2M + 0{,}9 = 3{,}35$ und damit $5B + 2M = 2{,}45$. Nach dem gleichen Prinzip wie eben versuche ich jetzt wieder, eine Unbekannte loszuwerden, und zu diesem Zweck multipliziere ich die erste Gleichung mit zwei. Es ergibt sich

$$\begin{array}{rcrcr} 4B & + & 2M & = & 2{,}2 \\ 5B & + & 2M & = & 2{,}45. \end{array}$$

Abziehen der ersten Gleichung von der zweiten liefert dann $B = 0{,}25$, und schon ist der Brötchenpreis berechnet. Setzt man ihn noch in die allererste Gleichung $2B + M = 1{,}1$ ein, so folgt $0{,}5 + M = 1{,}1$, also $M = 0{,}6$. Insgesamt habe ich also $B = 0{,}25$; $M = 0{,}6$; $E = 0{,}15$. ■

Das Verfahren, das ich hier gerade vorgeführt habe, wird Ihnen vielleicht bekannt vorgekommen sein: Es ist das klassische Additionsverfahren, mit dem man Sie sicher auch schon in der Schule geplagt hat. In diesem Abschnitt will ich nichts anderes tun, als dieses Verfahren etwas zu systematisieren und ein Schema daraus zu machen, das man ohne großartig nachdenken zu müssen anwenden kann. Dazu muss ich erst einmal aufschreiben, was genau man unter einem linearen Gleichungssystem versteht. Das ist aber nicht weiter schwer. Eine einzelne lineare Gleichung hat einige Unbekannte, man weiß aber nicht unbedingt, wie viele, und deshalb sagt man, sie hat die Unbekannten $x_1, x_2, ..., x_n$ Im Beispiel 5.21 war also $n = 3$. Da die Gleichung linear sein soll, darf man mit den Unbekannten nur das anstellen, was ich auch im Beispiel mit ihnen gemacht habe: Man darf sie mit Konstanten multiplizieren und dann addieren oder subtrahieren. Jetzt muss ich nur noch Namen für die konstanten Vorfaktoren der Unbekannten erfinden, und das mache ich genauso wie bei den Einträgen für Matrizen. Wenn es sich bei meiner linearen Gleichung um die erste Gleichung des Systems handelt, schreibe ich

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1.$$

Die Konstante a_{11} gehört also in der *ersten* Gleichung zur *ersten* Unbekannten, a_{12} gehört in der *ersten* Gleichung zur *zweiten* Unbekannten und so weiter, bis dann schließlich mit b_1 die rechte Seite der ersten Gleichung erscheint. Man liest daher auch hier wieder a_{11} als a eins eins und nicht etwa als a elf. Die zweite Gleichung heißt dann natürlich nach dem gleichen Muster

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_.$$

Die Zusammenfassung von einigen – sagen wir m – linearen Gleichungen mit denselben n Unbekannten ist dann ein lineares Gleichungssystem.

Lineares Gleichungssystem

Das aus m linearen Gleichungen mit n Unbekannten $x_1, .., x_n$ bestehende System

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m, \end{array}$$

wobei die Zahlen a_{ij} und b_i bekannt sind, heißt lineares Gleichungssystem. Mit der **Koeffizientenmatrix**

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

der rechten Seite $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ und dem Vektor $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ schreibt man dafür auch oft

$$A \cdot \mathbf{x} = \mathbf{b}.$$

Sie sehen, dass ich das Gleichungssystem auf zwei verschiedene Arten aufgeschrieben habe. Die erste Methode ist die konventionelle, die Sie sicher aus der Schule kennen. Die zweite Schreibweise ist eine Matrixschreibweise, die auf dem beruht, was Sie im Abschnitt über Matrizen gelernt haben. Da nämlich jede Gleichung auf demselben Schema beruht und vor allem immer die gleichen Unbekannten vorkommen, erspart man sich die Mühe, immer wieder die gleichen Unbekannten aufzuschreiben, und sammelt nur die Koeffizienten in einer großen Matrix A und die rechte Seite in einem Vektor $\mathbf{b}$. Wenn Sie dann das Produkt $A \cdot \mathbf{x}$ so ausrechnen, wie wir es im zweiten Abschnitt besprochen haben, werden Sie feststellen, dass genau die Versammlung der linken Seiten meines Gleichungssystems heraus kommt. Die etwas knappe Schreibweise $A \cdot \mathbf{x} = \mathbf{b}$ ist also nichts weiter als eine verkürzende Schreibweise für ein lineares Gleichungssystem.

Nun gut, aufschreiben können Sie jetzt so ein lineares Gleichungssystem, aber wie kann man es systematisch lösen? Das zeige ich Ihnen gleich im nächsten Beispiel. Vorher muss ich allerdings noch kurz etwas anderes mit Ihnen durchgehen. Im Beispiel 5.21 habe ich die gegebenen Gleichungen immer wieder manipuliert, ohne damit Schaden anzurichten, weil sich durch diese Manipulationen nichts an den Lösungen des Gleichungssystems geändert hat. Am besten notieren wir einmal kurz, welche Manipulationen ohne jedes Risiko durchgeführt werden können:

Erlaubte Manipulationen an Gleichungssystemen

Die Lösungsmenge eines linearen Gleichungssystems wird durch die folgenden Operationen nicht verändert:
a) Multiplikation einer Gleichung mit einer von Null verschiedenen Zahl.
b) Addieren eines Vielfachen einer Gleichung auf eine andere.
c) Vertauschen zweier Gleichungen.

Dem ist wohl nicht viel hinzuzufügen, außer vielleicht die Tatsache, dass man mit b) natürlich auch Gleichungen von anderen Gleichungen abziehen kann, da Sie auch mit negativen Zahlen multiplizieren dürfen. Im nächsten Beispiel zeige ich Ihnen, wie man die eben erteilte Lizenz zum Manipulieren ausnützt und auch noch kurz und prägnant aufschreibt. Der Trick wird darin bestehen, ein Gleichungssystem in Matrixform zu schreiben und dann die Zeilen dieser Matrix so zu verändern, dass man am Ende die Lösung direkt ablesen kann.

Beispiel 5.23:
Ich betrachte das lineare Gleichungssystem

$$\begin{array}{rcrcrcr} -x & + & 2y & + & z & = & -2 \\ 3x & - & 8y & - & 2z & = & 4 \\ x & & & + & 4z & = & -2. \end{array}$$

mit den Unbekannten x, y und z. Die Koeffizientenmatrix lautet dann

$$A = \begin{pmatrix} -1 & 2 & 1 \\ 3 & -8 & -2 \\ 1 & 0 & 4 \end{pmatrix}$$

und auf der rechten Seite steht der Vektor

$$\mathbf{b} = \begin{pmatrix} -2 \\ 4 \\ -2 \end{pmatrix}.$$

Zu einem Gleichungssystem gehören aber immer vollständige Gleichungen einschließlich der rechten Seite, weshalb ich die eben aufgeschriebene rechte Seite der Matrix A hinzufüge. Das ergibt dann die um eine Spalte breitere Matrix

$$(A, b) = \begin{pmatrix} -1 & 2 & 1 & -2 \\ 3 & -8 & -2 & 4 \\ 1 & 0 & 4 & -2 \end{pmatrix}.$$

Jede Zeile dieser Matrix entspricht nun einer Gleichung des gegebenen Gleichungssystems: Die ersten drei Zahlen geben die Koeffizienten der Unbekannten an, während die letzte Zahl die rechte Seite beschreibt. Nun will ich dafür sorgen, dass eine Unbekannte verschwindet, und das bedeutet in der Matrixschreibweise, dass zum Beispiel die Koeffizienten der Unbekannten x zu Null werden, denn $0 \cdot x$ ist nichts. Ich addiere deshalb das Dreifache der ersten Zeile auf die zweite sowie die erste Zeile auf die dritte. Damit habe ich eigentlich nichts anderes getan als die dreifache erste Gleichung auf die zweite zu addieren und die erste Gleichung selbst auf die dritte, und solche Operationen sind immer erlaubt. Die neue Matrix lautet:

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

In der zweiten und dritten Zeile stehen jetzt jeweils Nullen am Anfang. Der Koeffizient von x ist also zu Null geworden, mit anderen Worten: x ist aus diesen beiden Gleichungen eliminiert und die zweite und die dritte Zeile entsprechen zusammen zwei Gleichungen mit zwei Unbekannten, nämlich $-2y + z = -2$ und $2y + 5z = -4$. Aus der letzten Gleichung sollte nun auch noch y verschwinden, damit nur eine Gleichung mit einer Unbekannten übrig bleibt. Dazu addiere ich einfach die zweite Zeile auf die dritte und erhalte

$$\begin{pmatrix} -1 & 2 & 1 & -2 \\ 0 & -2 & 1 & -2 \\ 0 & 0 & 6 & -6 \end{pmatrix}.$$

Die letzte Zeile entspricht jetzt der Gleichung $6z = -6$, und das macht schon klar, was ich jetzt tun muss. Dividieren der letzten Zeile durch 6 ergibt:

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

Nun kann ich natürlich an der dritten Zeile direkt ablesen, dass $z = -1$ gilt. Es wäre angenehm, die gleiche Situation auch für die anderen Unbekannten herstellen zu können, aber das ist jetzt nicht mehr schwer. Um in der zweiten Zeile nur noch mit y zu rechnen, muss der Eintrag 1 von der dritten Stelle verschwinden, weil in diesem Fall sowohl der Vorfaktor von x als auch der von z gleich null ist. Zu diesem Zweck ziehe ich die dritte Zeile von der zweiten ab und finde:

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

Die zweite Zeile entspricht nun der Gleichung $-2y = -1$, die man natürlich durch -2 teilen sollte. Daher dividiere ich die zweite Zeile durch -2 mit dem Resultat:

$$\begin{pmatrix} -1 & 2 & 1 & -2 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Sie wissen wohl schon, was jetzt passieren muss. Außer der -1 ganz vorne müssen alle Koeffizienten aus der ersten Zeile eliminiert werden, was ich leicht erreichen kann, indem ich die doppelte zweite Zeile und die unveränderte dritte Zeile von der ersten abziehe:

$$\begin{pmatrix} -1 & 0 & 0 & -2 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Multiplikation der ersten Zeile mit -1 ergibt schließlich:

$$\begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

Endlich ist es geschafft. Übersetzt man jetzt wieder die einzelnen Zeilen in Gleichungen, dann kommt in der ersten Gleichung nur noch x vor, in der zweiten nur noch y und in der dritten nur noch z, und man sieht auf einen Blick:

$$x = 2, y = \frac{1}{2}, z = -1.$$

■

Haben Sie gesehen, worauf das Ganze hinausläuft? Aus drei Gleichungen mit drei Unbekannten habe ich zwei Gleichungen mit zwei Unbekannten gemacht, indem ich in der zweiten und dritten Zeile meiner Matrix die Koeffizienten von x zu Null werden ließ. Das Addieren oder Subtrahieren von Zeilen entspricht dabei genau dem Addieren oder Subtrahieren von Gleichungen, und das war ja ausdrücklich erlaubt. Aus den beiden entstandenen Gleichungen habe ich dann eine Gleichung mit nur noch einer Unbekannten extrahiert, indem ich in der dritten Zeile der Matrix den Koeffizienten von y zur Bedeutungslosigkeit verdammt habe. Bei diesem Stand der Dinge hatte ich eine Matrix errreicht, die links unterhalb der Hauptdiagonalen nur noch aus Nullen bestand, und das ist immer das erste Ziel beim Lösen von Gleichungssystemen. Anschließend habe ich mich dann mühsam wieder von unten nach oben gehangelt, um auch rechts oberhalb der Hauptdiagonalen (mit Ausnahme natürlich der rechten Seite) nur noch Nullen zu haben und daraus am Ende die Lösungen direkt ablesen zu können.

Dieses Verfahren heißt **Gauß-Algorithmus**. Es beruht darauf, die Matrix (A, b) mit einfachen Zeilenumformungen auf die Form

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & x_1 \\ 0 & 1 & \cdots & 0 & x_2 \\ \vdots & 0 & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & x_n \end{pmatrix}$$

zu bringen, wobei $x_1, ..., x_n$ die Lösungen des Gleichungssystems sind. Der Gauß-Algorithmus existiert in verschiedenen Varianten, die aber eigentlich alle auf das Gleiche hinauslaufen. Zunächst sehen wir ihn uns für den Fall an, dass das Gleichungssystem für jede der Unbekannten zu genau einer Lösung führt:

Gauß-Algorithmus

Es sei $A\mathbf{x} = \mathbf{b}$ ein lineares Gleichungssystem mit n Gleichungen und n Unbekannten, das eine eindeutige Lösung hat. Man startet mit der Matrix

$$(A|b) = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & b_n \end{pmatrix},$$

in der die Koeffizienten des Gleichungssystems und seine rechte Seite aufgelistet sind. Dann sind die folgenden Schritte durchzuführen:

a) Falls $a_{11} = 0$, suche man eine Zeile, deren erstes Element von Null verschieden ist, vertausche diese Zeile mit der ersten Zeile und benenne sie um. In jedem Fall gilt dann $a_{11} \neq 0$.

b) Man subtrahiere ein geeignetes Vielfaches der ersten Zeile von der zweiten, dritten, ..., letzten Zeile, sodass diese Zeilen jeweils mit Null beginnen. Die neue Matrix hat dann in der ersten Spalte unterhalb von a_{11} nur noch Nullen.

c) Man subtrahiere ein geeignetes Vielfaches der zweiten Zeile von der dritten, ..., letzten Zeile, so dass diese Zeilen jeweils mit Null beginnen und auch an der zweiten Stelle eine Null haben. Die neue Matrix hat dann in der ersten Spalte unterhalb von a_{11} und in der zweiten Spalte unterhalb von a_{22} nur noch Nullen.

d) Man wiederhole dieses Verfahren für die nachfolgenden Spalten der Matrix so lange, bis links unterhalb der Hauptdiagonalelemente nur noch Nullen stehen.

e) Man teile die letzte Zeile durch den letzten von Null verschiedenen Koeffizienten, sodass am Ende der Hauptdiagonale eine Eins steht. Anschließend gehe man mit den passenden Zeilenumformungen von unten nach oben vor, sodass die Matrix zum Schluss die folgende Form hat:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & x_1 \\ 0 & 1 & \cdots & 0 & x_2 \\ \vdots & 0 & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 & x_n \end{pmatrix}.$$

In der letzten Spalte steht dann die Lösung des linearen Gleichungssystems. Falls dieses Verfahren nicht zum Ziel führt, hat das lineare Gleichungssystem keine eindeutige Lösung.

Auf den ersten Blick ziemlich abstrakt, ich weiß. Es bleibt einem aber gar nichts anderes übrig, als ein wenig abstrakt zu werden, wenn man ein einigermaßen aufwendiges Verfahren irgendwie allgemein gültig aufschreiben will. Ein weiteres Beispiel wird Ihnen zeigen, dass eigentlich gar nicht so viel dahinter steckt.

Beispiel 5.24:
Ich will das lineare Gleichungssystem

$$\begin{array}{ccccccccc} x & & & + & z & + & u & = & 0 \\ x & + & y & + & 2z & + & u & = & 1 \\ & - & y & & & + & u & = & 0 \\ x & & & & & + & 2u & = & 0 \end{array}$$

lösen. Die Matrizendarstellung dieses Gleichungssystems lautet:

$$(A,b) = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 2 & 1 & 1 \\ 0 & -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 & 0 \end{pmatrix}.$$

Nun muss ich mir in der ersten Spalte unterhalb der obersten Eins ausschließlich Nullen verschaffen. Das Abziehen der ersten Zeile von der zweiten und vierten liefert die erwünschte Matrix:

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \end{pmatrix}.$$

Addieren der zweiten auf die dritte Zeile liefert:

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & 1 & 0 \end{pmatrix},$$

und schon habe ich auch in der zweiten Spalte unterhalb des Diagonalelementes nur Nullen und kann mich der dritten Spalte widmen. Addieren der dritten auf die vierte Zeile liefert:

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 & 1 \end{pmatrix}.$$

Sie sehen, dass nun unterhalb der Hauptdiagonalen nur noch Nullen stehen, aber das letzte Element der Hauptdiagonalen hat noch den ungünstigen Wert Zwei. Deshalb dividiere ich die vierte Zeile durch zwei und erhalte:

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}.$$

Der erste Teil des Verfahrens ist damit erledigt, ich habe mich von oben nach unten gearbeitet und unterhalb der Hauptdiagonalen Nullen erzeugt. Jetzt wechsle ich die

Blickrichtung und kämpfe mich von unten nach oben durch, um auch rechts oberhalb der Diagonalen alles zu vernullen. Die vierte Spalte erledige ich dabei in einem Schritt, indem ich die vierte Zeile von der dritten und der ersten abziehe mit dem Ergebnis:

$$\begin{pmatrix} 1 & 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & \frac{1}{2} \end{pmatrix}.$$

Anschließend ziehe ich die dritte Zeile von der ersten und der zweiten ab, was mich am Ende zu der Matrix

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

führt, an der ich direkt die Ergebnisse ablesen kann, denn nun ist

$$x = -1, y = z = u = \frac{1}{2}.$$ ■

Es ist also im Grunde genommen keine große Sache. Zuerst arbeitet man sich von links oben nach rechts unten und sorgt dafür, dass der Reihe nach in allen Spalten unterhalb der Hauptdiagonalen nur noch Nullen stehen, wobei man mit der ersten Spalte anfängt. Ist das erst einmal erledigt, dreht man die Arbeitsrichtung um und kümmert sich darum, dass der Reihe nach in allen Spalten auch oberhalb der Hauptdiagonalen nur noch Nullen stehen, wobei man jetzt hinten anfängt und vorne aufhört.

Dazu wieder einmal Übungen.

Übungsaufgabe 5.9:
Lösen Sie die folgenden linearen Gleichungssysteme:

$$\begin{array}{rcrcrcr} x & + & y & - & 5z & = & -10 \\ 3x & - & 4y & + & z & = & 9 \\ 6x & + & 2y & + & 4z & = & 12 \end{array} \quad \text{und} \quad \begin{array}{rcrcrcrcr} x & + & 3y & + & 2z & - & 6u & = & 1 \\ x & + & 2y & - & 2z & - & 5u & = & -1 \\ 2x & + & 4y & - & 2z & - & 9u & = & 0 \\ 2x & + & 4y & - & 6z & - & 9u & = & 2 \end{array}$$

■

Das ist schon mal augezeichnet, aber leider reicht es noch nicht ganz, denn nicht alle linearen Gleichungssysteme haben eine eindeutige Lösung.

Beispiel 5.25:
Ich suche nach den Lösungen des linearen Gleichungssystems

$$\begin{array}{rcrcrcrcr} x & + & 2y & + & z & + & 3u & = & 3 \\ 2x & + & 4y & + & 2z & + & 6u & = & 5 \\ 3x & + & 6y & + & 2z & + & 10u & = & -3 \end{array}$$

mit drei Gleichungen und den vier Unbekannten x, y, z und u. Inzwischen ist das Aufschreiben der Matrixform für Sie nur noch Routine; sie lautet:

$$(A, b) = \begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 2 & 4 & 2 & 6 & 5 \\ 3 & 6 & 2 & 10 & -3 \end{pmatrix}.$$

Nach meinem Gauß-Schema muss ich das Zweifache der ersten Zeile von der zweiten und das Dreifache der ersten Zeile von der dritten abziehen. Das ergibt:

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 1 & -12 \end{pmatrix}.$$

Die zweite Zeile macht einen etwas eigenartigen Eindruck, da sie im Bereich der Koeffizienten nur noch Nullen enthält und wir Zeilen mit so vielen Nullen ja eigentlich erst am Ende der Matrix gewöhnt sind. Ich vertausche daher die zweite mit der dritten Zeile, was nur dem Vertauschen zweier Gleichungen entspricht und daher erlaubt ist. Damit bekomme ich die Matrix

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 3 \\ 0 & 0 & -1 & 1 & -12 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

Und jetzt richten Sie Ihr Augenmerk einmal auf die dritte Zeile und übersetzen Sie sie wieder zurück in eine Gleichung. Da ich bei den Koeffizienten hier nur Nullen habe, lautet die Gleichung:

$$0x + 0y + 0z + 0u = -1, \text{ also } 0 = -1.$$

Kommt so etwas vor? Doch eher selten, ich habe hier eine unsinnige Gleichung erhalten, die niemals erfüllt werden kann. Welche Werte auch immer Sie für die vier Unbekannten einsetzen mögen, Sie werden es nie schaffen, dass die Gleichung $0 = -1$ erfüllt wird. Wenn ich aber schon weiß, dass hier eine Gleichung vorliegt, die ich nie erfüllen kann, dann kann ich mir das weitere Suchen auch sparen und mich auf die Couch oder in die Badewanne legen: Es würde ja ohnehin zu nichts führen. Keine Wertekombination der Welt kann diese unmögliche Gleichung $0 = -1$ retten, und das heißt einfach nur, dass es keine Lösung gibt! Dieses lineare Gleichungssystem ist schlicht und ergreifend unlösbar. Man sagt dann entweder genau diesen Satz „Das Gleichungssystem ist unlösbar“ oder gibt die Lösungsmenge als die leere Menge $\emptyset$ an, also als die etwas trostlose Menge, in der man nicht ein einziges Element findet. ■

Es gibt also auch lineare Gleichungssysteme, die keine Lösung erlauben und deren Lösungsmenge deshalb leer ist. Das Schöne ist, dass Sie sich keine grauen Haare über die Frage wachsen lassen müssen, wie man solche unzuverlässigen Gleichungssysteme erkennt, denn das Beispiel hat es ganz klar gezeigt: Sie starten den Gauß-Algorithmus genauso, wie Sie ihn kennen gelernt haben, und wenn er auf eine Zeile führt, bei der nur in der letzten Spalte keine Null steht, dann hat das Gleichungssystem keine Lösung.

Unlösbare lineare Gleichungssysteme

Ist $A \cdot \mathbf{x} = \mathbf{b}$ ein lineares Gleichungssystem, bei dem der Gauß-Algorithmus zu einer Zeile der Form $\begin{pmatrix} 0 & 0 & \ldots & 0 & b \end{pmatrix}$ mit $b \neq 0$ führt, in der alle Einträge außer dem letzten gleich null sind, so hat das lineare Gleichungssystem keine Lösung.

Wo diese unpassende Zeile steht, ist dabei völlig egal; sobald sie auftaucht, können Sie getrost jedes weitere Rechnen unterlassen und sich damit begnügen, dass das Gleichungssystem unlösbar ist.

Übungsaufgabe 5.10:
Bestimmen Sie die Lösung des Gleichungssystems

$$\begin{array}{rcrcrcl} x & + & y & + & 2z & = & 1 \\ x & + & 5y & + & 2z & = & 3 \\ 2x & - & 2y & + & 4z & = & 5\,. \end{array}$$

■

Es gibt also Gleichungssysteme, bei denen Sie für jede Unbekannte genau eine Lösung finden, und es gibt Gleichungssysteme, die überhaupt keine Lösung haben. War's das schon? Leider nein, eine dritte Möglichkeit gibt es auch noch, nämlich unendlich viele Lösungen. Das sieht man schon an dem übersichtlichen Beispiel des Gleichungssystems, das nur aus der einen Gleichung $x + y = 0$ besteht, denn offenbar kann ich hier $x = y = 0$ oder $x = 1, y = -1$ oder $x = 2, y = -2$ usw. als Lösungen heranziehen. An einem etwas größeren Beispiel zeige ich Ihnen jetzt, wie man mit so einer Situation umgeht.

Beispiel 5.26:
Ich untersuche das Gleichungssystem

$$\begin{array}{rcrcrcl} x & + & y & + & z & = & 0 \\ 2x & + & y & - & z & = & 3 \\ 3x & + & 2y & & & = & 3\,. \end{array}$$

In Matrixform heißt es

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

Das sieht noch ganz normal aus, also werde ich wie üblich den Gauß-Algorithmus starten. Um in der ersten Spalte die nötigen Nullen zu produzieren, ziehe ich die doppelte erste Zeile von der zweiten ab und die dreifache erste von der dritten. Das ergibt:

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

Noch immer treten keine Probleme auf, aber die werden sich gleich einfinden, wenn ich das Schema fortführe und die zweite Zeile von der dritten abziehe, um die störende

-1 in der dritten Zeile zu eliminieren. Die Matrix lautet dann:

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

Das ist nun eine neue Situation, denn in der dritten Zeile treten nur noch Nullen auf, was ich sonst nur von Sitzungen kenne. In eine Gleichung übersetzt, lautet die dritte Zeile nur $0 = 0$, und das ist keine sehr hilfreiche Information. Mit anderen Worten: Durch die Anwendung des Gauß-Algorithmus hat sich herausgestellt, dass in dem Gleichungssystem eigentlich nur die Information aus zwei Gleichungen mit drei Unbekannten verpackt war und nicht aus drei Gleichungen. Ein genauerer Blick auf die Gleichungen bestätigt diese Diagnose, denn die dritte Gleichung ist nur die Summe aus den ersten beiden Gleichungen und konnte daher nichts Neues bieten.

Die Nullzeile ist als Informant nicht zu gebrauchen. Es ist auch nicht mehr möglich, entlang der Hauptdiagonalen lauter Einsen zu produzieren, weil in der letzten Zeile nun mal nur noch Nullen stehen. Das ursprüngliche Ziel des Gauß-Algorithmus kann ich also nicht mehr erreichen. Das macht aber nichts. Bisher haben wir im Zuge des Gauß-Algorithmus versucht, in der Hauptdiagonalen Einsen und ansonsten Nullen zu produzieren, und das bedeutet nur, dass eine Einheitsmatrix hergestellt werden sollte. Und eine Einheitsmatrix kann ich immer noch auf die Reihe bringen; die wird dann zwar nicht mehr ganz so groß, aber besser eine kleine Einheitsmatrix als gar keine. Wie die Sache aussieht, kann ich beim vorliegenden Gleichungssystem keine Einheitsmatrix mit drei Zeilen und drei Spalten bekommen, sondern nur eine mit zwei Zeilen und zwei Spalten. Dazu multipliziere ich die zweite Zeile mit -1 und subtrahiere sie anschließend von der ersten. Das ergibt:

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

Wie Sie sehen, habe ich jetzt links oben eine etwas kleinere Einheitsmatrix aus zwei Zeilen und zwei Spalten, und mehr kann der Mensch nicht verlangen. Übersetzt man die verbliebenen zwei Zeilen nämlich in Gleichungen, so lauten sie $x - 2z = 3$ und $y + 3z = -3$, mehr Information ist nicht da. Da ich schon eine kleine Einheitsmatrix hergestellt hatte, kann ich nun ganz leicht nach den beiden Unbekannten x und y auflösen, denn beide haben bereits den Vorfaktor 1. Es gilt also:

$$x = 3 + 2z \text{ und } y = -3 - 3z.$$

Weitere Informationen gibt dieses Gleichungssystem nicht her, so lange Sie auch suchen mögen. Offenbar ist es aufgrund dieser zwei Gleichungen mit drei Unbekannten nicht möglich, eindeutige Lösungen für x, y und z auszurechnen, dafür habe ich zu wenige Informationen. Aber Lösungen gibt es trotzdem: Ist beispielsweise $z = 0$, so folgt aus den beiden berechneten Gleichungen sofort $x = 3$ und $y = -3$. Ist dagegen $z = 1$, so finden Sie $x = 5$ und $y = -6$. Was auch immer Sie für z einsetzen, Sie können aus den beiden errechneten Gleichungen problemlos sofort x und y ausrechnen. Deshalb hat dieses lineare Gleichungssystem unendlich viele Lösungen, denn man kann für z einsetzen, was man will, und daraus die Werte für x und y berechnen. Die Lösungen haben also die Form

$$x = 3 + 2z, y = -3 - 3z, z \in \mathbb{R} \text{ beliebig.}$$

Da man in diesem Fall für z einsetzen kann, was man will, nennt man z auch einen freien Parameter. ■

So funktioniert das immer. Wenn im Verlauf des Gauß-Algorithmus Nullzeilen auftreten, dann kann man sie getrost ignorieren und eine möglichst große Einheitsmatrix aufbauen. Hat diese Einheitsmatrix zum Beispiel zwei Zeilen und zwei Spalten und gibt es mehr als zwei Unbekannte, dann kann man für die überschüssigen Unbekannten (im Beispiel 5.26 war das z) beliebige Werte einsetzen und die Werte für die anderen Unbekannten direkt aus den Gleichungen ausrechnen, die aus der letzten errechneten Matrix resultieren.

Gauß-Algorithmus

Zur Lösung des linearen Gleichungssystems $A \cdot \mathbf{x} = \mathbf{b}$ bringe man die Matrix (A, b) durch die bekannten Operationen auf eine Form, in der links oben eine möglichst große Einheitsmatrix steht, wobei unterhalb dieser Einheitsmatrix nur noch reine Nullzeilen oder aber Zeilen der Form $\begin{pmatrix} 0 & 0 & \dots & 0 & b \end{pmatrix}$ vorkommen.

Stehen unterhalb der Einheitsmatrix Zeilen der Form $\begin{pmatrix} 0 & 0 & \dots & 0 & b \end{pmatrix}$ mit $b \neq 0$, so ist das Gleichungssystem unlösbar. Kommen reine Nullzeilen vor, so kann man sie ignorieren. Reicht die Einheitsmatrix nicht bis ganz an die letzte Spalte (also an die rechte Seite) heran, so kann man die „überschüssigen“ Unbekannten frei wählen und die anderen Unbekannten aus der neuen Matrix berechnen.

Auch das ist wieder eine recht abstrakte Formulierung, weshalb ein weiteres Beispiel nicht schaden kann:

Beispiel 5.27:
Gegeben sei das lineare Gleichungssystem

$$\begin{array}{rcrcrcrcr} x & + & 5y & + & 2z & + & 3u & = & 4 \\ 4x & + & 18y & + & 2z & + & 8u & = & 12 \\ 3x & + & 11y & - & 6z & + & u & = & 4 \\ & & 2y & + & 6z & + & 4u & = & 4\,. \end{array}$$

In Matrixform lautet das System

$$(A, b) = \begin{pmatrix} 1 & 5 & 2 & 3 & 4 \\ 4 & 18 & 2 & 8 & 12 \\ 3 & 11 & -6 & 1 & 4 \\ 0 & 2 & 6 & 4 & 4 \end{pmatrix}.$$

Wieder einmal kommt der Gauß-Algorithmus zum Einsatz, der unterhalb der links oben stehenden Eins für Nullen sorgen muss. Zu diesem Zweck ziehe ich die vierfache erste Zeile von der zweiten ab und die dreifache erste von der dritten. Das ergibt die Matrix:

$$\begin{pmatrix} 1 & 5 & 2 & 3 & 4 \\ 0 & -2 & -6 & -4 & -4 \\ 0 & -4 & -12 & -8 & -8 \\ 0 & 2 & 6 & 4 & 4 \end{pmatrix}.$$

Da ich in der Hauptdiagonalen am liebsten Einsen haben möchte, teile ich die zweite Zeile durch -2. Das Resultat ist:

$$\begin{pmatrix} 1 & 5 & 2 & 3 & 4 \\ 0 & 1 & 3 & 2 & 2 \\ 0 & -4 & -12 & -8 & -8 \\ 0 & 2 & 6 & 4 & 4 \end{pmatrix}.$$

Unterhalb der neu geschaffenen Eins will ich jetzt wieder ganz nach Plan Nullen erzeugen. Das mache ich, indem ich die vierfache zweite Zeile auf die dritte addiere und die doppelte zweite Zeile von der vierten abziehe. Damit habe ich:

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

Zwei Zeilen und damit auch zwei Gleichungen haben sich als überflüssig erwiesen, die gesuchte Einheitsmatrix kann also nur noch aus zwei Zeilen und zwei Spalten bestehen. Sie ist auch leicht zu finden, denn ich muss nur noch die fünffache zweite Zeile von der ersten abziehen. Das führt zu:

$$\left(\begin{array}{cc|ccc} 1 & 0 & -13 & -7 & -6 \\ 0 & 1 & 3 & 2 & 2 \\ \hdashline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Hier habe ich die erreichte Einheitsmatrix deutlich vom Rest der Welt abgegrenzt, damit Sie sie besser im Blickfeld haben. Die letzten beiden Zeilen bestehen nur aus Nullen, spielen also keine Rolle mehr. Die kleine Einheitsmatrix aus zwei Zeilen und zwei Spalten zeigt, dass man die Unbekannten x und y aus z und u berechnen kann. Die entsprechenden Gleichungen lauten nämlich:

$$x - 13z - 7u = -6 \text{ und } y + 3z + 2u = 2.$$

Daraus folgt:

$$x = 13z + 7u - 6 \text{ sowie } y = -3z - 2u + 2.$$

Man kann also die Werte für z und u frei wählen – das sind die schon erwähnten freien Parameter – und erhält damit die Lösungen:

$$x = 13z + 7u - 6, y = -3z - 2u + 2, z, u \in \mathbb{R} \text{ beliebig.}$$

Auch hier gibt es also unendlich viele Lösungen. Die Werte für die Unbekannten z und u können Sie beliebig aus der Luft greifen, die Werte für x und y mit den errechneten Gleichungen leicht ausrechnen. Nimmt man zum Beispiel $z = 0, u = 1$, so erhält man $x = 1, y = 0$, während man für $z = 1, u = 0$ sofort $x = 7, y = -1$ erhält. Der Anzahl der Lösungen sind somit keine Grenzen gesetzt. ■

Wie Sie feststellen mussten, gibt es drei Möglichkeiten für ein lineares Gleichungssystem: Es kann gar keine Lösung haben oder für jede Unbekannte genau eine oder sogar unendlich viele. Das macht aber gar nichts, denn Sie haben auch gesehen, dass man all diese Möglichkeiten auf die gleiche Weise mithilfe des Gauß-Algorithmus testen und die Lösungen ausrechnen kann.

Nun wird wieder etwas gerechnet.

Übungsaufgabe 5.11:
Bestimmen Sie die Lösungen der folgenden linearen Gleichungssysteme.

$$\begin{array}{rcrcrcr} x & - & y & & & = & 2 \\ x & + & y & - & 2z & = & 0 \\ 2x & - & y & - & z & = & 3 \end{array} \quad \text{und} \quad \begin{array}{rcrcrcrcr} x & - & y & & & + & u & = & 0 \\ & & -2y & - & z & + & 2u & = & 0 \\ -2x & + & 5y & + & 2z & - & 4u & = & 0 \\ 4x & - & 13y & - & 6z & + & 10u & = & 0\,. \end{array}$$ ■

Sie werden sich erinnern, dass ich in den Abschnitten über Vektoren und über Matrizen jeweils ein Problem offen lassen musste, weil wir noch nicht so recht mit den linearen Gleichungssystemen umgehen konnten. Jetzt steht der engültigen Lösung dieser Probleme nichts mehr im Wege. Fangen wir an mit der linearen Abhängigkeit von Vektoren. Im ersten Abschnitt hatte ich Ihnen berichtet, dass die Vektoren $\mathbf{x}_1, ..., \mathbf{x}_m$ genau dann linear unabhängig sind, wenn aus

$$c_1 \cdot \mathbf{x}_1 + c_2 \cdot \mathbf{x}_2 + \cdots + c_m \cdot \mathbf{x}_m = \mathbf{0}$$

stets folgt:

$$c_1 = c_2 = \cdots = c_m = 0,$$

wobei unter $\mathbf{0}$ der Nullvektor verstanden wird. Kann man die Vektoren dagegen zum Nullvektor kombinieren, ohne dass jeder der Vorfaktoren $c_1, c_2, ..., c_m$ zu Null werden muss, so sind die Vektoren linear abhängig. Will man das nun für konkrete gegebene Vektoren testen, so gelangt man ganz schnell wieder einmal zu einem linearen Gleichungssystem.

Beispiel 5.28:
Ich will feststellen, ob die Vektoren

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$$

linear abhängig oder linear unabhängig sind. Dazu stelle ich die Gleichung

$$c_1 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_2 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + c_3 \cdot \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

auf und muss nachrechnen, ob daraus $c_1 = c_2 = c_3 = 0$ folgt oder ob es auch noch andere Möglichkeiten gibt. Diese Vektorengleichung übersetze ich in ein lineares Gleichungssystem, indem ich aufschreibe, was für die einzelnen Komponenten der Vektoren gelten muss. Das ergibt dann das Gleichungssystem

$$\begin{array}{ccccccc} c_1 & + & c_2 & + & c_3 & = & 0 \\ c_1 & & & + & 3c_3 & = & 0 \\ & & c_2 & - & 2c_3 & = & 0. \end{array}$$

In Matrixform heißt das:

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

Ab jetzt ist alles Routine. Um in der ersten Spalte unterhalb der Hauptdiagonalen nur noch Nullen zu haben, subtrahiere ich die erste von der zweiten Zeile. Das ergibt:

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

Nun addiere ich die zweite Zeile auf die dritte und erhalte:

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

Damit habe ich schon mal eine Nullzeile gefunden, die mich im Folgenden nicht mehr belasten muss. Ich möchte aber gerne eine möglichst große Einheitsmatrix haben und werde mich dazu auf zwei Zeilen und zwei Spalten beschränken müssen. Die Addition der zweiten Zeile auf die erste und anschließende Multiplikation der zweiten Zeile mit -1 führt schließlich zu:

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

Übersetzt man das wieder in zwei Gleichungen, so lauten sie

$$c_1 + 3c_3 = 0 \text{ und } c_2 - 2c_3 = 0.$$

Man kann also die Unbekannte c_3 beliebig wählen und erhält

$$c_1 = -3c_3, c_2 = 2c_3, c_3 \in \mathbb{R} \text{ beliebig.}$$

So ist zum Beispiel $c_1 = -3, c_2 = 2, c_3 = 1$ eine Lösung, die offenbar nicht nur aus Nullen besteht. Die drei Vektoren sind deshalb linear abhängig. ■

Die Vektoren aus Beispiel 5.11b) können Sie in der folgenden Übungsaufgabe selbst auf lineare Abhängigkeit testen.

Übungsaufgabe 5.12:
Stellen Sie fest, ob die im Beispiel 5.11b) aufgeführten Vektoren linear abhängig oder linear unabhängig sind. ■

Ein Problem ist damit erledigt. Ein weiteres Problem trat im zweiten Abschnitt auf, als es darum ging, die Inverse einer gegebenen Matrix auszurechnen, denn auch das hatte etwas mit linearen Gleichungssystemen zu tun. Wie so oft, werde ich Ihnen auch diesmal die Sache anhand eines Beispiels erklären. Vorher sollte ich Sie aber kurz daran erinnern, worum es eigentlich geht: Hat man eine quadratische Matrix A gegeben, also eine Matrix, die genauso viele Zeilen wie Spalten hat, dann sucht man nach einer Matrix A^{-1}, die sozusagen dem Kehrwert dieser Matrix A entspricht. Genauer gesagt soll die Gleichung $A \cdot A^{-1} = I_n$ erfüllt sein, wobei man unter I_n die n-reihige Einheitsmatrix versteht, die in der Hauptdiagonalen aus Einsen und ansonsten nur aus Nullen besteht. Das alles hatten wir im zweiten Abschnitt besprochen. Wie man nun an die Inverse A^{-1} herankommt, werden Sie im Verlauf des Beispiels 5.29 sehen.

Beispiel 5.29:
Ich suche nach der Inversen der Matrix $A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$, wobei ich mich zunächst einmal der Hoffnung hingebe, dass eine inverse Matrix überhaupt existiert. Wenn es sie denn geben sollte, dann hat sie das gleiche Format wie A selbst, besteht also aus zwei Zeilen und zwei Spalten. Da ich noch keine Ahnung habe, was in der inversen Matrix stehen könnte, schreibe ich sie auf als $A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Nun weiß ich aber, dass $A \cdot A^{-1}$ genau die Einheitsmatrix ergeben soll, denn so war die Inverse definiert. Und ich weiß auch, wie die zweireihige Einheitsmatrix aussieht, nämlich $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Also muss gelten:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = A \cdot A^{-1} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+2c & b+2d \\ a+3c & b+3d \end{pmatrix}.$$

Was in aller Welt soll mir das jetzt helfen? Eine ganze Menge, denn ich habe herausgefunden, dass zwei Matrizen gleich sein sollen, nämlich

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a+2c & b+2d \\ a+3c & b+3d \end{pmatrix}.$$

Wenn aber die beiden Matrizen gleich sein sollen, dann müssen ihre Einträge gleich sein, und das bedeutet:

$$\begin{array}{rcrl} a & + & 2c & = 1 \\ a & + & 3c & = 0 \end{array} \quad \text{und} \quad \begin{array}{rcrl} b & + & 2d & = 0 \\ b & + & 3d & = 1. \end{array}$$

Das ist praktisch, denn jetzt habe ich zwei lineare Gleichungssysteme bekommen, das eine mit den Unbekannten a und c, das andere mit den Unbekannten b und d.

Wie man so etwas löst, wissen Sie, und ich werde jetzt das erste der beiden Systeme lösen. Seine Matrixform lautet:

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 0 \end{pmatrix}.$$

Abziehen der ersten Zeile von der zweiten liefert

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

Jetzt ziehe ich die doppelte zweite Zeile von der ersten ab. Das führt zu:

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

Folglich ist $a = 3$ und $c = -1$, und schon habe ich zwei Einträge der inversen Matrix berechnet. Auf die gleiche Weise bekomme ich natürlich b und d. Die Matrixform des Gleichungssystems lautet:

$$\begin{pmatrix} 1 & 2 & 0 \\ 1 & 3 & 1 \end{pmatrix}.$$

Abziehen der ersten Zeile von der zweiten liefert

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Jetzt ziehe ich die doppelte zweite Zeile von der ersten ab. Das führt zu:

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

Folglich ist $b = -2$ und $d = 1$. Damit ist schon alles erledigt und ich muss nur noch die inverse Matrix aufschreiben. Sie lautet

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

Eigentlich bin ich jetzt fertig, aber bei der Rechnung sollte Ihnen etwas aufgefallen sein: Egal ob ich a und c oder b und d ausgerechnet habe, die Koeffizientenmatrix war in jedem Fall die gleiche. Auch die Operationen waren jeweils die gleichen, was Sie schon daran sehen können, dass ich für beide Gleichungssysteme die gleichen Zwischentexte geschrieben habe. Nur die rechten Seiten waren natürlich verschieden. Wenn also die Koeffizientenmatrix immer die gleiche ist und grundsätzlich die gleichen Operationen durchzuführen sind, dann kann man sich doch das Leben erleichtern und gleich alles auf einmal machen, indem man beide rechte Seiten auf einmal behandelt und nicht nur eine. Die Startmatrix sieht dann so aus:

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 1 & 3 & 0 & 1 \end{pmatrix}.$$

Abziehen der ersten Zeile von der zweiten liefert

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

Jetzt ziehe ich die doppelte zweite Zeile von der ersten ab. Das führt zu:

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

Schon wieder habe ich das Gleiche gemacht, nur eben mit zwei rechten Seiten auf einmal. Und was stellen Sie fest? In der rechten Hälfte der letzten Matrix steht genau die inverse Matrix! In einem Arbeitsgang habe ich also die inverse Matrix berechnet, für die ich vorher zwei Arbeitsgänge gebraucht habe. ■

Auf diese Weise funktioniert das Invertieren jeder Matrix, sofern sie überhaupt eine Inverse besitzt. Man schreibt die Matrix in die linke Hälfte einer großen Matrix, schreibt die Einheitsmatrix in die rechte Hälfte und manipuliert die gesamte Matrix mit den vertrauten Methoden des Gauß-Algorithmus so lange, bis links die Einheitsmatrix erschienen ist. Sobald Sie diesen angenehmen Zustand erreicht haben, finden Sie in der rechten Hälfte die gesuchte inverse Matrix. Und es kommt sogar noch besser: Schließlich gibt es ja Matrizen, die keine Inverse haben, was ist denn mit denen? Ganz einfach. Falls die Matrix nicht invertierbar ist, wird es Ihnen beim besten Willen nicht gelingen, auf der linken Seite der großen Matrix eine komplette Einheitsmatrix zu zaubern. Der Gauß-Algorithmus ist also gleichzeitig eine Möglichkeit, die Invertierbarkeit einer Matrix festzustellen: Wenn die Rechnung nicht zum Ende führt, kann es keine Inverse geben.

Invertierung von Matrizen

Es sei $A \in \mathbb{R}^{n\times n}$, eine Matrix, deren inverse Matrix A^{-1} berechnet werden soll. Man fasst A und die Einheitsmatrix I_n in einer Matrix zusammen, indem man in die linke Hälfte die Matrix A und in die rechte Hälfte die Einheitsmatrix schreibt, und formt diese Matrix mithilfe der bekannten Operationen aus dem Gauß-Algorithmus so um, dass die linke Hälfte in die Einheitsmatrix I_n verwandelt wird. Dann steht die inverse Matrix A^{-1} in der rechten Hälfte der umgeformten großen Matrix. Falls es nicht möglich ist, in der linken Hälfte die Einheitsmatrix herzustellen, ist A nicht invertierbar.

Mit einem letzten Beispiel schließe ich diesen Abschnitt ab.

Beispiel 5.30:
Gesucht ist die Inverse der Matrix

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

Wie gewohnt schreibe ich die dreireihige Einheitsmatrix I_3 neben A in eine große Matrix mit drei Zeilen und sechs Spalten:

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

Diese neue Matrix muss so umgeformt werden, dass in ihrer linken Hälfte die Einheitsmatrix I_3 steht, denn sobald das geschafft ist, habe ich in der rechten Hälfte die Inverse A^{-1} vor mir, sofern alle Umformungen gut gehen. Nun wird die verdreifachte erste Zeile von der zweiten Zeile abgezogen und anschließend die erste Zeile von der dritten. Das Resultat ist die Matrix

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

Die verdoppelte zweite Zeile wird von der dritten abgezogen. Das ergibt:

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

Die dritte Zeile wird mit -1 multipliziert und danach die neue dritte Zeile auf die erste Zeile addiert. Das ergibt dann die Matrix:

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

Und schon sehen Sie die inverse Matrix vor sich, denn links steht jetzt die Einheitsmatrix, sodass rechts ohne jede Frage die inverse Matrix von A aufgeführt wird. Sie lautet also:

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

■

Übungsaufgabe 5.13:
Berechnen Sie die Inverse der Matrix

$$A = \begin{pmatrix} 1 & 0 & -1 \\ -8 & 4 & 1 \\ -2 & 1 & 0 \end{pmatrix}$$

aus Beispiel 5.21. ■

Nun haben wir erstens alles Nötige über lineare Gleichungssysteme besprochen und zweitens in einem Aufwasch auch noch die offen gebliebenen Punkte aus früheren Abschnitten erledigt. Im nächsten Abschnitt werde ich Ihnen zeigen, wie man Vektoren und lineare Gleichungssysteme auf das eine oder andere geometrische Problem anwenden kann.

5.4 Analytische Geometrie

Erschrecken Sie nicht vor dem etwas hochtrabenden Titel, ich will hier nur einen sehr kleinen Teil der analytischen Geometrie vorstellen, um die bisher gelernten Techniken ein wenig zu illustrieren. Während die allgemeine Vektoren- und Matrizenrechnung fast jedem einmal begegnen kann, vom Betriebswirt bis zum Ingenieur, vom Psychologen bis zum Physiker, wendet sich die eher geometrische Interpretation der Vektorrechnung, über die ich jetzt reden will, mehr an die technisch-naturwissenschaftlich orientierten Leser. Zunächst muss ich einiges über die Darstellung von Vektoren im Koordinatensystem sagen und dann werde ich über die Gleichung einer Ebene im Raum sprechen.

Vektoren habe ich bisher einfach nur als Ansammlung von Zahlen betrachtet und dabei völlig ignoriert, dass man sie auch geometrisch interpretieren kann, sofern man sich auf zwei- und dreidimensionale Vektoren beschränkt. Die Idee ist einfach genug und ich werde sie Ihnen am Beispiel der zweidimensionalen Vektoren kurz zeigen. Werfen Sie einmal einen Blick auf die Abbildung 5.1.

Ich habe hier zwei Punkte A und B in ein zweidimensionales Koordinatensystem eingetragen, wobei Sie solche Koordinatensysteme schon aus dem Kapitel über Funktionen kennen. Bezeichnet man wie üblich die waagrechte Richtung als x-Richtung und die senkrechte als y-Richtung, dann hat A offenbar die **Koordinaten** $A = (1, 2)$

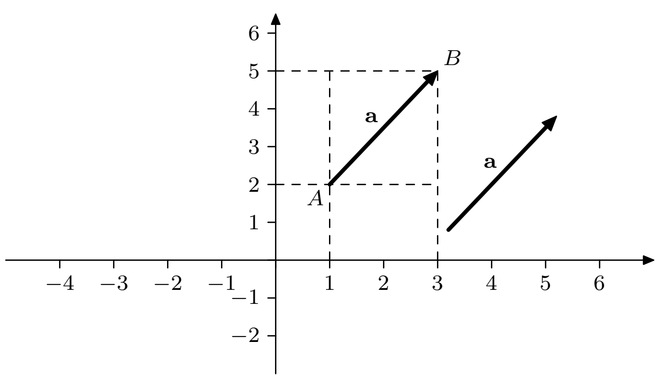

Abb. 5.1 Vektor in der Ebene

und B die Koordinaten $B = (3, 5)$. Nun kann man sich dafür interessieren, wie man auf dem schnellsten Weg von A nach B kommt, und die kürzeste Verbindung zwischen den beiden Punkten ist die in der Abbildung eingetragene Strecke. Ich hatte aber nach dem Weg von A nach B gefragt und nicht etwa umgekehrt von B nach A, die Strecke muss also eine **Richtung** haben, und die wird durch den kleinen Pfeil, der in Richtung B zeigt, symbolisiert. Die Strecke hat daher eine Richtung und natürlich auch eine **Länge**, denn sie soll schließlich nicht über B hinauszeigen.

Was hat das nun mit den altbekannten zweidimensionalen Vektoren zu tun? Ganz einfach: Die eben besprochene gerichtete Strecke $\mathbf{a}$ können Sie mithilfe eines Vektors beschreiben. Um von A nach B zu kommen, muss ich doch zwei Schritte in die waagrechte Richtung gehen und drei Schritte in die senkrechte. Damit sind Richtung und Länge der gerichteten Strecke $\mathbf{a}$ vollständig beschrieben, denn wenn ich bei A starte, zwei Schritte nach rechts und drei Schritte nach oben gehe, bin ich genau da gelandet, wo mich meine gerichtete Strecke hinführen soll. Die waagrechte und die senkrechte Koordinate beschreiben also meine gerichtete Strecke, und daher setzt man $\mathbf{a} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, womit nur gemeint ist, dass man zwei Schritte nach rechts laufen muss und drei Schritte nach oben. Es ist dabei völlig unerheblich, von wo aus ich mit dieser Lauferei anfange. Durch die Koordinatenschreibweise wird nur angegeben, in welche Richtungen man laufen muss, nicht aber, wo der Startpunkt ist. Deshalb kann ich $\mathbf{a}$ beliebig durch die Gegend schieben; solange ich nichts an Richtung und Länge ändere, also nur Parallelverschiebungen vornehme, bleibt es der gleiche Vektor, was Sie auch in Abbildung 5.1 erkennen können.

Damit haben wir schon geklärt, wie man einen zweidimensionalen Vektor anschaulich interpretieren kann: als gerichtete Strecke in der Ebene, wobei die erste Komponente angibt, wie weit man sich in horizontaler Richtung bewegt, und die zweite Komponente die Anzahl der Schritte in die vertikale Richtung verrät. Der Startpunkt ist dabei egal, nur die Angabe der Schrittzahlen ist wichtig.

Zweidimensionale Vektoren

Ein zweidimensionaler Vektor $\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ kann in ein zweidimensionales Koordinatensystem eingezeichnet werden, indem man von einem beliebigen Startpunkt aus sich um a_1 Schritte in die waagrechte Richtung bewegt und um a_2 Schritte in die senkrechte. Der Vektor ist dann eine gerichtete Strecke vom gewählten Anfangspunkt bis zum erreichten Endpunkt.

Sind umgekehrt zwei Punkte mit den Koordinaten $A = (x_A, y_A)$ und $B = (x_B, y_B)$ gegeben, so wird der **Verbindungsvektor** $\overrightarrow{AB}$ zwischen den beiden Punkten berechnet durch $\overrightarrow{AB} = \begin{pmatrix} x_B - x_A \\ y_B - y_A \end{pmatrix}$.

In Abbildung 5.1 war beispielsweise $A = (1,2)$ und $B = (3,5)$. Daher gilt $\overrightarrow{AB} = \begin{pmatrix} 3-1 \\ 5-2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$, genau wie Sie es vorhin schon gesehen hatten. Übrigens gibt es natürlich einen ganz besonderen Anfangspunkt im Koordinatensystem, mit dem sich am leichtesten rechnen lässt, und das ist der Nullpunkt $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Ist nämlich irgendein Punkt $A = (x_A, y_A)$ gegeben und sucht man den Verbindungsvektor zwischen dem Nullpunkt und A, so lautet er $\overrightarrow{\mathbf{0}A} = \begin{pmatrix} x_A - 0 \\ y_A - 0 \end{pmatrix} = \begin{pmatrix} x_A \\ y_A \end{pmatrix}$. Der Vektor vom Nullpunkt zum Punkt A hat also die gleichen Koordinaten wie der Punkt A selbst, und er gibt an, wie man vom Ursprung des Koordinatensystems aus laufen muss, um an den gewünschten Ort A zu kommen. Deshalb nennt man ihn auch den **Ortsvektor** von A.

Für den Moment war das genug über zweidimensionale Vektoren, um die dreidimensionalen muss ich mich ja auch noch kümmern. Das ist jetzt aber kein Problem, denn da geht alles genauso wie bei den zweidimensionalen, nur eben mit einer Dimension mehr. Hat man nämlich ein dreidimensionales Koordinatensystem, so haben die Punkte auch drei Koordinaten und nicht mehr nur zwei, und um festzustellen, wie ich beispielsweise vom Punkt $A = (1, 2, 3)$ zum Punkt $B = (-1, 3, 1)$ komme, muss ich wieder nur die einzelnen Differenzen ausrechnen. In der ersten Komponente muss ich zwei Schritte in die negative Richtung gehen, in der zweiten einen Schritt in die positive und in der dritten zwei Schritte in die negative: das ist das gleiche Spiel wie eben mit dem einzigen Unterschied, dass noch eine räumliche Komponente dazu kommt. Ich werde deshalb auch gleich alles Nötige zusammen fassen.

Dreidimensionale Vektoren

Ein dreidimensionaler Vektor $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ kann in ein dreidimensionales Koordinatensystem eingezeichnet werden, indem man von einem beliebigen Startpunkt aus sich um a_1 Schritte in die erste Richtung bewegt, um a_2 Schritte in die zweite Richtung und um a_3 Schritte in die dritte. Der Vektor ist dann eine gerichtete Strecke vom gewählten Anfangspunkt bis zum erreichten Endpunkt.

Sind umgekehrt zwei Punkte mit den Koordinaten $A = (x_A, y_A, z_A)$ und $B = (x_B, y_B, z_B)$ gegeben, so wird der **Verbindungsvektor** $\overrightarrow{AB}$ zwischen den beiden Punkten berechnet durch $\overrightarrow{AB} = \begin{pmatrix} x_B - x_A \\ y_B - y_A \\ z_B - z_A \end{pmatrix}$.

Bis auf die Anzahl der Komponenten hat sich also nichts geändert. Kurz zusammengefasst kann man sagen, dass ein zweidimensionaler Vektor eine gerichtete Strecke in der Ebene, ein dreidimensionaler Vektor dagegen eine gerichtete Strecke im Raum ist. Der Ortsvektor eines zwei- oder dreidimensionalen Punktes ist dann der Vektor, der vom Nullpunkt zu dem gewählten Punkt zeigt.

Dazu wieder die eine oder andere Übung:

Übungsaufgabe 5.14:

a) Berechnen Sie die Vektoren zwischen den Punkten

$$A = (2, 3, 1) \text{ und } B = (-1, 0, 3) \text{ bzw. zwischen } C = (8, 17) \text{ und } D = (9, 16).$$

b) Wie lauten die Ortsvektoren der Punkte $A = (3, 6, -1)$ und $B = (4, 7, 9)$? ■

Nun habe ich ja mit den Vektoren nicht nur herumgespielt, sondern sehr konkret gerechnet, und ich sollte Ihnen zeigen, wie sich diese Rechenoperationen auf die Darstellung im Koordinatensystem auswirkt. Fangen wir damit an, dass aus einem Vektor $\mathbf{a}$ der Vektor $-\mathbf{a}$ berechnet wird. Während $\mathbf{a}$ beispielsweise die Punkte A und B verbindet, sollte $-\mathbf{a}$ genau in die umgekehrte Richtung zeigen und den Punkt B mit dem Punkt A verbinden, und das ist auch der Fall: Man erhält $-\mathbf{a}$ aus $\mathbf{a}$, indem man die Richtung von $\mathbf{a}$ umdreht, also sozusagen den Vektorpfeil einfach an der anderen Seite des Vektors anbringt.

Negativer Vektor

Ist $\mathbf{a}$ ein zwei- oder dreidimensionaler Vektor, dann ist $-\mathbf{a}$ der Vektor gleicher Länge, der genau in die entgegengesetzte Richtung zeigt. Ist also $\mathbf{a}$ der Vektor zwischen den Punkten A und B, dann ist $-\mathbf{a}$ der Vektor zwischen B und A.

Ich hatte aber nicht nur negative Vektoren ausgerechnet, sondern beliebige Vektoren mit beliebigen reellen Zahlen multipliziert. Auch das ist im Koordinatensystem kein Problem. Nehmen Sie beispielsweise den Vektor $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, der vom Nullpunkt zum Punkt $(1, 2)$ zeigt. Wie Sie im ersten Abschnitt gelernt haben, ist dann $2 \cdot \mathbf{a} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$, und das ist der Vektor, der vom Nullpunkt zum Punkt $(2, 4)$ zeigt. Anstatt einen Schritt nach rechts und zwei Schritte nach oben zu gehen, werde ich jetzt also zwei Schritte nach rechts und vier Schritte nach oben gehen: Alles hat sich verdoppelt, ist ja auch kein Wunder, wenn man mit zwei multipliziert. Und da sich alles verdoppelt hat, muss ich den im Koordinatensystem eingezeichneten Vektor auch nur verdoppeln, indem ich ihn einfach in die gleiche Richtung zeigen lasse, aber mit dem Faktor 2 strecke.

Multiplikation eines Vektors mit einer Zahl

Ist $\mathbf{a}$ ein zwei- oder dreidimensionaler Vektor und λ eine positive reelle Zahl, dann ist $\lambda \cdot \mathbf{a}$ der Vektor, der in die gleiche Richtung wie $\mathbf{a}$ zeigt, aber mit dem Faktor λ gestreckt ist. Ist λ eine negative reelle Zahl, dann ist $\lambda \cdot \mathbf{a}$ der Vektor, der in die entgegengesetzte Richtung wie $\mathbf{a}$ zeigt und mit dem Faktor $|\lambda|$ gestreckt ist.

Bei einem positiven Faktor wird also in die gegebene Richtung gestreckt, bei einem negativen in die entgegengesetzte, und mehr steckt nicht dahinter. Auch die nächste Operation, die Addition von zwei Vektoren, ist nichts wirklich Aufregendes, wie Sie sofort an der Abbildung 5.2 erkennen können.

Der Vektor $\mathbf{a}$ zeigt hier von $A = (1, 2)$ nach $B = (3, 5)$, hat also die Koordinaten $\mathbf{a} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$. Dagegen ist $\mathbf{b}$ der Vektor, der von $B = (3, 5)$ nach $C = (4, 1)$ zeigt und deshalb mit den Koordinaten $\mathbf{b} = \begin{pmatrix} 1 \\ -4 \end{pmatrix}$ beschrieben werden kann. Nach den Regeln aus Abschnitt 1 ist dann $\mathbf{a} + \mathbf{b} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ -4 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$. Wenn ich also den Vektor $\mathbf{a} + \mathbf{b}$ an den Punkt A ansetze, dann muss ich drei Schritte nach rechts und einen Schritt nach unten gehen, um seinen Endpunkt zu erreichen, und Sie sehen, wo ich dann ankomme: genau im Punkt C.

So geht das immer. Sobald Sie zwei zweidimensionale oder zwei dreidimensionale Vektoren addieren wollen, zeichnen Sie den ersten in ein Koordinatensystem ein, hängen den zweiten an den ersten an, und die Summe der beiden ist dann der Vektor, der vom Anfangspunkt des ersten Vektors zum Endpunkt des zweiten Vektors zeigt.

Addition von Vektoren

Gegeben seien zwei zweidimensionale oder zwei dreidimensionale Vektoren $\mathbf{a}$ und $\mathbf{b}$. Zeichnet man $\mathbf{a}$ und $\mathbf{b}$ so in ein Koordinatensystem ein, dass der Anfangspunkt von $\mathbf{b}$ mit dem Endpunkt von $\mathbf{a}$ übereinstimmt, dann ist $\mathbf{a} + \mathbf{b}$ der Vektor, der vom Anfangspunkt des Vektors $\mathbf{a}$ zum Endpunkt des Vektors $\mathbf{b}$ zeigt.

Ein Beispiel dafür sehen Sie in Abbildung 5.2 vor sich. Bedenken Sie übrigens, dass es bei Vektoren völlig egal ist, wo Sie anfangen, nur ihre Richtung und ihre Länge ist entscheidend. Deshalb kann ich den Vektor $\mathbf{b}$ ohne Probleme an das Ende von $\mathbf{a}$ anhängen.

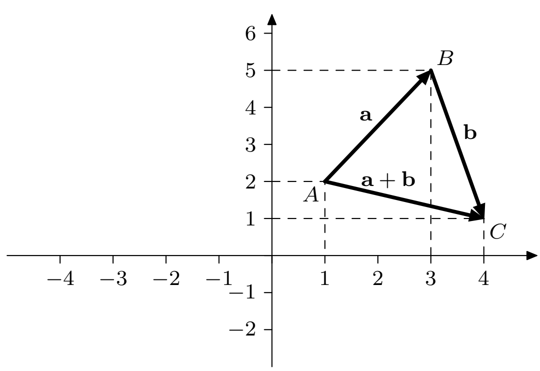

Abb. 5.2 Addition von zwei Vektoren

Hier noch etwas zum Üben:

Übungsaufgabe 5.15:

Zeichnen Sie die Vektoren $\mathbf{a} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \mathbf{c} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ in ein Koordinatensystem und bestimmen Sie die Vektoren $\mathbf{a} + 2\mathbf{b} - \mathbf{c}$ und $3\mathbf{a} + \mathbf{b} + \mathbf{c}$. ■

Jetzt habe ich genug Material zusammengetragen, um das eigentliche Thema dieses Abschnitts anzugehen: die Gleichung einer Ebene im Raum. Ich gehe davon aus, dass Sie sich eine Ebene im Raum vorstellen können, Sie müssen dabei nur an ein Blatt Papier denken, das bewegungslos im Raum herumschwebt und nach allen Seiten beliebig weit ausgedehnt ist. Mein Ziel ist es, so eine Ebene mit einer einfachen Gleichung zu beschreiben. Nun haben Sie ja im Kapitel über Funktionen sich zwar noch nicht mit Ebenen herumgeschlagen, aber doch immerhin mit Geraden, die man normalerweise durch Gleichungen der Form $y = ax + b$ beschreibt. Diese Geraden lagen alle innerhalb eines zweidimensionalen Koordinatensystems, also innerhalb einer Ebene. Jetzt gehe ich einen Schritt weiter und betrachte einen Raum, also ein dreidimensionales Koordinatensystem, in das ich eine Ebene legen will – die Situation ist also recht ähnlich, und vielleicht habe ich Glück und finde auch eine ähnliche Gleichung. Tatsächlich wird auch genau das passieren. Bezeichnet man die drei Raumkoordinaten mit x, y und z, so haben Ebenengleichungen immer die Form $4x - 3y + 2z = 17$ oder $-2x + 4y + z = 9$, also allgemein $ax + by + cz = d$. Darauf will ich hinaus; Sie werden sehen, dass das gar nicht so wild wird.

Ausgangspunkt der ganzen Sache ist die altbekannte Beobachtung, dass ein Tisch mit drei Beinen niemals wackelt, sofern die Beine stabil sind. Anders gesagt: Durch drei Punkte im Raum kann man immer genau eine Ebene legen, wie Sie auch problemlos feststellen können, indem Sie drei Finger hochhalten und ein Blatt Papier darauf legen. Ich gehe also davon aus, dass ich durch drei beliebige Punkte A, B und C, die sich irgendwo in einem dreidimensionalen Koordinatensystem befinden, eine Ebene legen soll. Wie das ungefähr aussieht, zeigt die Abbildung 5.3. Für den Moment interessiert mich an dieser Abbildung nur, dass die Ebene irgendwo im Raum liegt und durch die drei Punkte A, B und C geht. Aber zum Bestimmen einer Gleichung reicht das noch nicht, weshalb ich noch einen beliebigen Ebenenpunkt (x, y, z) und den Nullpunkt eingetragen habe. Wie kommt man nun vom Nullpunkt zum Punkt (x, y, z)? Ganz einfach: Man geht zuerst zum Punkt A, denn der war schließlich gegeben. Und dann hat man ja schon die Ebene erreicht und kann ganz gemütlich von A nach (x, y, z) laufen. Bezeichne ich für den Moment meinen Punkt (x, y, z) mit P, dann erhalte ich also den Ortsvektor von P durch $\overrightarrow{\mathbf{0}P} = \overrightarrow{\mathbf{0}A} + \overrightarrow{AP}$.

Der erste Vektor $\overrightarrow{\mathbf{0}A}$ ist kein Grund zur Aufregung, denn er ist nur der Ortsvektor des Punktes A, den Sie bekommen, indem Sie die Koordinaten von A als Vektor schreiben, also senkrecht anstatt waagrecht. Anders sieht es aus bei $\overrightarrow{AP}$. Er verläuft von A nach P, liegt also komplett in der Ebene, um die es geht, und wie Sie an Abbildung 5.3 sehen können, wird diese Ebene von den beiden Vektoren $\mathbf{b} = \overrightarrow{AB}$ und $\mathbf{c} = \overrightarrow{AC}$ aufgespannt. Wenn Sie nämlich in der Situation von Abbildung 5.3 sowohl

$\mathbf{b}$ als auch $\mathbf{c}$ ein klein wenig mit dem jeweils passenden Faktor stauchen und dann den gestauchten $\mathbf{b}$-Vektor an den gestauchten $\mathbf{c}$-Vektor hängen, dann kommen Sie ebenfalls von A nach P. Mit anderen Worten: Es gibt reelle Zahlen λ und μ, so dass $\overrightarrow{AP} = \lambda \cdot \mathbf{b} + \mu \cdot \mathbf{c}$ gilt. Man kann den Vektor von A nach P aus den beiden gegebenen Vektoren $\mathbf{b}$ und $\mathbf{c}$ kombinieren, weil alle diese Vektoren in derselben Ebene verlaufen und diese Ebene von $\mathbf{b}$ und $\mathbf{c}$ aufgespannt wird.

Setze ich nun diese neue Erkenntnis in die alte Gleichung $\overrightarrow{\mathbf{0}P} = \overrightarrow{\mathbf{0}A} + \overrightarrow{AP}$ ein, so ergibt sich:

$$\overrightarrow{\mathbf{0}P} = \overrightarrow{\mathbf{0}A} + \lambda \cdot \mathbf{b} + \mu \cdot \mathbf{c} = \mathbf{a} + \lambda \cdot \mathbf{b} + \mu \cdot \mathbf{c},$$

denn den Ortsvektor von A hatte ich mit $\mathbf{a}$ bezeichnet. Und da der Punkt P gerade die Koordinaten (x, y, z) hat und es sich bei $\overrightarrow{\mathbf{0}P}$ um den Ortsvektor von P handelt, haben wir

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \overrightarrow{\mathbf{0}A} + \lambda \cdot \mathbf{b} + \mu \cdot \mathbf{c} = \mathbf{a} + \lambda \cdot \mathbf{b} + \mu \cdot \mathbf{c}$$

für jeden Punkt (x, y, z) auf der Ebene. Die reellen Zahlen λ und μ bezeichnet man dabei als **Parameter**, und deshalb heißt diese Form der Ebenengleichung die **parametrisierte Form** der Ebenengleichung.

Sehen wir uns einmal ein Beispiel dazu an.

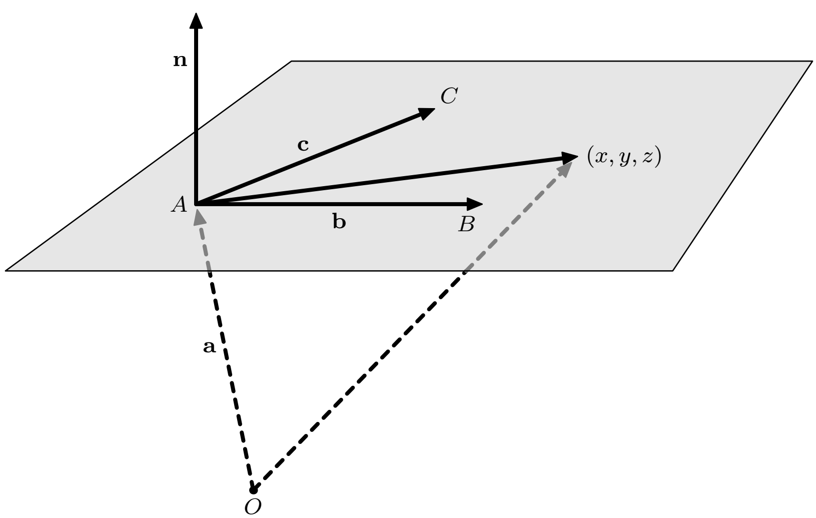

Abb. 5.3 Ebene im Raum

Beispiel 5.31:
Es geht um die Ebene durch die Punkte $A = (1, 2, 3), B = (0, 1, 2)$ und $C = (-1, 2, 1)$. Dann ist natürlich $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, und für die Vektoren $\mathbf{b}$ und $\mathbf{c}$ gilt, da sie den Punkt A mit den Punkten B bzw. C verbinden:

$$\mathbf{b} = \begin{pmatrix} 0-1 \\ 1-2 \\ 2-3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} \text{ und } \mathbf{c} = \begin{pmatrix} -1-1 \\ 2-2 \\ 1-3 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix}.$$

Also lautet die Gleichung der Ebene:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \lambda \cdot \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} + \mu \cdot \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix} \text{ mit } \lambda, \mu \in \mathbb{R}.$$

Nimmt man zum Beispiel $\lambda = \mu = 0$, so erhält man den Ebenenpunkt $(1, 2, 3)$, was nicht sehr überraschend ist. Sie können aber jede beliebige Kombination von Zahlen einsetzen und werden immer wieder einen Punkt auf der Ebene erreichen; für $\lambda = -1$ und $\mu = 1$ findet man zum Beispiel den Ortsvektor $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} + \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}$, woraus sofort folgt, dass auch der Punkt $(0, 3, 2)$ auf der Ebene liegt. ■

Als kleine Auflockerung wieder mal eine Übungsaufgabe.

Übungsaufgabe 5.16:
Berechnen Sie die parametrisierte Ebenengleichung der Ebene durch die drei Punkte $A = (2, 1, 3), B = (-1, 0, 5)$ und $C = (1, 1, 1)$. ■

Das ist schon mal ganz gut, reicht aber noch nicht, denn eigentlich wollte ich ja eine andere Form der Ebenengleichung erreichen. Das haben wir gleich. Eine entscheidende Rolle wird dabei der Vektor $\mathbf{n}$ spielen, den ich vorsorglich in Abbildung 5.3 eingezeichnet und bisher völlig ignoriert hatte. Er hat nämlich die angenehme Eigenschaft, senkrecht auf der Ebene zu stehen, weshalb ich Ihnen jetzt erst einmal zeige, wie man so etwas wie „senkrecht stehen“ nachrechnet.

Beispiel 5.32:
Werfen Sie einen kurzen Blick auf die Abbildung 5.4. Durch schlichtes Abzeichnen und Nachmessen können Sie feststellen, dass die Vektoren $\mathbf{a} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ und $\mathbf{b} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$ senkrecht aufeinander stehen, dass also ein Winkel von $90°$ zwischen ihnen liegt.

Berechnet man nun das Skalarprodukt beider Vektoren, so findet man:

$$\mathbf{a} \cdot \mathbf{b} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 2 \end{pmatrix} = 2 \cdot (-3) + 3 \cdot 2 = 0.$$

Zufall? Nein. Schicksal? Ja, denn genauso ist das immer: Wann immer zwei Vektoren senkrecht aufeinander stehen, wird ihr Skalarprodukt gleich null sein und umgekehrt. ■

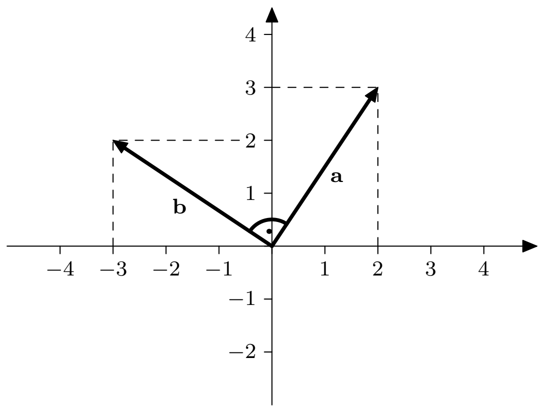

Abb. 5.4 Senkrecht stehende Vektoren

Senkrecht stehende Vektoren

Genau dann stehen zwei zweidimensionale oder zwei dreidimensionale Vektoren $\mathbf{a}$ und $\mathbf{b}$ senkrecht aufeinander, wenn ihr Skalarprodukt null ergibt, das heißt, wenn $\mathbf{a} \cdot \mathbf{b} = 0$ gilt.

Was hat das nun mit meiner Ebene zu tun? Der Vektor $\mathbf{n}$ aus Abbildung 5.3 steht senkrecht auf den beiden Vektoren $\mathbf{a}$ und $\mathbf{b}$, und da diese beiden Vektoren die ganze Ebene aufspannen, steht er gleichzeitig senkrecht auf der ganzen Ebene. Das wird sich gleich als sehr praktisch erweisen, aber für den Moment muss ich noch kurz der Frage nachgehen, wie man so einen Vektor $\mathbf{n}$ ausrechnen kann:

Beispiel 5.33:
Ich verwende wieder die Situation aus Beispiel 5.31; die Vektoren, auf denen $\mathbf{n}$ senkrecht stehen soll, heißen also $\mathbf{b} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$ und $\mathbf{c} = \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix}$. Bezeichne ich das

ominöse $\mathbf{n}$ mit $\mathbf{n} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$, so muss nach dem eben ermittelten Prinzip das Skalarprodukt von $\mathbf{n}$ mit $\mathbf{b}$ und auch das Skalarprodukt von $\mathbf{n}$ mit $\mathbf{c}$ zu null werden. Also muss gelten:

$$0 = \mathbf{n} \cdot \mathbf{b} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} = -n_1 - n_2 - n_3$$

und

$$0 = \mathbf{n} \cdot \mathbf{c} = \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix} = -2n_1 - 2n_3.$$

Ich habe somit zwei Gleichungen erhalten, nämlich $-n_1 - n_2 - n_3 = 0$ und $-2n_1 - 2n_3 = 0$. Wie nennt man so etwas? Richtig, das ist ein lineares Gleichungssystem aus zwei Gleichungen mit den drei Unbekannten n_1, n_2 und n_3, und ich werde es jetzt auf die im dritten Abschnitt dieses Kapitels besprochene Weise lösen. Multipliziert man der Einfachheit halber beide Gleichungen sofort mit -1, so lautet die Matrixform des Gleichungssystems

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 2 & 0 & 2 & 0 \end{pmatrix}.$$

Ich ziehe die doppelte erste Zeile von der zweiten ab, teile dann die zweite durch -2 und ziehe anschließend die zweite von der ersten ab. Das ergibt die Matrizenfolge:

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & -2 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Übersetzt man das wieder in Gleichungen, so heißt das $n_1 + n_3 = 0$ und $n_2 = 0$. Die erreichte kleine Einheitsmatrix bezieht sich auf die Unbekannten n_1 und n_2, weshalb die Unbekannte n_3 frei wählbar ist und die anderen beiden berechnet werden können – in diesem Fall sogar sehr übersichtlich, denn schließlich ist n_2 sogar gleich null. Ich habe also die Lösungen:

$$n_1 = -n_3, n_2 = 0, n_3 \in \mathbb{R} \text{ ist frei wählbar.}$$

Das Ziel meiner Rechnung war aber, einen Vektor herauszubekommen, der auf den beiden Vektoren $\mathbf{b}$ und $\mathbf{c}$ senkrecht steht. Und davon gibt es tatsächlich eine ganze Menge, denn wie ich gerade ausgerechnet habe, macht das jeder Vektor der Form $\begin{pmatrix} -n_3 \\ 0 \\ n_3 \end{pmatrix}$, also beispielsweise $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ oder auch $\begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}$. ■

Genau wie in Beispiel 5.33 können Sie immer einen Vektor ausrechnen, der auf zwei gegebenen Vektoren und damit auch auf der von ihnen aufgespannten Ebene senkrecht steht, indem Sie das richtige lineare Gleichungssystem lösen. Sie können es aber auch lassen, denn im Prinzip ist es ja immer die gleiche Rechnung, und deshalb hat sich schon früher jemand eine allgemeine Formel überlegt, in die man nur noch einsetzen muss, sodass man nicht immer und immer wieder das grundsätzlich gleiche System lösen muss. Was dabei herauskommt, nennt man das **Vektorprodukt** oder auch **Kreuzprodukt** zweier Vektoren.

Vektorprodukt

Sind $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ und $\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ zwei dreidimensionale Vektoren, so nennt man den Vektor

$$\mathbf{b} \times \mathbf{c} = \begin{pmatrix} b_2c_3 - b_3c_2 \\ b_3c_1 - b_1c_3 \\ b_1c_2 - b_2c_1 \end{pmatrix}$$

das Vektorprodukt oder auch Kreuzprodukt von $\mathbf{b}$ und $\mathbf{c}$. Der Vektor $\mathbf{b} \times \mathbf{c}$ steht senkrecht auf $\mathbf{b}$ und auf $\mathbf{c}$.

Schön ist das nicht, aber praktisch. Sie können sich nämlich das Lösen des Gleichungssystems aus Beispiel 5.33 ersparen und gleich ausrechnen:

$$\begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} \times \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} (-1)\cdot(-2) - (-1)\cdot 0 \\ (-1)\cdot(-2) - (-1)\cdot(-2) \\ (-1)\cdot 0 - (-1)\cdot(-2) \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix},$$

und das ist tatsächlich eine der Lösungen des Gleichungssystems, die ich oben ausgerechnet hatte.

Auch dieses neue Produkt können Sie jetzt ein wenig einüben:

Übungsaufgabe 5.17:

Rechnen Sie nach, dass das Vektorprodukt $\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} \times \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}$ senkrecht auf den Vektoren $\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}$ und $\begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}$ steht. ■

Jetzt habe ich alles zusammen, um kurz und schmerzlos die gesuchte Ebenengleichung auszurechnen. Nur noch ein Begriff, dann lege ich los: Einen Vektor, der auf einer Ebene senkrecht steht, nennt man auch einen **Normalenvektor** der Ebene, und das erklärt, warum ich den senkrecht stehenden Vektor mit $\mathbf{n}$ bezeichne. Anhand meines alten Beispiels zeige ich Ihnen jetzt, wie all die besprochenen Einzelheiten zusammenpassen:

Beispiel 5.34:
Wie schon in Beispiel 5.31 geht es um die Ebene durch die Punkte $A = (1,2,3)$, $B = (0,1,2)$ und $C = (-1,2,1)$; vielleicht sollten Sie noch einmal zur Orientierung einen Blick auf die Abbildung 5.3 werfen. Die Vektoren **b** und **c** hatte ich schon in Beispiel 5.31 berechnet als

$$\mathbf{b} = \begin{pmatrix} 0-1 \\ 1-2 \\ 2-3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} \text{ und } \mathbf{c} = \begin{pmatrix} -1-1 \\ 2-2 \\ 1-3 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ -2 \end{pmatrix}.$$

Als Vektor **n**, der senkrecht auf **b** und **c** und damit senkrecht auf der ganzen Ebene steht, kann ich den gerade eben bestimmten Normalenvektor

$$\mathbf{n} = \mathbf{b} \times \mathbf{c} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}$$

verwenden. Da aber dieser Vektor **n** senkrecht auf der Ebene steht, wird er sicher senkrecht auf jedem Vektor stehen, der ganz in der betrachteten Ebene liegt, z. B. auf **b** und **c**, aber auch auf dem Vektor, der von A aus zu dem Ebenenpunkt (x,y,z) verläuft. Diesen Vektor kann ich nun leicht ausrechnen, denn A hat die Koordinaten $(1,2,3)$, während (x,y,z) natürlich die Koordinaten (x,y,z) hat und daher der verbindende Vektor $\begin{pmatrix} x-1 \\ y-2 \\ z-3 \end{pmatrix}$ lautet. Auf diesem Vektor soll nun **n** senkrecht stehen, und wie Sie gelernt haben, bedeutet das nur, dass das Skalarprodukt der beiden Vektoren gleich null sein muss. Also weiß ich:

$$0 = \mathbf{n} \cdot \begin{pmatrix} x-1 \\ y-2 \\ z-3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} x-1 \\ y-2 \\ z-3 \end{pmatrix}.$$

Sie wissen aber auch, wie man so ein Skalarprodukt ausrechnet. Es gilt nämlich:

$$\begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \cdot \begin{pmatrix} x-1 \\ y-2 \\ z-3 \end{pmatrix} = 2\cdot(x-1)+0\cdot(y-2)-2\cdot(z-3) = 2x-2-2z+6 = 2x-2z+4.$$

Da das Skalarprodukt gleich null ist, folgt daraus:

$$2x - 2z + 4 = 0, \text{ also } 2x - 2z = -4 \text{ und damit } x - z = -2,$$

wobei ich im letzten Schritt einfach die Gleichung durch zwei geteilt habe.

Wie Sie sehen, bin ich fertig. Die gesuchte Gleichung lautet $x - z = -2$, und das heißt, dass alle Punkte (x,y,z) auf meiner Ebene diese Gleichung erfüllen müssen. Welchen Ebenenpunkt auch immer Sie sich aussuchen: Die Differenz von erster und dritter Koordinate wird auf jeden Fall gleich -2 sein. ■

So geht das immer, und damit Sie das Schema nicht mühselig aus den Beispielen zusammensuchen müssen, schreibe ich es jetzt noch einmal auf:

Berechnung der parameterfreien Ebenengleichung

Sind $A = (a_1, a_2, a_3), B = (b_1, b_2, b_3)$ und $C = (c_1, c_2, c_3)$ drei Raumpunkte, die nicht alle auf einer Geraden liegen, so berechnet man die parameterfreie Gleichung der Ebene, die durch alle drei Punkte geht, nach dem folgenden Schema:

a) Man bestimme

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \\ b_3 - a_3 \end{pmatrix}, \mathbf{c} = \begin{pmatrix} c_1 - a_1 \\ c_2 - a_2 \\ c_3 - a_3 \end{pmatrix}.$$

b) Man berechne mit dem Kreuzprodukt den Normalenvektor

$$\mathbf{n} = \mathbf{b} \times \mathbf{c} = \begin{pmatrix} b_2c_3 - b_3c_2 \\ b_3c_1 - b_1c_3 \\ b_1c_2 - b_2c_1 \end{pmatrix}.$$

c) Für die Ebenenpunkte (x, y, z) mache man den Ansatz

$$\mathbf{n} \cdot \begin{pmatrix} x - a_1 \\ y - a_2 \\ z - a_3 \end{pmatrix} = 0.$$

d) Man berechne das Skalarprodukt aus c) und vereinfache die entstehende Gleichung, bis eine Gleichung der Form $ax + by + cz = d$ entsteht. Diese Gleichung ist dann die parameterfreie Gleichung der Ebene.

Die Gleichung heißt übrigens deshalb parameterfrei, weil in ihr – im Gegensatz zur parametrisierten Gleichung – keine Parameter λ und μ mehr auftreten, sondern eben nur noch die Koordinaten der Punkte (x, y, z) selbst. Und warum habe ich vorausgesetzt, dass die Punkte nicht auf einer Geraden liegen? Ganz einfach: Wenn Sie drei Punkte auf einer geraden Linie haben, dann gibt es natürlich keine eindeutige Ebene durch diese drei Punkte, sondern eine Unmenge davon. Das muss man aber nicht vor dem Beginn der Rechnung testen; falls nämlich alle drei Punkte auf einer Geraden liegen, werden Sie beim Rechnen feststellen, dass das Vektorprodukt $\mathbf{b} \times \mathbf{c}$ genau den Nullvektor ergibt, und mit dem kann man nun mal nicht viel anfangen.

Ihre neuen Kenntnisse können Sie an der folgenden Aufgabe testen.

Übungsaufgabe 5.18:

Bestimmen Sie die parameterfreie Gleichung der Ebene, die durch die drei Punkte $A = (4, 0, 0), B = (5, 1, -1)$ und $C = (-3, 2, 1)$ geht. ■

Damit beende ich die Lineare Algebra; im nächsten Kapitel werden Sie einiges über die Differenzialrechnung erfahren.

6 Differenzial- und Integralrechnung

In Kapitel 2 wurden wichtige mathematische Objekte behandelt: **Funktionen**. Die dort beschriebenen Dinge haben Sie sich inzwischen sicherlich noch einmal etwas verinnerlicht. Das ist auch gut so, denn ich möchte mich jetzt mit den **analytischen** Eigenschaften von reellwertigen Funktionen beschäftigen. Damit meine ich, dass ich gemeinsam mit Ihnen klären möchte, wie man natürliche Fragen der folgenden Art beantworten kann: Wie kann man die Steigung einer Funktion mathematisch beschreiben? An welchen Stellen werden die Werte einer Funktion am größten? Wie kann der Graph einer Funktion gezeichnet werden? Wie groß ist die Fläche, die der Graph einer Funktion mit der x-Achse umschließt?

Insbesondere werden Sie gemeinsam mit mir erkennen, dass die mathematischen Antworten auf die erste und die letzte eben gestellte Frage in engster Art und Weise verbunden sind, obgleich diese vielleicht auf den ersten Blick wenig miteinander zu tun zu haben scheinen. Dies besagt der Hauptsatz der Differenzial- und Integralrechnung, über den ich mit Ihnen sprechen werde – aber so weit bin ich noch lange nicht.

6.1 Erste Ableitung von Funktionen und Ableitungsregeln

Es erscheint mir sinnvoll, mit einer Wanderung durch bergiges Gebiet zu beginnen. Die Steigungen, die Sie im Verlauf einer solchen Wanderung bewältigen müssen, hängen bekanntlich von der Stelle ab, an der Sie sich gerade befinden. Schauen Sie sich einmal meinen Wandervorschlag „Pfälzer Wald“ in Abbildung 6.1 an, der Ihnen einen Querschnitt durch das dortige Gebirgsprofil zeigt. Unter Mathematikern ist es durchaus normal, sich eine solche Wanderung als Funktion $f : [a, b] \mapsto \mathbb{R}$ vorzustellen. Jeder Stelle x im Intervall $[a, b]$ wird hier der zugehörigen Höhenwert $f(x)$ zugeordnet. Abbildung 6.1 zeigt Ihnen alle Punkte $P = (x|f(x))$, wobei x das Intervall $[a, b]$ durchläuft – also den **Graphen von** f. Die Wanderung beginnt am niedrigsten Punkt (an der Stelle a mit der Höhe $f(a)$). Am Anfang (an den Stellen im Bereich zwischen a und x_1) sind die Steigungen relativ groß. Danach (zwischen x_1 und x_2) sind die Steigungen etwas geringer – während das letzte Stück (zwischen x_2 und x_3) zunächst unnötigerweise sehr große Steigungen aufweist, dann aber – zum ersten Gipfel (an der Stelle x_3) hin – zunehmend flacher wird. Bis dahin ist der Wert der Steigungen stets positiv – am Gipfel selbst (an der Stelle x_3) ist zum Glück keine Steigung

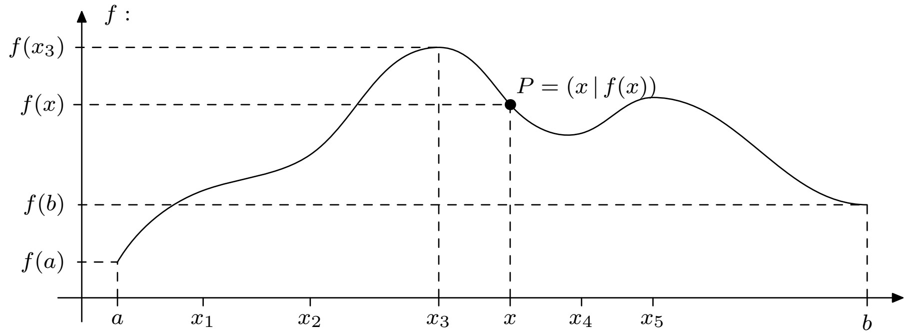

Abb. 6.1 Wandervorschlag „Pfälzer Wald“: Die Steigung von f ist nicht konstant, denn diese hängt von der Stelle x ab, an der man sich befindet.

mehr vorhanden, der Wert der Steigung ist dort 0. Keine Angst – ich gehe jetzt nicht alle Details dieser Wanderung mit Ihnen durch – im nächsten Abschnitt werde ich auf diese sowieso noch einmal zurückkommen, weil ich dort mit Ihnen unter anderem über Maxima und Minima von Funktionen plaudern werde. Wichtig ist mir im Moment, dass Sie erkennen: Die Steigung von dieser Funktion f ist *nicht konstant*. Das ist für sehr viele Funktionen so.

Sie haben längst gemerkt, um was es mir hier geht: Mein Ziel ist es, Ihnen den Begriff der **Steigung von Funktionen** f näher zu bringen. Der Wert der Steigung von f hängt normalerweise von der Stelle x ab, die ich betrachte. Die variierenden Werte der Steigung von f werden somit im Allgemeinen von einer weiteren Funktion beschrieben. Diese nennt man **erste Ableitung von** f und bezeichnet sie oft mit f' (sprich: „f Strich“). Ich schlage vor, dass ich zunächst die Berechnung der ersten Ableitungen einfacher Funktionen anschaue, bevor ich Ihnen die genaue und allgemeine Definition von ableitbaren Funktionen gebe.

Die einfachsten Funktionenklasse, die mir einfällt, sind **konstante Funktionen**, also $f(x) = c$, wobei c eine feste, reelle Zahl ist. Wie immer ist hier x eine reelle Variable. Die „Wattwanderung nördlich von Aurich“(siehe Abbildung 6.2) ist durch eine solche Funktion beschreibbar, wobei ich dort wieder davon ausgehe, dass x lediglich das Intervall $[a, b]$ durchläuft. Nehme ich an, dass Ebbe ist und die ganze Wan-

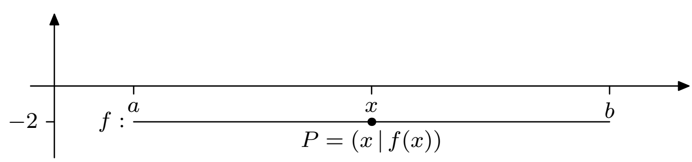

Abb. 6.2 Wandervorschlag „Wattwanderung nördlich von Aurich“: Die zugehörige Funktion f ist konstant -2 und somit ist 0 der Wert der Steigung von f an jeder Stelle x. Für dieses f verschwindet also die Steigung überall.

derung 2 Meter unter der Meeresoberfläche stattfindet, so kann ich die entsprechende Funktion hinschreiben: $f(x) = -2$ für alle $x \in [a, b]$. Sie erkennen unschwer, dass die Steigung dieser Funktion hier deutlich einfacher ist als im „Pfälzer Wald“: Der Wert der Steigung von f ist nämlich immer 0. Also: $f'(x) = 0$ für alle x im Intervall $[a, b]$. Genauso ist es, wenn ich eine andere Konstante für c wählen würde, also zum Beispiel $c = 2\,240$, $c = 4/7$ oder gar $c = -2/\pi$ – stellen Sie sich einfach vor, die Wattwanderung findet beispielsweise $2/\pi \approx 0.64$ Meter unter der Meeresoberfläche statt.

Erste Ableitung konstanter Funktionen

Der Wert der Steigung einer **konstanten Funktion** $f(x) = c$ ist stets 0. Die erste Ableitung f' von f **verschwindet** somit für alle x: $f'(x) = 0$.

Ich schaue mir nun die nächstschwierigere Klasse, nämlich die **linearen Funktionen** $f(x) = c\,x + d$, an. Wie immer ist $x \in \mathbb{R}$ die Variable und d und c sind feste, reelle Zahlen. Der Wert der Steigung solcher Funktionen ist glücklicherweise ebenfalls sehr einfach zu bestimmen. In Kapitel 2 wurde Ihnen erzählt, dass der Graph einer solchen Funktionen stets eine Gerade ist. Der *Wert der Steigung einer Geraden* ist aber bekanntlich *konstant*, das heißt *unabhängig* von der Stelle x, an der ich mich befinde. Eine *Ausnahmesituation* – aber warum ist das so?

Schauen Sie sich doch einmal zum Beispiel die in Abbildung 6.3 dargestellte lineare Funktion $f(x) = x$ an, die ich erhalte, indem ich $d = 0$ und $c = 1$ oben einsetze. Diese Funktion wird manchmal **Identität** genannt. Wenn ich von der Stelle 1 aus ein kleines Stückchen h ($\neq 0$) nach rechts gehe, also an die Stelle $1 + h$, so ist die Differenz der entsprechenden Funktionswerte, $f(1 + h) = 1 + h$ und $f(1) = 1$, ebenfalls h. Die Steigung ergibt sich durch Bildung des Quotienten der Differenz der Funktionswerte mit der Differenz der x-Werte:

$$\frac{f(1+h) - f(1)}{(1+h) - 1} = \frac{(1+h) - 1}{h} = \frac{h}{h} = 1.$$

Dasselbe Ergebnis erhalte ich, wenn ich von einer beliebigen anderen Stelle x aus ein kleines Stückchen h nach rechts gehe, also an die Stelle $x + h$, denn dann ist die Differenz der entsprechenden Funktionswerte $f(x + h) = x + h$ und $f(x) = x$ ebenfalls h. Sicherheitshalber schreibe ich Ihnen den Quotienten aus dieser Differenz der Funktionswerte und der Differenz der Stellen x und $x + h$ hin:

$$\frac{f(x+h) - f(x)}{(x+h) - x} = \frac{(x+h) - x}{h} = \frac{h}{h} = 1.$$

Ich stelle zweierlei fest. Erstens: Es ist hier im Ergebnis egal, wie h gewählt wurde, das heißt, wie weit ich von der festen Stelle x, die mich gerade interessiert, weggehe. Zweitens, unabhängig von x, hat der obige Quotient immer den Wert 1. Damit ist die erste Ableitung f' von f die Funktion: $f'(x) = 1$.

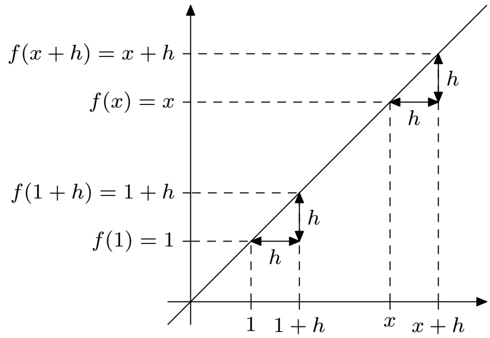

Abb. 6.3 Der Wert der Steigung der Identität $f(x) = x$ ist für jedes x gleich 1: Geht man von einer beliebigen Stelle x um h auf die Seite an die Stelle $x+h$, so ist der Unterschied der Funktionswerte $f(x)$ und $f(x+h)$ ebenfalls h.

Wenn ich dasselbe für die etwas schwierigere Funktion $f(x) = 3\,x$ (also $d = 0$ und $c = 3$) versuche, passiert Ähnliches. Wenn ich von einer festen (aber beliebigen) Stelle x aus um h auf die Seite gehe – also an der Stelle $x+h$ lande –, ist die Differenz der Funktionswerte $f(x+h) - f(x) = 3\,(x+h) - 3\,x = 3 \cdot h$. Der Quotient ist also $3 \cdot h/h = 3$. Somit ist $f'(x) = 3$ die erste Ableitung von $f(x) = 3\,x$. Genauso ist das, wenn ich den Graphen dieses fs um – sagen wir – 2 nach oben verschiebe, also $f(x) = 3\,x + 2$ betrachte – ich berechne dann ebenfalls $f'(x) = 3$ – versuchen Sie einmal für sich selbst, den Quotienten der Differenzen für dieses f zu bilden.

Jetzt betrachte ich eine allgemeine lineare Funktion $f(x) = c\,x + d$. Schnappe ich mir ein beliebiges, festes x und gebe mir ein h vor, so ist wie immer $(x+h) - x = h$. Zudem rechne ich die Differenz $f(x+h) - f(x)$ der Funktionswerte von f in x und $x+h$ folgendermaßen aus:

$$(c\,(x+h) + d) - (c\,x + d) = c\,x + c\,h + d - c\,x - d = c\,h.$$

Der Quotient aus den beiden Differenzen ist somit allgemein wie folgt:

$$\frac{f(x+h) - f(x)}{(x+h) - x} = \frac{c\,h}{h} = c.$$

Die Steigung hat also unabhängig von der Stelle x immer den Wert c und es ist auch egal, wie weit ich durch die Wahl von h ($\neq 0$) von der Stelle x weggehe: Das h konnte ich hier nämlich locker *rauskürzen*.

Erste Ableitung linearer Funktionen

Der Wert der Steigung einer **linearen Funktion** $f(x) = c\,x + d$ ist stets c. Die erste Ableitung f' von f ist somit für alle x **konstant**: $f'(x) = c$.

Konstante und lineare Funktionen stellen eine große Ausnahmesituation dar, denn deren erste Ableitung ist konstant. Das ist im Allgemeinen anders. Bereits der Wandervorschlag „Pfälzer Wald“ zeigte uns ja, dass dort die Werte der Steigung von der

Stelle x abhängen und somit *keine konstante* Funktion bilden. Es ist nun an der Zeit, eine konkrete Funktion zu betrachten, bei der die Werte der Steigung für verschiedene Stellen x auch *tatsächlich unterschiedlich* sind.

Dazu betrachte ich für $x \in \mathbb{R}$ die Funktion $f(x) = x^2$, deren Graph **Normalparabel** genannt wird. Diese kennen Sie schon aus Kapitel 2 – hier habe ich sie in Abbildung 6.4 als durchgezogenen Bogen veranschaulicht. Versuchsweise werde ich nun so vorgehen, wie ich es oben für lineare Funktionen gemacht habe. Mal schauen, wie weit ich damit kommen werde. Für jedes feste x und vorgegebenes h ($\neq 0$) ist die Differenz der Funktionswerte in x und $x + h$ dann $f(x + h) - f(x) = (x + h)^2 - x^2$, während die Differenz der Stellen $(x + h) - x$ wie immer h ergibt. Wenn ich mich jetzt an die **Binomischen Formeln** erinnere (schauen Sie vielleicht noch einmal in Kapitel 1 nach), so kann ich die Differenz der Funktionswerte weiter ausrechnen: $f(x + h) - f(x) = x^2 + 2\,h\,x + h^2 - x^2 = 2\,h\,x + h^2$. Der Quotient aus der Differenz der Funktionswerte mit der Differenz der Stellen ist somit:

$$\frac{f(x+h) - f(x)}{(x+h) - x} = \frac{2\,h\,x + h^2}{h} = \frac{h\,(2\,x + h)}{h} = 2\,x + h.$$

Nun sehe ich zwei *entscheidende Unterschiede* im Vergleich zu dem, was ich oben für lineare Funktionen herausbekommen habe. Erstens: Der Quotient hängt von h ab – h lässt sich leider nicht vollkommen wegkürzen. Zweitens: Hier tritt x auf der rechten Seite auf – der Quotient *hängt von der Stelle x ab.* Um hier weiterzukommen, sollte ich mir jetzt klar machen, was ich da gerade ausgerechnet habe: Der berechnete Quotient $2\,x + h$ ist, nach dem was ich Ihnen gerade über die Steigung von linearen Funktionen erzählt habe, die Steigung der Geraden durch die Punkte $(x|f(x)) = (x|x^2)$ und $(x + h|f(x + h)) = (x + h|(x + h)^2)$. In Abbildung 6.4 habe ich dies für $x = 1$ und $h = 2$ veranschaulicht: Die gestrichelte Gerade s ist der Graph der linearen Funktion

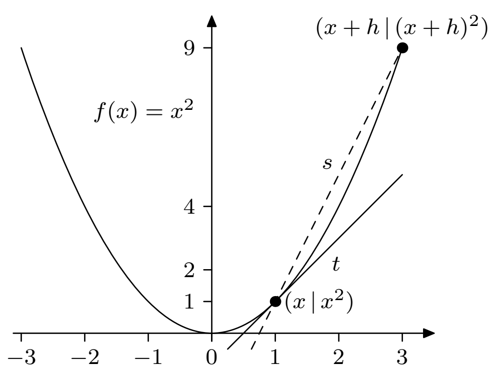

Abb. 6.4 Der Wert der Steigung von $f(x) = x^2$ an der Stelle $x = 1$ ist 2. Das ist auch der Wert der Steigung der Tangente t an f in $x = 1$, welche sich als Grenzwert der Steigungen der Sekanten $2 + h$ für $h \to 0$ ergibt. Für diese Funktion f ist die erste Ableitung f' nicht konstant, sondern es gilt: $f'(x) = 2\,x$.

$s(x) = 4\,x - 3$, die in $x = 1$ und $x = 3$ denselben Wert wie f hat: $f(1) = 1 = s(1)$ und $f(3) = 9 = s(3)$. Die Steigung dieser Gerade s ist 4. Dies ist lediglich *eine erste Näherung* an den tatsächlichen Wert der Steigung von f in $x = 1$.

Damit ich genauere Näherungen an den tatsächlichen Wert der Steigung von f in $x = 1$ erhalten kann, stelle ich mir nun vor, dass h *immer näher bei* 0 *liegt, ohne je* 0 *zu werden* – man sagt, dass h **gegen** 0 **strebt** und schreibt: $h \to 0$. Die Steigung der Geraden durch die Punkte $(1|1)$ und $(1+h|(1+h)^2) = (1+h|1+2\,h+h^2)$ ist $2+h$ und stellt für solche, immer kleiner werdenden h *immer genauere Näherungen* des *tatsächlichen Werts* der Steigung von $f(x) = x^2$ in $x = 1$ dar. Setze ich für h beispielsweise $1/10, 1/100, 1/1\,000, \ldots$, so erhalte ich der Reihe nach für diese Näherung $2 + h$ die Ergebnisse 2,1, 2,01 und 2,001. Ich denke, Sie sehen jetzt, dass der Term $2 + h$ **im Grenzübergang** $h \to 0$ gegen den **Grenzwert** 2 **konvergiert**. Viele Mathematiker verwenden hierfür zuweilen gerne die übersichtliche Schreibweise

$$\lim_{h \to 0} (2 + h) = 2$$

und meinen damit, dass der Term $2 + h$ *beliebig nah* an die Grenze (lat.: **Limes**) 2 herankommt, falls h nur nah genug bei 0 liegt. Ich habe somit herausgefunden, dass 2 der tatsächliche Wert der Steigung der Funktion $f(x) = x^2$ an der Stelle $x = 1$ ist: $f'(1) = 2$. Dieser Wert ist gerade die Steigung der als durchgezogene Linie dargestellten Gerade t in Abbildung 6.4. Diese „**tangiert**“ (lat.: „berührt“) den Graph der Funktion f im Punkt $(1|1)$. Ich kann mir vorstellen, dass diese Gerade durch den Grenzübergang $h \to 0$ aus den Geraden durch die beiden Punkte $(1|1)$ und $(1+h|(1+h)^2)$ hervorgeht.

Ich sollte noch eine *wichtige Bemerkung* über die Bedeutung des **Limeszeichens** $\lim\limits_{h \to 0}$ machen, denn dieses tritt im Folgenden noch öfter auf. *Eigentlich* ist dies so gemeint, dass man schauen sollte, ob der Grenzwert *existiert und jeweils gleich ist*, wenn ich *alle* Möglichkeiten betrachte, mit h gegen 0 zu streben. Das ist oben natürlich der Fall – insbesondere kann ich dort auch *negative* hs mit $h \to 0$ anschauen, also zum Beispiel $-1/10, -1/100, -1/1\,000, \ldots$, oder vielleicht hs mit $h \to 0$ und **alternierenden Vorzeichen**, also zum Beispiel $-1/2, 1/4, -1/8, 1/16, \ldots$. Sie sehen: Im obigen Term $2 + h$ kommt auch dann immer 2 als Grenzwert heraus.

Was ist der Wert der Steigung der Funktion $f(x) = x^2$ nun an einer beliebigen Stelle x? Der oben berechnete Quotient $2\,x + h$ ist lediglich eine Näherung an diesen Wert, den ich als Grenzwert durch Bildung des Grenzübergangs $h \to 0$ fast genauso wie oben bestimmen kann, indem ich (sehr) *salopp* gesprochen 0 für h in $2\,x + h$ *einsetze*:

$$\lim_{h \to 0} (2\,x + h) = 2\,x.$$

Ich sehe, dass $2\,x$ der *von x abhängige* Wert der Steigung von $f(x) = x^2$ an der Stelle x ist: In $x = 1$ ist 2 der Wert der Steigung von f: $f'(1) = 2$, während ich in $x = -3$ den Wert der Steigung wie folgt berechne: $f'(-3) = 2\,(-3) = -6$.

Erste Ableitung der quadratischen Potenzfunktion $f(x) = x^2$

Der Wert der Steigung der **quadratischen Potenzfunktion** $f(x) = x^2$ ist $2\,x$. Die erste Ableitung f' von f ist somit für alle x festgelegt durch: $f'(x) = 2\,x$.

Ein bisschen Übung kann jetzt bestimmt nicht schaden:

Übungsaufgabe 6.1:
a) Bestimmen Sie den Wert der Steigung von $f(x) = x^2$ an den Stellen $x = 3$ und $x = -2$. b) Bestimmen Sie die Stelle x, an der die erste Ableitung von $f(x) = x^2$ den Wert -2 hat. ■

Übungsaufgabe 6.2:
Zeigen Sie wie oben, dass $p_3'(x) = 3\,x^2$ die erste Ableitung von $p_3(x) = x^3$ ist. Hinweis: $(x+h)^3 = x^3 + 3\,x^2\,h + 3\,x\,h^2 + h^3$. ■

Vielleicht ist Ihnen bei der Bildung der bisherigen ersten Ableitungen schon eine gewisse Regelmäßigkeit aufgefallen. Damit Sie sehen, worauf ich nun hinauswill, verrate ich Ihnen, dass $4\,x^3$ der Wert der Steigung von $p_4(x) = x^4$ an einer beliebigen Stelle x ist. Außerdem sage ich Ihnen, dass die erste Ableitung von $p_5(x) = x^5$ die Funktion $p_5'(x) = 5\,x^4$ ist. Die *Grobregel* zur Bestimmung der ersten Ableitung einer allgemeinen **Potenzfunktion** $p_i(x) = x^i$, wobei $i \in \mathbb{N}$, ist also: Schreibe den Exponenten i multiplikativ nach vorne und ziehe eins von diesem Exponenten ab.

Erste Ableitung der Potenzfunktion $p_i(x) = x^i$

Der Wert der Steigung der **Potenzfunktion** $p_i(x) = x^i$ ist $i\,x^{i-1}$. Die erste Ableitung p_i' von p_i ist somit für alle x festgelegt durch: $p_i'(x) = i\,x^{i-1}$.

Beispiel 6.1:
Sie können jetzt also locker die erste Ableitung p_{128}' von $p_{128}(x) = x^{128}$ berechnen: $p_{128}'(x) = 128\,x^{127}$. Oder wenn Sie morgen jemand auf der Straße anspricht und fragen sollte: „Was ist der Wert der Steigung der Funktion $p_7(x) = x^7$ an der Stelle $x = -\sqrt{2}$?“, so können Sie souverän antworten: Die erste Ableitung p_7' von $p_7(x) = x^7$ ist $p_7'(x) = 7\,x^6$. Setze ich $x = -\sqrt{2}$ in diese allgemeine Darstellung von p_7' ein, so erhalte ich: $p_7'(-\sqrt{2}) = 7\,(-\sqrt{2})^6 = 7\,(-1)^6(\sqrt{2})^6 = 7\,(2^{1/2})^6 = 7\,(2^{(1/2)6}) = 7\,(2^{6/2}) = 7\,(2^3) = 7 \cdot 8 = 56$. Prima. ■

Natürlich ist die obige Vorgehensweise zur Bestimmung der Werte der Steigungen einer Funktion, beziehungsweise deren erste Ableitung, nicht nur auf Potenzfunktionen $p_i(x) = x^i$ beschränkt. Ich erkläre Ihnen jetzt, wie man dieses Konzept der ersten Ableitung *verallgemeinert* – hierzu empfehle ich noch einmal Abbildung 6.4 anzuschauen und bitte darum, die nachfolgenden Beschreibungen mit der obigen Vorgehensweise für $f(x) = x^2$ zu vergleichen. Ich schaue mir jetzt also allgemeinere Funktionen f einer reellen Variablen x an. Wenn Sie möchten, können Sie sich solche

fs vorstellen wie das f der Wanderung im „Pfälzer Wald“ in Abbildung 6.1. Ich betrachte im Folgenden ein beliebiges, aber festes x und geeignete h ($\neq 0$) nahe bei 0. Dies bedeutet, dass ich annehmen möchte, dass x und $x+h$ beide stets im **Definitionsbereich** von f (siehe Kapitel 2) liegen.

Sekante von f durch x und $x+h$

Die Gerade durch die beiden Punkte $(x|f(x))$ und $(x+h|f(x+h))$ bezeichnet man als die **Sekante von f durch x und $x+h$**.

Für das feste x hängt die Sekante von f durch x und $x+h$ natürlich auch von der Wahl von h ab. In Abbildung 6.4 habe ich eine Sekante mit s bezeichnet und als gestrichelte Linie dargestellt. Die *Steigung dieser Sekante* ist der Quotient

$$\frac{f(x+h)-f(x)}{h}$$

aus der Differenz der Funktionswerte $f(x+h)$ und $f(x)$ mit der Differenz der Stellen $x+h$ und x. Diesen Quotienten bezeichnet man gerne als **Differenzenquotienten von f in x für h**. Solche Quotienten stellen im Allgemeinen lediglich *Näherungen* an die tatsächliche Steigung von f in x dar. Den Übergang dieser näherungsweisen Steigungen zum Wert der *tatsächlichen Steigung* von f in x wird konzeptuell durch eine Grenzwertbildung beschrieben.

Differenzialquotient und erste Ableitung von f in x

Falls der Grenzwert der Differenzenquotienten von f in x für h gegen 0

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$$

existiert, so heißt dieser **Differenzialquotient von f in x**, und wird mit $f'(x)$ (sprich: „f Strich von x“) bezeichnet. Ist dies der Fall, so ist die *reelle Zahl* $f'(x)$ der Wert der Steigung von f in x und wird **erste Ableitung von f an der Stelle x** genannt. Man sagt dann, dass **f in x (einmal) ableitbar** ist.

Manche Leute sprechen auch davon, dass f **an der Stelle** x **(einmal) differenzierbar** ist – dies meint exakt dasselbe. *Falls* f an der Stelle x ableitbar ist, so stimmt der Wert $f'(x)$ der Steigung von f in x mit dem Wert der Steigung einer besonderen Geraden überein – diese tritt als die durch den Grenzübergang $h \to 0$ aus den Sekanten hervorgehende Gerade auf und „**tangiert**“ den Graph der Funktion f im Punkt $(x|f(x))$. Im Beispiel von Abbildung 6.4 habe ich diese besondere Gerade t als durchgezogene Linie dargestellt.

Tangente von f in x

Die lineare Funktion t durch den Punkt $(x|f(x))$ mit der Steigung $f'(x)$ bezeichnet man als **Tangente von f in x**.

Die Tangente t von f in x kann ich explizit hinschreiben, indem ich in die obige allgemeine Form von linearen Funktionen $c = f'(x)$ und $d = -f'(x)\,x + f(x)$ einsetze. Das sieht folgendermaßen aus

$$t(\bar{x}) = f'(x)\,\bar{x} - f'(x)\,x + f(x),$$

wobei ich hier zur Unterscheidung von der festen Stelle x ausnahmsweise die Variablenbezeichnung $\bar{x}$ verwende. Wenn ich hier x für $\bar{x}$ einsetze, kommt $f(x)$ heraus, und ich erkenne auch, dass natürlich $t'(\bar{x}) = c = f'(x)$ gilt.

Beispiel 6.2:
Im Beispiel von Abbildung 6.4 sind $c = f'(1) = 2$ und $d = -f'(1)\,1 + f(1) = -(2 \cdot 1) + 1 = -1$ und die Tangente von $f(x) = x^2$ in $x = 1$ ist die lineare Funktion $t(\bar{x}) = 2\,\bar{x} - 1$. ■

Besondere Freude erlebt der Mathematiker, wenn die Funktion f an *allen* Stellen x ableitbar ist. In diesem Fall sagt man, dass **die Funktion f ableitbar** ist und man hat dann – genau wie ich dies eingangs dieses Abschnitts schon geschrieben habe – eine *weitere* reellwertige Funktion f' in der Variablen x, die für jede Stelle x den Wert der Ableitung $f'(x)$ von f liefert. Diese nennt man dann **Ableitungsfunktion von f**, kurz: **Ableitung von f**. Das habe ich übrigens im Hinterkopf gehabt, als ich oben für die feste Stelle x den Differenzialquotienten mit dem Symbol $f'(x)$ bezeichnete: Man nimmt hier sozusagen den Zusammenhang zur Funktion f' vorweg, weil man hofft, dass f überall ableitbar ist.

Eine kleine Übung zur Verinnerlichung der Begriffe schadet sicherlich nicht:

Übungsaufgabe 6.3:
Berechnen Sie die Differenzenquotienten von $f(x) = x^2 + x$ in x für h und zeigen Sie damit, dass $f'(x) = 2\,x + 1$ die erste Ableitung von f für alle x ist. Hinweis: Schauen Sie, wie ich oben für $p_2(x) = x^2$ vorgegangen bin. ■

Im Folgenden diskutiere ich mit Ihnen noch drei weitere Beispiele ableitbarer Funktionen. Dies sind Standard-Funktionen, die Sie aus den vorherigen Kapiteln kennen.

Erste Ableitung der Exponentialfunktion $\exp(x) = e^x$

Der Wert der Steigung der **Exponentialfunktion** $\exp(x) = e^x$ ist e^x. Die erste Ableitung $\exp'$ von $\exp$ ist für alle x festgelegt durch: $\exp'(x) = e^x$.

Sie sehen, dass die Ableitung $\exp'$ der Exponentialfunktion $\exp$ die Exponentialfunktion selbst ist: $\exp' = \exp$. Das ist recht übersichtlich – aber wie kann man dies nun wieder erkennen? Ich skizziere Ihnen hier lediglich eine Grundidee:

Zunächst betrachte ich ein beliebiges x und berechne für $h \neq 0$ die Differenzenquotienten:

$$\frac{\exp(x+h)-\exp(x)}{h} = \frac{e^{x+h}-e^x}{h} = \frac{e^x\,e^h-e^x}{h} = e^x\,\frac{e^h-1}{h}.$$

Dabei habe ich im vorletzten Schritt das Potenzgesetz $e^{x+h} = e^x \cdot e^h$ angewandt und dann e^x im Zähler ausgeklammert. Jetzt muss ich den Differenzialquotienten ausrechnen – also den folgenden Grenzwert:

$$\lim_{h\to 0}\frac{\exp(x+h)-\exp(x)}{h} = \lim_{h\to 0} e^x\,\frac{e^h-1}{h} = e^x \lim_{h\to 0}\frac{e^h-1}{h} = e^x \cdot 1 = e^x.$$

Ich erkenne, dass für diesen wieder e^x herauskommt – also: $\exp'(x) = e^x = \exp(x)$. Den Term e^x kann ich hier nach dem zweiten „=“-Zeichen, ohne etwas falsch zu machen, vor den Grenzwert schreiben, denn dieser hat nichts mit der Grenzwertbildung für $h \to 0$ zu tun, weil h *nicht* in diesem auftritt. Das ist intuitiv verständlich. Aber danach habe ich zugegebenermaßen etwas gemacht, was eigentlich eine *genauere Analyse* erfordert – ich verwende nämlich:

$$\lim_{h\to 0}\frac{e^h-1}{h} = 1.$$

Ich beweise Ihnen dies hier nicht – Sie können mir aber voll vertrauen: Das läßt sich aus der Definition der **Euler'schen Zahl** e ableiten. Die obige saloppe Vorgehensweise, 0 einfach für h in den Differenzenquotienten einzusetzen, tut es hier übrigens nicht: Ich würde dann durch 0 teilen, was ich *nie nie nie* tun darf. Ich will Sie aber damit auch nicht ganz alleine lassen und habe Ihnen deshalb Tabelle 6.1 angefertigt, aus der Sie die Werte von $(e^h-1)/h$ für hs nahe 0 ablesen können.

Tab. 6.1 Werte des Terms $(e^h-1)/h$ für Wahlen von h nahe 0.

$h=$	$-0{,}1$	$-0{,}01$	$-0{,}001$	$0{,}001$	$0{,}01$	$0{,}1$
$(e^h-1)h \approx$	0,951 63	0,995 02	0,999 50	1,000 50	1,005 02	1,051 71

Meine nächsten beiden Beispiele ableitbarer Funktionen sind die **Sinusfunktion** $\sin(x)$ und die **Cosinusfunktion** $\cos(x)$. Hier geht man normalerweise genauso vor: Erst bildet man die Differenzenquotienten und dann den Differenzialquotienten. Man benötigt dabei die **Additionstheoreme** und ähnliche Grenzwertbetrachtungen wie bei der Exponentialfunktion. Ich mache das jetzt nicht, denn dies werden vermutlich später andere in Ihrem Studium für mich tun.

Erste Ableitung der Sinusfunktion $\sin(x)$ und Cosinusfunktion $\cos(x)$

Die erste Ableitung $\sin'$ der **Sinusfunktion** $\sin(x)$ ist für alle x festgelegt durch: $\sin'(x) = \cos(x)$. Die erste Ableitung $\cos'$ der **Cosinusfunktion** $\cos(x)$ ist für alle x festgelegt durch: $\cos'(x) = -\sin(x)$.

Ich muss Ihnen jetzt noch etwas sagen, was Ihnen vielleicht nicht so gut gefällt: Es gibt leider auch Funktionen, die *nicht* ableitbar sind. Eine Funktion f heißt **nicht ableitbar**, falls es (mindestens) *eine Stelle x gibt* für die f nicht ableitbar ist. Vielleicht ist Ihnen aufgefallen, dass ich oben bei der Formulierung des Begriffs des Differenzialquotienten *sehr vorsichtig* war („*Falls der Grenzwert ... existiert*, ...“), und das hat auch seinen guten Grund: Es könnte nämlich sein, dass für gewisse Funktionen f an manchen Stellen x mit der Bildung des Grenzwerts der Differenzenquotienten für $h \to 0$ etwas schief läuft. Damit meine ich, dass es einem – wenn es dumm läuft – passieren kann, dass der Differenzialquotient *nicht existiert*. Dies bedeutet, dass dieser entweder gar *keine reelle Zahl* ist oder aber nicht so klar ist, welche Zahl dieser Grenzwert sein sollte – dieser also *nicht eindeutig festlegbar* ist. Bevor ich beginne Sie zu verwirren, zeige ich Ihnen am besten zwei klassische Beispiele von Funktionen, die nicht ableitbar sind.

Ich beginne mit der **(Quadrat-)Wurzelfunktion** $f(x) = \sqrt{x}$, wobei $x \geq 0$. Ich darf hier keine negativen Zahlen x verwenden, denn wenn ich die Wurzel aus einer negativen Zahlen ziehe würde, so wäre dies leider keine reelle Zahl mehr, und ich behandle hier nur Funktionen, die **reellwertig** sind: $f(x) \in \mathbb{R}$ für alle x. Testweise verfahre ich nun wie üblich: Erst bilde ich die Differenzenquotienten für ein beliebiges, festes $x \geq 0$, wobei h so gewählt ist, dass $x + h > 0$ gilt. Die Differenzenquotienten sind

$$\frac{f(x+h) - f(x)}{h} = \frac{\sqrt{x+h} - \sqrt{x}}{h} = \frac{h}{h\,(\sqrt{x+h} + \sqrt{x})} = \frac{1}{(\sqrt{x+h} + \sqrt{x})}.$$

Hier habe ich die Quotienten nach dem ersten „=“-Zeichen zunächst mit $\sqrt{x+h} + \sqrt{x}$ erweitert, dann nach dem zweiten „=“-Zeichen die dritte **Binomische Formel** (siehe Kapitel 1) angewandt $(\sqrt{x+h} - \sqrt{x})\,(\sqrt{x+h} + \sqrt{x}) = x + h - x = h$ und schließlich nach dem dritten „=“-Zeichen das h gekürzt. Jetzt versuche ich, den Grenzwert für $h \to 0$ herauszufinden – also den Differenzialquotienten zu berechnen:

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{1}{(\sqrt{x+h} + \sqrt{x})} = \frac{1}{(\sqrt{x} + \sqrt{x})} = \frac{1}{2\,\sqrt{x}}.$$

Ich habe hierzu nicht zu viel überlegt und – wie oben – einfach salopp 0 für h in den Term $1/(\sqrt{x+h} + \sqrt{x})$ eingesetzt, um den Grenzwert zu bestimmen. Sieht eigentlich ganz in Ordnung aus, denn ich erhalte $f'(x) = 1/(2\sqrt{x}) = (1/2)\,x^{-1/2}$, und Sie fragen sich jetzt vielleicht, was denn hier schief gelaufen sein sollte. In der Tat ist hier alles in bester Butter, *solange* x nicht ausgerechnet gleich 0 ist. Die Wurzelfunktion ist an allen positiven Stellen $x > 0$ ableitbar. Die verbleibende Stelle $x = 0$ macht allerdings enorme Schwierigkeiten. Wenn ich mir das nochmal für $x = 0$ anschaue (h sollte dann > 0 sein) bekomme ich nämlich ein recht ärgerliches Problem:

$$\lim_{h \to 0} \frac{f(0+h) - f(0)}{h} = \lim_{\substack{h \to 0 \\ h > 0}} \frac{1}{(\sqrt{0+h} + \sqrt{0})} = \lim_{\substack{h \to 0 \\ h > 0}} \frac{1}{\sqrt{h}}.$$

Wenn hier h immer näher an 0 herangeht, wird der Term im Nenner $\sqrt{h}$ ebenfalls gegen 0 gehen und der Differenzenquotient $1/\sqrt{h}$ wird immer größer. Ich versuche das einmal: Testweise nehme ich $h = 1/100 = 10^{-2},\ 1/10\,000 = 10^{-4},\ 1/1\,000\,000 = 10^{-6}$ und erhalte der Reihe nach die folgenden Werte durch Einsetzen in den Term $1/\sqrt{h}$: $1/\sqrt{10^{-2}} = 1/(10^{-2})^{\frac{1}{2}} = 1/10^{(-2)\frac{1}{2}} = 1/10^{-1} = 10$, $1/\sqrt{10^{-4}} = 100$, $1/\sqrt{10^{-6}} = 1\,000$. Ich erkenne, dass der Term $1/\sqrt{h}$ immer größer wird, umso näher h an 0 herangeht. Der Term nähert sich somit leider keinesfalls einer *reellen Zahl* an – im Gegenteil, er wächst über alle positive Schranken hinaus an. Als grobe Regel kann man sich vorstellen, dass man im Grenzübergang $h \to 0$ *durch* 0 *teilen würde*, was man bekanntlich *nie nie nie* machen darf: Ich darf hier also *nicht* salopp $h = 0$ in den Term $1/\sqrt{h}$ einfach einsetzen um den Grenzwert zu ermitteln. Man schreibt dies nun folgendermaßen hin:

$$\lim_{\substack{h\to 0\\ h>0}} \frac{1}{\sqrt{h}} = \infty.$$

Das Zeichen ∞ steht als Abkürzung für **unendlich groß**: Ich finde keine reelle Zahl, die ich auf die rechte Seite neben das „=“-Zeichen ohne zu mogeln hinschreiben könnte. Der Grenzwert des Differenzenquotienten existiert für die Stelle $x = 0$ nicht. Anschaulich bedeutet dies, dass der Wert der Steigung der Funktion größer als jede vorgegebene, positive Schranke wird, wenn ich nur nah genug an die Stelle $x = 0$ herangehe (siehe Abbildung 6.5). Die Wurzelfunktion ist somit an der Stelle $x = 0$ nicht ableitbar. Damit ist diese Funktion nicht ableitbar – obgleich es an allen anderen Stellen, die größer 0 sind, mit der Grenzwertbildung geklappt hat. Dies hat übrigens nicht damit zu tun, dass es sich hier zufällig um die spezielle Stelle $x = 0$ handelt: So ist die Funktion $f(x) = \sqrt{x-2}$, $x \geq 2$, deren Graph man erhält, indem man die Wurzelfunktion um zwei Einheiten auf der x-Achse nach rechts verschiebt, in $x = 2$ nicht ableitbar.

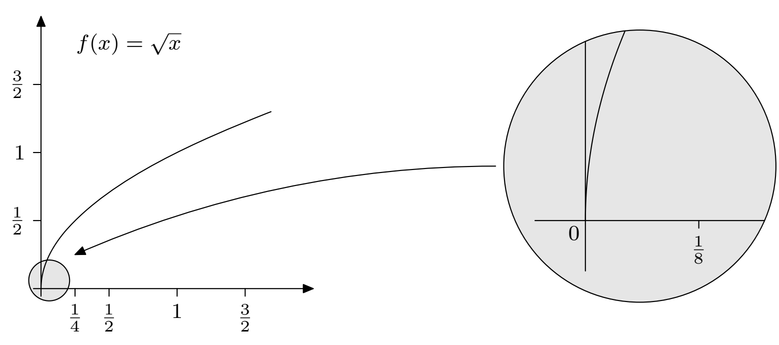

Abb. 6.5 Die Wurzelfunktion $f(x) = \sqrt{x}$ ist nicht ableitbar, denn sie ist an der Stelle $x = 0$ nicht ableitbar: Der Wert der Steigung von f wird unendlich groß, wenn man sich $x = 0$ von rechts kommend nähert. Ein Zoom auf die Stelle $x = 0$ (rechts) illustriert, dass der Wert der Steigung über alle positive Schranken hinauswächst.

Es gibt hier aber zum Glück auch etwas Positives zu berichten:

Erste Ableitung der Wurzelfunktion $f(x) = \sqrt{x}$, $x \in [a, b]$ mit $a > 0$

Falls $a > 0$ ist, so ist $1/(2\sqrt{x})$ der Wert der Steigung der **(Quadrat-) Wurzelfunktion** $f(x) = \sqrt{x}$ für alle $x \in [a, b]$. Die erste Ableitung f' von f ist dann für alle $x \in [a, b]$ festgelegt durch: $f'(x) = 1/(2\sqrt{x})$.

Wenn ich also ein sinnvolles Intervall $[a, b]$ betrachte, das die Stelle $x = 0$ nicht enthält, so ist die Wurzelfunktion dort ableitbar und ich kann die erste Ableitung so berechnen, wie ich es nach der oben genannten Grobregel (für Potenzfunktionen) vielleicht schon vermuten würde: $(x^{1/2})' = 1/2\, x^{-1/2}$. Zur Information flüstere ich Ihnen kurz zu, dass die **m-te-Wurzelfunktion** $f(x) = \sqrt[m]{x} = x^{1/m}$, wobei $m \in \mathbb{N}$, in den zulässigen Bereichen genauso abgeleitet wird: $f'(x) = (1/m)\, x^{(1/m)-1} = (1/m)\, x^{(1-m)/m} = (1/m)\, x^{-((m-1)/m)} = (1/m)\, (1/\sqrt[m]{x^{m-1}})$. Sie sehen also, dass die obige Grobregel zur Bildung der Ableitung hier auch gilt. Nicht vergessen sollte ich zu sagen, dass die m-te-Wurzelfunktion in $x = 0$ nicht ableitbar ist.

Sie sehen also ein, dass man bei der Bildung des Grenzwerts des Differenzenquotienten für $h \to 0$ im Allgemeinen gut aufpassen muss. Wichtig ist, dass für jedes x auch tatsächlich eine reelle Zahl für den Differenzialquotienten herauskommt und nicht etwa ∞. Damit aber nicht genug: Ich muss Sie noch auf eine weitere Problematik hinweisen, die ebenfalls mit der Bildung des Differenzialquotienten zusammenhängt. In Ausnahmesituationen kann es nämlich passieren, dass der Differenzialquotient nicht existiert, *obwohl* die zugehörigen Differenzenquotienten nicht unendlich groß werden, wenn h gegen 0 geht. Oben habe ich bereits angedeutet, dass diese Schwierigkeit auftreten kann, falls nicht klar ist, welchen Wert der Differenzialquotient bekommen soll – dieser also *nicht eindeutig festlegbar* ist: Man hat dann *zwei Kandidaten* für diesen Wert, kann sich aber, ähnlich wie dies in vielen Situationen unseres täglichen Lebens ist, leider nicht entscheiden.

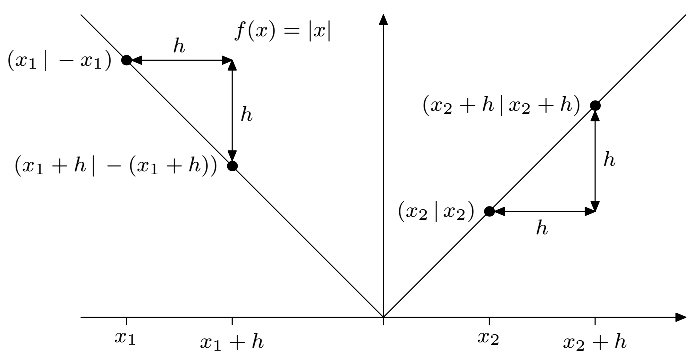

Abb. 6.6 Die Betragsfunktion $f(x) = |x|$ ist nicht ableitbar, denn sie ist an der Stelle $x = 0$ nicht ableitbar: Der Graph von f hat dort einen „Knick“.

Um klar zu machen, was ich meine, schaue ich mir den angekündigten zweiten Klassiker an. Dies ist die für $x \in \mathbb{R}$ definierte **Betragsfunktion**

$$f(x) = |x| = \begin{cases} x, & \text{falls } x \geq 0 \\ -x, & \text{falls } x < 0 \end{cases},$$

die ich Ihnen in Abbildung 6.6 veranschaulicht habe. Beachten Sie, dass die Betragsfunktion *keine* lineare Funktion ist: Diese Funktion ist nur **stückweise linear**. Wenn ich mir nun ein beliebiges $x > 0$ schnappe und den Grenzwert der Differenzenquotienten ausrechnen möchte, so geht dies genauso wie oben für die Identität:

$$\lim_{h \to 0} \frac{x + h - x}{h} = \lim_{h \to 0} \frac{h}{h} = \lim_{h \to 0} 1 = 1.$$

Genauso funktioniert die Berechnung des Differenzialquotienten für beliebige $x < 0$:

$$\lim_{h \to 0} \frac{-(x+h) - (-x)}{h} = \lim_{h \to 0} \frac{-x - h + x}{h} = \lim_{h \to 0} \frac{(-h)}{h} = \lim_{h \to 0} (-1) = -1.$$

Hier habe ich stillschweigend vorausgesetzt, dass h so nahe bei 0 liegt, dass $x + h < 0$ ist und deshalb $|x + h| = -(x + h)$ gemäß der obigen Definition des Betrags gilt. Wunderbar – ich stelle also fest, dass die Betragsfunktion für $x > 0$ sowie $x < 0$ ableitbar ist. Die Ableitung hat den Wert 1 beziehungsweise -1. Prima, aber wie sieht es mit der verbleibenden Stelle $x = 0$ aus? Das muss ich mir noch ein bisschen genauer anschauen. Ich setze in die beiden obigen Differenzenquotienten jeweils $x = 0$ ein und betrachte sowohl den **rechtsseitigen Grenzwert** (Differenzenquotienten werden nur für $h > 0$ gebildet):

$$\lim_{\substack{h \to 0 \\ h > 0}} \frac{0 + h - 0}{h} = \lim_{\substack{h \to 0 \\ h > 0}} \frac{h}{h} = 1$$

als auch den **linksseitigen Grenzwert** (Differenzenquotienten für $h < 0$ gebildet):

$$\lim_{\substack{h \to 0 \\ h < 0}} \frac{-(0 + h) - 0}{h} = \lim_{\substack{h \to 0 \\ h < 0}} \frac{(-h)}{h} = -1.$$

Jetzt sehen Sie, was ich oben mit dem Problem der *zwei Kandidaten* für den Wert des Differenzialquotienten gemeint habe. Da ist offenbar eine Schwierigkeit: Wenn ich von rechts aus auf die Stelle $x = 0$ zugehe ($h > 0$), dann ist mein Kandidat für die Ableitung von $f(x) = |x|$ in $x = 0$ der Wert 1, während von links kommend ($h < 0$) ein anderer Kandidat für die Ableitung von f in $x = 0$ auftritt: -1. Die Kandidaten für den Differenzialquotienten sind zwar nicht unendlich groß, wie das noch mit der Wurzelfunktion in $x = 0$ war. Trotzdem ist das hier aber auch nicht so toll, denn ich kann jetzt gar nicht so genau sagen, welcher der beiden Kandidaten nun der *tatsächliche* Grenzwert des Differenzenquotienten von f in $x = 0$ sein sollte. Das ist ähnlich wie im Kasino: Möglicherweise gefallen Ihnen die Farben „Rot“ und „Schwarz“ beide sehr gut und Sie können sich partout nicht festlegen, auf welche

Farbe Sie Ihren Chip beim Roulett setzen sollen. In der Mathematik macht man es sich nun einfach: Der Differenzialquotient von f in 0 ist *nicht eindeutig festlegbar* und deshalb ist f in $x = 0$ nicht ableitbar. Somit ist die Betragsfunktion nicht ableitbar – im übertragenen Sinne bedeutet dies, dass Sie Ihren Chip in Geld umtauschen und nach Hause gehen. Damit haben Sie sicherlich nichts falsch gemacht, und auf diesen Standpunkt stellt man sich auch in der Mathematik, wenn die eindeutige Festlegung des Differenzialquotientens partout nicht gelingen will. Anschaulich können Sie dies an dem „Knick“ erkennen, den der Graph der Betragsfunktion in $x = 0$ hat. Die Funktionen in den Abbildungen 6.1 bis 6.4 haben alle keinen solchen „Knick“ – die Graphen sehen *glatt* aus. Dies ist für ableitbare Funktionen so – und oben haben Sie ja gesehen, dass diese Funktionen glücklicherweise an jeder Stelle x und damit tatsächlich insgesamt ableitbar sind.

Training ist wichtig. Deshalb bitte ich Sie, sich an den folgenden Aufgaben zu versuchen. Achtung: Der b)-Teil ist gar nicht so einfach.

Übungsaufgabe 6.4:
Untersuchen Sie, ob a) $f(x) = |x - 2|$ und b) $f(x) = \frac{1}{2}\left((x-2)^3 + |x-2|^3\right)$ an der Stelle $x = 2$ ableitbar ist. Sind die Funktionen ableitbar? ■

Ich möchte mich jetzt wieder in eine positive Richtung bewegen und langfristig versuchen, mir das Bestimmen von Ableitungen etwas einfacher zu gestalten. Wenn ich mich stur an die obige Vorgehensweise klammere, so kann ich die Ableitung einer vorgegebenen Funktion ermitteln, indem ich zunächst die Differenzenquotienten an jeder Stelle bestimme und dann versuche, den Differenzialquotienten zu berechnen. Dabei ist darauf zu achten, dass dieser stets existiert und insbesondere eindeutig ist. Dies ist im Allgemeinen recht mühselig, wie Sie vielleicht schon selbst bei der Bearbeitung von Übungsaufgabe 6.3 festgestellt haben.

Beispiel 6.3:
Ein anderes etwas mühseliges Beispiel ist die Funktion $f(x) = -5\,x^2 + 4\sqrt{x}$, wobei ich wegen des Wurzelterms annehme, dass x aus einem Intervall $[a, b]$ mit $a > 0$ ist. Wenn ich die obige Vorgehensweise anwende, so würde ich hier erst einmal die Differenzenquotienten bilden:

$$\begin{aligned}\frac{f(x+h)-f(x)}{h} &= \frac{(-5)\,(x+h)^2 + 4\sqrt{x+h} - (-5\,x^2 + 4\sqrt{x})}{h}\\ &= \frac{-5\,x^2 - 10\,x\,h - 5\,h^2 + 4\sqrt{x+h} + 5\,x^2 - 4\sqrt{x}}{h}\ .\end{aligned}$$

Wende ich dieselben Rechentricks, die ich Ihnen oben bei der Berechnung der Differenzenquotienten der quadratischen Potenzfunktion und der Wurzelfunktion gezeigt habe, noch einmal (simultan) mit kleiner Variation an, so erhalte ich

$$\begin{aligned}\frac{f(x+h)-f(x)}{h} &= \frac{(-10\,x - 5\,h)\,h}{h} + \frac{4\sqrt{x+h} - 4\sqrt{x}}{h}\\ &= (-5)\,(2\,x+h) + \frac{4}{\sqrt{x+h}+\sqrt{x}}.\end{aligned}$$

Jetzt bilde ich den Grenzwert für $h \to 0$, also den Differenzialquotienten, indem ich – salopp gesprochen – 0 für h in die Differenzenquotienten einsetze:

$$\begin{aligned}\lim_{h\to 0}\frac{f(x+h)-f(x)}{h} &= \lim_{h\to 0}\left((-5)\,(2x+h)+\frac{4}{\sqrt{x+h}+\sqrt{x}}\right)\\ &= (-5)\,(2x)+\frac{4}{\sqrt{x}+\sqrt{x}} = -10\,x+\frac{2}{\sqrt{x}}.\end{aligned}$$

Das darf ich, denn ich habe aufgepasst, dass das Intervall $[a, b]$ die Stelle $x = 0$ nicht enthält. Mein Ergebnis ist hier, dass f ableitbar ist, denn f ist an allen betrachteten Stellen $x \in [a, b]$ ableitbar und es gilt: $f'(x) = -10\,x + 2/\sqrt{x}$. ∎

Wie Sie anhand dieses Beispiels erkennen, habe ich hier im Wesentlichen die Rechentricks verwendet, die ich weiter oben schon ähnlich für die quadratische Potenzfunktion und die Wurzelfunktion benutzt hatte. Es stellt sich an dieser Stelle die berechtigte Frage, ob ich nicht viel *einfacher hätte vorgehen* können, indem ich das, was ich bereits über die Ableitung dieser beiden – sagen wir *elementaren* – Funktionen gelernt habe, direkter verwendet hätte. Genau das habe ich nun vor. Ich werde Ihnen also im Folgenden zeigen, dass es oftmals reicht, einen viel *einfacheren Weg* zu bestreiten, um die Ableitung einer Funktion zu bestimmen. Ich stelle Ihnen jetzt eine *Maschinerie* von **Ableitungsregeln** vor, die es ermöglicht, für große Klassen von *Standardfunktionen* die Ableitung zu bestimmen, *ohne explizit* den zuweilen mühseligen Weg über Differenzen- und Differenzialquotienten zu gehen oder sich gar nochmals tiefere Gedanken über die Existenz der Letztgenannten zu machen.

Erinnern Sie sich noch an mein erstes ernsthaftes Beispiel für die Bestimmung der Ableitung einer Funktion? Stimmt – das war die in Abbildung 6.3 dargestellte Identität $f(x) = x$, deren Ableitung $f'(x) = 1$ ist. Oben hatte ich Ihnen dann gesagt, dass ich die etwas schwierigere Funktion $f(x) = 3\,x$ ähnlich behandeln kann und erhielt: $f'(x) = 3$. Etwas schnuddelig, aber korrekt kann ich das wie folgt hinschreiben: $f'(x) = (3\,x)' = 3\,(x)' = 3 \cdot 1 = 3$. Damit zeige ich Ihnen, dass ich für diese Funktion erst den **Faktor** 3 *ausklammern* kann (zweites „="-Zeichen) und dann lediglich verwende, was ich bereits über die Ableitung der Identität gelernt habe.

Beispiel 6.4:
Ein mehr allgemeines Beispiel erhalte ich, wenn ich eine Potenzfunktion mit einem reellen Faktor a_i multipliziere: $f(x) = a_i\,x^i$. Erinnere ich mich an die Ableitung der Potenzfunktion, so kann ich genauso vorgehen: $f'(x) = (a_i x^i)' = a_i\,(x^i)' = a_i\,i x^{i-1} = i a_i\,x^{i-1}$. ∎

Beispiel 6.5:
Ein weiteres Beispiel ist die Funktion $f(x) = 4\sqrt{x}$, wobei ich davon ausgehe, dass x aus $[a, b]$ mit $a > 0$ ist. Die Ableitung von diesem f rechne ich wie folgt aus: $f'(x) = (4\sqrt{x})' = 4\,(\sqrt{x})' = 4\,(1/(2\sqrt{x})) = 2/\sqrt{x}$. Hier habe ich erst den Faktor 4 ausgeklammert (zweites „="-Zeichen) und dann lediglich verwendet, was ich bereits über die Ableitung der Wurzelfunktion gelernt habe. ∎

Dieser kleine Trick mit dem Ausklammern klappt immer, wenn ich eine ableitbare Funktion f mit einem reellen Faktor λ (sprich: „Lambda“) multipliziere und die so entstehende Funktion $\lambda \cdot f$, festgelegt durch $(\lambda \cdot f)(x) = \lambda \cdot f(x)$, ableiten möchte. Schauen Sie, warum dies so ist. Erst bilde ich die Differenzenquotienten

$$\frac{(\lambda \cdot f)(x+h) - (\lambda \cdot f)(x)}{h} = \frac{\lambda\, f(x+h) - \lambda\, f(x)}{h} = \lambda\, \frac{f(x+h) - f(x)}{h}$$

und dann den Differenzialquotienten

$$\begin{aligned} \lim_{h \to 0} \frac{(\lambda \cdot f)(x+h) - (\lambda \cdot f)(x)}{h} &= \lim_{h \to 0} \lambda\, \frac{f(x+h) - f(x)}{h} \\ &= \lambda \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lambda\, f'(x), \end{aligned}$$

wobei ich beim vorletzten „=“-Zeichen ausnutze, dass ich hier den konstanten Faktor λ vor den Grenzwert schreiben darf (denn dieser ist eine feste Zahl und hat somit nichts mit h zu tun), und beim letzten „=“-Zeichen davon profitiere, dass f in x ableitbar ist.

Erste Ableitung des Faktorprodukts $\lambda \cdot f$ – Faktorregel

Ist f eine ableitbare Funktion und λ eine reelle Konstante, so ist das **Faktorprodukt** $\lambda \cdot f$ eine ableitbare Funktion und die erste Ableitung $(\lambda \cdot f)'$ von $\lambda \cdot f$ ist für alle x festgelegt durch: $(\lambda \cdot f)'(x) = \lambda\, f'(x)$.

Erinnern Sie sich noch an Übungsaufgabe 6.3? Dort sollten Sie durch Verwendung der Differenzenquotienten zeigen, dass $f'(x) = 2\,x + 1$ die erste Ableitung von $f(x) = x^2 + x$ ist. Bestimmt haben Sie die Aufgabe inzwischen souverän gelöst. Zur Belohnung verrate ich Ihnen jetzt, dass man dies noch deutlich einfacher machen kann, indem man *direkt* benutzt, was ich Ihnen oben über die Ableitung der Identität und der quadratischen Potenzfunktion erzählt habe. Diese beiden elementaren Funktionen tauchen als Summanden von f auf. Leite ich diese jetzt einzeln ab und summiere ich danach die beiden Ableitungen auf, so erhalte ich nämlich ebenfalls f' – etwas schnuddelig kann ich das wie folgt hinschreiben:

$$f'(x) = (x^2 + x)' = (x^2)' + (x)' = 2\,x + 1.$$

Damit habe ich Ihnen (durch das zweite „=“-Zeichen) gezeigt, dass das *Auseinanderziehen von Summen* beim Ableiten erlaubt ist: Die Ableitung der Summe ist hier also die Summe der Ableitungen. Dasselbe kann ich auch bei obigem Beispiel $f(x) = -5\,x^2 + 4\,\sqrt{x}$ anwenden, wobei ich noch die eben genannte Faktorregel benutze:

$$f'(x) = (-5\,x^2)' + (4\,\sqrt{x})' = (-5)(x^2)' + 4\,(\sqrt{x})' = (-5)\,2x + \frac{4}{2\,\sqrt{x}} = -10\,x + 2\,x^{-1/2}.$$

Ich finde, das geht einfacher als direkt über die Differenzen- und Differenzialquotienten zu argumentieren. Versuchen Sie doch auch einmal vorab ein Beispiel:

Übungsaufgabe 6.5:
Bestimmen Sie, wie in den eben gezeigten Beispielen, die erste Ableitung von $f(x) = 3\,\cos(x) - 2\,\sin(x)$. ■

Dieser Trick mit dem Auseinanderziehen der Summe klappt immer, wenn ich zwei ableitbare Funktionen f und g habe und die Summe der beiden Funktionen $f + g$, festgelegt durch $(f + g)(x) = f(x) + g(x)$, ableiten möchte. Schauen Sie, warum dies so ist: Erst bilde ich die Differenzenquotienten

$$\begin{aligned}\frac{(f+g)(x+h)-(f+g)(x)}{h} &= \frac{f(x+h)+g(x+h)-(f(x)+g(x))}{h}\\ &= \frac{f(x+h)-f(x)}{h}+\frac{g(x+h)-g(x)}{h}\end{aligned}$$

und dann den Differenzialquotienten

$$\lim_{h\to 0}\frac{(f+g)(x+h)-(f+g)(x)}{h} = \lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}+\frac{g(x+h)-g(x)}{h}\right)$$
$$= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}+\lim_{h\to 0}\frac{g(x+h)-g(x)}{h} = f'(x)+g'(x),$$

wobei ich beim vorletzten „="-Zeichen schlicht benutze, dass ich hier die Summen von Grenzwerten auseinander ziehen darf, und beim letzten „="-Zeichen davon profitiere, dass f und g ableitbar sind.

Erste Ableitung der Summe von Funktionen $f + g$ – Summenregel

Sind f und g ableitbare Funktionen, so ist deren **Summe** $f + g$ eine ableitbare Funktion und die erste Ableitung $(f + g)'$ von $f + g$ ist für alle x festgelegt durch: $(f + g)'(x) = f'(x) + g'(x)$.

Zuweilen wird nun noch die **Differenzregel** formuliert: Schreibe ich die Differenz $f - g$ zweier ableitbarer Funktionen f und g als Summe von f mit dem Faktorprodukt $(-1) \cdot g$, so erhalte ich aus den beiden obigen Regeln (Faktor- und Summenregel):

$$(f-g)' = (f+(-1)\cdot g)' = f' + ((-1)\cdot g)' = f' + (-1)\,g' = f' - g'.$$

Somit ist die Ableitung der Differenz von zwei ableitbaren Funktionen die Differenz der Ableitungen dieser Funktionen.

Die bisher behandelten, einfachen Ableitungsregeln erlauben bereits die Bestimmung der Ableitung für relativ große Klassen von Funktionen. Wenn ich mir beispielsweise noch einmal lineare Funktionen $p(x) = c\,x + d$ anschaue (auch **Polynomfunktionen vom Grad eins**, bzw. kurz **Polynom vom Grad eins**, genannt), so kann ich alternativ zu dem, was ich Ihnen ziemlich am Anfang dieses Abschnitts gezeigt habe, die Ableitung wie folgt ausrechnen: $p'(x) = (c\,x)' + (d\,1)' = c\,(x)' + d\,(1)' = c\,1 + d\,0 = c$. Sie sehen, es kommt wieder die Konstante c heraus. Ähnlich kann ich nun aber auch vorgehen, falls ich es mit etwas schwierigeren Funktionen zu tun habe, welche aus *mehr als zwei Summanden* bestehen:

Beispiel 6.6:
Ein Beispiel ist das folgende **Polynom vom Grad fünf**: $p(x) = 3\,x^5 - \sqrt{2}\,x^3 + \frac{8}{7}\,x^2 - 3$. Mehrfache Anwendung der Faktor- und Summenregel ergibt hier die Ableitung:

$$\begin{aligned} p'(x) &= (3\,x^5)' - (\sqrt{2}\,x^3)' + \left(\frac{8}{7}\,x^2\right)' - (3)' = 3\,(x^5)' - \sqrt{2}\,(x^3)' + \frac{8}{7}\,(x^2)' - 3\,(1)' \\ &= 3\,(5x^4) - \sqrt{2}\,(3x^2) + \frac{8}{7}\,(2x) - 3\,(0) = 15\,x^4 - 3\sqrt{2}\,x^2 + \frac{16}{7}\,x, \end{aligned}$$

wobei ich verwende, was ich oben über die Ableitungen von Potenzfunktionen gelernt habe: $(x^5)' = 5\,x^4$, $(x^3)' = 3\,x^2$, $(x^2)' = 2\,x$ und $(x^0)' = (1)' = 0$. ■

Beispiel 6.7:
Ein anderes Beispiel ist $f(x) = 3\,\exp(x) - 2\,\sin(x) + \cos(x)$. Hier treten die Exponentialfunktion und die Sinus- beziehungsweise Cosinusfunktion als Terme auf. Ich verwende nun meine oben erworbenen Kenntnisse über die Ableitung dieser elementaren Funktionen und rechne f' mit der Faktor- und Summenregel aus:

$$f'(x) = 3\,\exp'(x) - 2\,\sin'(x) + \cos'(x) = 3\,\exp(x) - 2\,\cos(x) - \sin(x).$$ ■

Besonders nah am Herzen liegen mir Polynomfunktionen. Ich zeige Ihnen zunächst, wie ich eine allgemeine **quadratische Polynomfunktion** (auch: **Polynom vom Grad zwei**) $p(x) = a_2\,x^2 + a_1\,x + a_0$ ableiten kann. Hier sind a_0, a_1 und a_2 feste reelle Zahlen. Mit der Faktor- und Summenregel leite ich diese so ab:

$$p'(x) = (a_2\,x^2)' + (a_1\,x)' + (a_0\,1)' = a_2\,(x^2)' + a_1\,(x)' + a_0\,(1)' = 2a_2\,x + a_1.$$

Hier habe ich, wie in dem obigen Beispiel mit der Polynomfunktion vom Grad fünf unser Wissen über die Ableitung der Potenzfunktionen verwendet. Ich erkenne somit, dass die Ableitung einer quadratischen Polynomfunktion immer eine lineare Polynomfunktion ist. Diese Vorgehensweise gelingt allgemein für **Polynomfunktionen vom Grad** n (kurz: **Polynom vom Grad n**).

Erste Ableitung von Polynomfunktionen p

Jede **Polynomfunktion p vom Grad** n, $p(x) = a_n\,x^n + \ldots + a_i\,x^i + \ldots a_1\,x + a_0$ ist ableitbar, und die erste Ableitung p' von p ist das Polynom vom Grad $n-1$, das für alle x wie folgt festgelegt ist:

$$p'(x) = n\,a_n\,x^{n-1} + \ldots + i\,a_i\,x^{i-1} + \ldots + a_1\,.$$

Das Ableiten von Polynomen sollten Sie können. Deshalb schlage ich vor, dass Sie sich an der folgenden Übung versuchen:

Übungsaufgabe 6.6:
Berechnen Sie die erste Ableitung der folgenden Polynomfunktionen: a) $f(x) = -x^3 - 2\,x^2 + \frac{1}{8}\,x - \sqrt[3]{2}$ und b) $f_t(x) = t^2\,x^4 - 2t^4\,x^2 + t^2$. Bestimmen Sie zudem c) die Ableitung der folgenden Polynome in der Variablen t: $f_x(t) = t^2\,x^4 - 2t^4\,x^2 + t^2 = -2x^2\,t^4 + (x^4+1)\,t^2$, die ich durch Vertauschung der Bedeutung des **Parameters** t und der Variablen x in b) festgelegt habe. ■

Vielleicht tritt nun die folgende Frage auf: Gibt es eigentlich so etwas Ähnliches wie die Summenregel auch für das *Produkt von ableitbaren Funktionen*? Die erfreuliche Antwort ist: Ja. Aber Achtung: Leider ist – wie man vielleicht bei erster Betrachtung vermuten würde – die Ableitung eines Produkts von Funktionen im Allgemeinen *nicht* das Produkt der jeweiligen Ableitungen.

Beispiel 6.8:
Schauen Sie sich hierzu mein Beispiel der kubischen Potenzfunktion x^3 an. Diese Funktion kann ich als Produkt der quadratischen Potenzfunktion $f(x) = x^2$ und der Identität $g(x) = x$ schreiben: $x^3 = x^2\,x = f(x)\,g(x)$. Oben habe ich Ihnen erklärt, dass $f'(x) = 2\,x$ und $g'(x) = 1$ gelten. Das Produkt der Ableitungen von f und g ist somit $f'(x)\,g'(x) = (2x)\,1 = 2x$. Sie sehen: Was hier herauskommt, stimmt *keinesfalls* immer mit $3\,x^2 = (x^3)'$ überein. ■

Es kann eben leider nicht immer alles so butterweich durchgehen, wie man es sich vielleicht bei erster Betrachtung vorstellt. Um die Ableitung $(f \cdot g)'$ eines Produkts $f \cdot g$ von zwei ableitbaren Funktionen f und g, festgelegt durch $(f \cdot g)(x) = f(x)\,g(x)$, richtig zu bestimmen, muss ich mich ein bisschen mehr anstrengen. Ich zeige Ihnen jetzt, dass es gelingt, $(f \cdot g)'$ durch die als bekannt anzunehmenden Funktionen f, g, f' und g' auszudrücken. Schauen Sie, wie ich dies allgemein mache: Erst bilde ich die Differenzenquotienten

$$\begin{aligned}
\frac{(f\cdot g)(x+h) - (f\cdot g)(x)}{h} &= \frac{f(x+h)\,g(x+h) - f(x)\,g(x)}{h} \\
&= \frac{f(x+h)\,g(x+h) - f(x)\,g(x+h) + f(x)\,g(x+h) - f(x)\,g(x)}{h} \\
&= \frac{(f(x+h) - f(x))\,g(x+h) + f(x)\,(g(x+h) - g(x))}{h} \\
&= \frac{f(x+h) - f(x)}{h}\,g(x+h) + f(x)\,\frac{g(x+h) - g(x)}{h}.
\end{aligned}$$

Hier habe ich nach dem zweiten „="-Zeichen den Term $(f(x)\,g(x+h))/h$ abgezogen und gleich wieder hinzuaddiert, was am Ergebnis bekanntlich nichts ändert. Mit diesem kleinen Trick habe ich es erreicht, dass bei der Bildung der Differenzenquotienten des Produkts $f \cdot g$ die Differenzenquotienten der beiden Funktion f und g auftreten – sehen Sie das? Jetzt rechne ich den Differenzialquotienten aus:

$$\begin{aligned}
&\lim_{h\to 0} \frac{(f\cdot g)(x+h)-(f\cdot g)(x)}{h}\\
&\quad = \lim_{h\to 0}\left(\frac{f(x+h)-f(x)}{h}\,g(x+h)+f(x)\,\frac{g(x+h)-g(x)}{h}\right)\\
&\quad = \left(\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}\right) g(x)+f(x)\,\lim_{h\to 0}\frac{g(x+h)-g(x)}{h}\\
&\quad = f'(x)\,g(x)+f(x)\,g'(x).
\end{aligned}$$

Hier habe ich beim vorletzten „=“-Zeichen benutzt, dass ich hier die Summen von Grenzwerten auseinander ziehen darf und konstante Faktoren vor den Grenzwert geschrieben werden dürfen. Darüber hinaus habe ich dort auch davon profitiert, dass wegen der Ableitbarkeit von g

$$\lim_{h\to 0} g(x+h) = \lim_{h\to 0}\,(g(x)+g'(x)\,h) = g(x)$$

gilt. Beim letzten „=“-Zeichen habe ich oben zudem davon profitiert, dass f und g ableitbar sind.

Erste Ableitung des Produkts von Funktionen $f\cdot g$ – Produktregel

Sind f und g ableitbare Funktionen, so ist deren **Produkt** $f\cdot g$ eine ableitbare Funktion und die erste Ableitung $(f\cdot g)'$ von $f\cdot g$ ist für alle x festgelegt durch: $(f\cdot g)'(x) = f'(x)\,g(x)+g'(x)\,f(x)$.

Ich betrachte jetzt ein paar Beispiele der Produktregel:

Beispiel 6.8 (Fortsetzung):
Zunächst möchte ich unsere Diskussion von oben fortsetzen und noch einmal die kubische Potenzfunktion x^3 anschauen, deren Ableitung bekanntlich $3\,x^2$ ist. Setze ich nun wie oben $x^3 = f(x)\,g(x)$, wobei $f(x)=x^2$ und $g(x)=x$, so berechnet sich die Ableitung von $f\cdot g$ nach der Produktregel wie folgt: $(f\cdot g)'(x) = (x^2)'\,x+(x)'\,x^2 = (2x)\,x+1\,(x^2) = 2\,x^2+x^2 = 3\,x^2$. Jetzt stimmt's. Aber ich sollte hier nicht den Überblick verlieren – dies ist *nur ein Beispiel*, um die Funktionsweise der Produktregel zu illustrieren – die Potenzfunktion x^3 werde ich also weiterhin ohne Verwendung der Produktregel direkt als $3\,x^2$ ableiten. ■

Beispiel 6.9:
Ein ernsthaftes Beispiel ist die für $x\in[a,b]$ mit $a>0$ festlegbare Funktion

$$(f\cdot g)(x) = \overbrace{(x^2+x)}^{f(x)}\,\overbrace{(-5\,x^2+4\,\sqrt{x})}^{g(x)}\,.$$

Natürlich könnte ich hier brutal ausmultiplizieren und danach die Faktor- und Summenregel anwenden. Allerdings hätte ich dann das Problemchen, dass beispielsweise der Term $4\,x^2\sqrt{x} = 4\,x^{5/2}$ auftauchen würde und ich Ihnen (noch) nicht verraten habe, wie man diesen ableitet (... spätestens im nächsten Abschnitt – versprochen). Außerdem habe ich oben bereits glücklicherweise die Ableitungen der Funktionen $f(x) = x^2 + x$ und $g(x) = -5\,x^2 + 4\,\sqrt{x}$ ausgerechnet: $f'(x) = 2\,x + 1$ und $g'(x) = -10\,x + 2/\sqrt{x}$. Deshalb bietet sich die Produktregel an:

$$(f \cdot g)'(x) \;=\; \overbrace{(2x+1)}^{f'(x)}\,\overbrace{(-5\,x^2 + 4\,\sqrt{x})}^{g(x)} + \overbrace{(-10\,x + \frac{2}{\sqrt{x}})}^{g'(x)}\,\overbrace{(x^2 + x)}^{f(x)}.$$

Jetzt kann ich ausmultiplizieren und erhalte mit den Potenzregeln:

$$\begin{aligned}(f \cdot g)'(x) &= -10\,x^3 + 8\,x^{\frac{3}{2}} - 5\,x^2 + 4\,x^{\frac{1}{2}} - 10\,x^3 + 2\,x^{\frac{3}{2}} - 10\,x^2 + 2\,x^{\frac{1}{2}} \\ &= -20\,x^3 - 15\,x^2 + 10\,x^{\frac{3}{2}} + 6\,x^{\frac{1}{2}}.\end{aligned}$$ ■

Beispiel 6.10:
Das nächste Beispiel ist die Funktion $f(x) = \sin^2(x)$. Diese ist das Produkt der Sinusfunktion $\sin(x)$ mit sich selbst: $f(x) = (\sin \cdot \sin)(x) = \sin(x)\,\sin(x)$. Zum Glück weiß ich von oben, was die Ableitung der Sinusfunktion ist: $\sin'(x) = \cos(x)$. Damit kann ich also die Produktregel anwenden und erhalte:

$$\begin{aligned}f'(x) &= (\sin(x)\,\sin(x))' = \sin'(x)\,\sin(x) + \sin'(x)\,\sin(x) \\ &= 2\,\sin'(x)\,\sin(x) = 2\,\cos(x)\,\sin(x).\end{aligned}$$ ■

Ein bisschen Training kann jetzt nicht schaden:

Übungsaufgabe 6.7:
Verwenden Sie die Produktregel zur Berechnung der Ableitung der folgenden Funktionen: a) $f(x) = -x^3\,(x^2 - 4x + 2)$, b) $f(x) = (x^2 + x)\,\cos(x)$, c) $f(x) = x^2\,e^x$, d) $f(x) = \cos(x)\,e^x$ und e) $f(x) = \cos^2(x)$. ■

Ein nettes Beispiel zur Anwendung der Produktregel ist die Bestimmung der Ableitung f' von $f(x) = 1/x = x^{-1}$ für $x \neq 0$. Um diese zu bestimmen, muss ich erst einmal erkennen, dass $x\,f(x) = x\,(1/x) = 1$ gilt. Die Ableitung der rechten Seite verschwindet hier: $(1)' = 0$. Jetzt leite ich die linke Seite (Produkt aus Identität und f) mit der Produktregel ab:

$$(x \cdot f)'(x) = (x)'\,x^{-1} + (x^{-1})'\,x = 1\,x^{-1} + (x^{-1})'\,x = x^{-1} + (x^{-1})'\,x.$$

Ich will wissen, was $f'(x) = (x^{-1})'$ ist. Ich weiß jetzt bereits, dass $x^{-1} + (x^{-1})'\,x = 0$ gilt. Das Letztgenannte forme ich nun um, indem ich zunächst x^{-1} auf die rechte Seite bringe: $(x^{-1})'\,x = -x^{-1}$ und danach durch x teile: $(x^{-1})' = (-x^{-1})\,(1/x)$. Es gilt aber $1/x = x^{-1}$ und somit erhalte ich

$$f'(x) = (x^{-1})' = (-x^{-1})\,x^{-1} = -x^{-1-1} = -x^{-2} = \tfrac{-1}{x^2}.$$

Auch hier gilt somit (für $x \neq 0$) die Grobregel zur Bildung der ersten Ableitung, die ich Ihnen für die Potenzfunktion p_i genannt habe: $(x^{-1})' = (-1)\,x^{-2}$. Es ist schon etwas erstaunlich, für wie viele Funktionen unterschiedlichster Art die Grobregel anwendbar ist.

Schön, dass es die Produktregel gibt. Aber was mache ich, wenn jemand zur Tür hereinkommt und die Frage stellt: „Was ist die Ableitung von $\sin^{27}(x)$?“ Oder: „Ich hätte gerne die Ableitung von $\exp(\sin(x)) = e^{\sin(x)}$ gewusst!“ Alle bisherigen Ableitungsregeln helfen uns ebenfalls herzlich wenig, wenn uns jemand partout dazu zwingt, die Ableitung von $(3\,x^2\!+\!1)^{71}$ zu bilden – hier auszumultiplizieren, um dann die Faktor- und Summenregel anzuwenden, wäre einfach viel zu aufwendig, während die Verwendung der Produktregel ebenfalls schnell unübersichtlich werden würde. Ich verrate Ihnen jetzt, wie man die Ableitung **verketteter** Funktionen (siehe Kapitel 2) berechnen kann, ohne auf die Details der genauen Herleitung einzugehen. Irgendetwas muss ja auch noch für später übrig bleiben.

Erste Ableitung verketteter Funktionen $f \circ g$ – Kettenregel

Sind f und g ableitbare Funktionen, so ist die **Verkettung $f \circ g$ von f mit g** eine ableitbare Funktion und die erste Ableitung $(f \circ g)'$ von $f \circ g$ ist für alle x festgelegt durch: $(f \circ g)'(x) = g'(x)\, f'(g(x))$.

Am besten verstehen Sie das vermutlich durch Beispiele:

Beispiel 6.11:
Ich betrachte zunächst $(3\,x^2\!+\!1)^{71}$. Viel einfacher wäre dieses Beispiel, wenn der Term in der Klammer, also $3\,x^2+1$, *nur eine Variable $\bar{x}$ wäre*, denn die Potenzfunktion $f(\bar{x}) = \bar{x}^{71}$ kann ich locker ableiten: $f'(\bar{x}) = 71\,\bar{x}^{70}$. Aber leider ist dies hier nicht so einfach: Der Term in der Klammer ist *selbst eine Funktion*, nämlich $g(x) = 3\,x^2 + 1$. Glücklicherweise weiß ich aber immerhin, wie ich diese Funktion ableiten kann: $g'(x) = 6\,x$. Die Verkettung von f mit g ist $(f \circ g)(x) = (3\,x^2+1)^{71}$. Ich erhalte diese also, indem ich die Variable $\bar{x}$ der Funktion f *durch die Funktion g ersetze*: $\bar{x} = g(x)$. Jetzt sagt mir die Kettenregel, wie ich die Ableitung von $f \circ g$ finde: Ich muss die Ableitung von g – also $g'(x) = 6\,x$ – multiplizieren mit der Ableitung von f – also $f'(\bar{x}) = 71\,\bar{x}^{70}$ – *wobei dort die Variable $\bar{x}$ durch $g(x)$ ersetzt werden soll.* Mache ich dies so, so erhalte ich:

$$(f \circ g)'(x) = \overbrace{6x}^{g'(x)}\ \overbrace{71\,(\underbrace{3\,x^2+1}_{\bar{x}=g(x)})^{70}}^{f'(\bar{x})=f'(g(x))} = 426\,x(3\,x^2+1)^{70}. \qquad \blacksquare$$

Beispiel 6.12:
Das nächste Beispiel ist $\sin^{27}(x)$. Einfacher wäre das hier, wenn der Term in der Klammer, also $\sin(x)$, nur eine Variable $\bar{x}$ wäre, denn die Potenzfunktion $f(\bar{x}) = \bar{x}^{27}$ kann ich locker ableiten: $f'(\bar{x}) = 27\,\bar{x}^{26}$. Aber leider ist dies hier wiederum nicht so: Der Term in der Klammer ist nämlich die Funktion $g(x) = \sin(x)$, für die ich aber immerhin weiß, wie ich sie ableiten kann: $g'(x) = \cos(x)$. Die Kettenregel zeigt mir, wie ich die Ableitung von $(f \circ g)(x) = \sin^{27}(x)$ finde: Ich muss die Ableitung g' von g multiplizieren mit der Ableitung f' von f, wobei dort die Variable $\bar{x}$ durch $g(x)$

ersetzt werden soll:

$$(f \circ g)(x) = \overbrace{\cos(x)}^{\sin'(x)} \overbrace{27\,(\underbrace{\sin(x)}_{\bar{x}=\sin(x)})^{26}}^{(\bar{x}^{27})'} = 27\,\cos(x)\,\sin^{26}(x). \qquad \blacksquare$$

Übrigens kann ich das oben im Rahmen der Produktregel diskutierte Beispiel $\sin^2(x)$ genauso behandeln: $(f \circ g)(x) = \sin^2(x)$, wobei $f(\bar{x}) = \bar{x}^2$ und $g(x) = \sin(x)$. Dann erhalte ich – $g'(x) = \cos(x)$ und $f'(\bar{x}) = 2\,\bar{x}$ verwendend – mit der Kettenregel $(f \circ g)(x) = \cos(x)\,2\,(\sin(x))^1 = 2\,\cos(x)\sin(x)$, und das Ergebnis stimmt natürlich mit dem, was ich oben ausgerechnet habe, überein.

Beispiel 6.13:
Die Funktion $(f \circ g)(x) = \exp(\sin(x)) = e^{\sin(x)}$ kann ich auch mit der Kettenregel ableiten. Ich setze hier $f(\bar{x}) = e^{\bar{x}}$ und $g(x) = \sin(x)$. Die entsprechenden Ableitungen kenne ich: $f'(\bar{x}) = e^{\bar{x}}$ und $g'(x) = \cos(x)$. Somit ergibt sich: $(f \circ g)'(x) = \cos(x)\,e^{\sin(x)}$.

$\blacksquare$

Jetzt sind Sie aber wieder dran.

Übungsaufgabe 6.8:
Berechnen Sie die Ableitung der folgenden Funktionen: a) $f(x) = (x^2 - 4x + 2)^{16}$, b) $f(x) = \exp(x^2 - x)$, c) $f(x) = x^{n/m}$, wobei $x > 0$ und $n, m \in \mathbb{N}$, d) $f(x) = (1 - x^2)\,(x^2 + 1)^{-2}$. $\blacksquare$

Vielleicht ist eben die Frage aufgetreten, wie es mit der Ableitung des Quotienten f/g ableitbarer Funktionen f und g, festgelegt durch $(f/g)(x) = f(x)/g(x)$, aussieht. Zunächst muss ich hier natürlich davon ausgehen, dass für alle betrachteten x der Nenner nicht verschwindet: $g(x) \neq 0$, denn sonst würde die Funktion f/g dort keinen Sinn machen, denn durch 0 darf ich bekanntlich *nie nie nie* teilen. Glücklicherweise kann ich dann die Ableitung des Quotienten $f/g = f \cdot (g)^{-1}$ mithilfe der Funktionen f, g, f' und g' bestimmen, indem ich die Produkt- und Kettenregel *kombiniere*. Das sehe ich folgendermaßen: Die Produktregel ergibt erst einmal

$$(f/g)' = (f \cdot (g)^{-1})' = f' \cdot (g)^{-1} + ((g)^{-1})' \cdot f.$$

Hier kenne ich schon $f', (g)^{-1} = 1/g$ und f – nur wie kann ich $((g)^{-1})'$ mithilfe von g und g' hinschreiben? Dazu leite ich jetzt $(g)^{-1}$ unter Verwendung der Kettenregel ab. Hierzu setze ich $\tilde{f}(\bar{x}) = \bar{x}^{-1}$ und verwende, dass $\tilde{f} \circ g$ tatsächlich $(g)^{-1}$ ergibt: $(\tilde{f} \circ g)(x) = (g(x))^{-1}$. Das sehe ich, indem ich $\bar{x}$ durch $g(x)$ ersetze. Zum Glück kenne ich die Ableitung von $\tilde{f}$ von unserer Diskussion direkt nach Übungsaufgabe 6.7: $\tilde{f}'(\bar{x}) = -\bar{x}^{-2}$. Somit erhalte ich nach der Kettenregel an allen Stellen x mit $g(x) \neq 0$:

$$((g)^{-1})'(x) = (\tilde{f} \circ g)'(x) = g'(x)\,\tilde{f}'(g(x)) = g'(x)\,(-(g(x))^{-2}) = -g'(x)\,(g(x))^{-2}.$$

Jetzt kenne ich auch $((g)^{-1})' = -g' \cdot (g)^{-2}$ und darf dies oben auch einsetzen. Die Kombination von Produkt- und Kettenregel ergibt somit

$$(f/g)' = f' \cdot (g)^{-1} + (-g' \cdot (g)^{-2}) \cdot f = f'/g - (g' \cdot f)/g^2 = (f' \cdot g - g' \cdot f)/g^2,$$

wobei ich im letzten Schritt den ersten Term mit g erweitert habe. Das war nicht ganz trivial (math.: einfach). Zur Belohnung weiß ich jetzt aber das Folgende:

Erste Ableitung des Quotienten von Funktionen f/g – Quotientenregel

Sind f und g ableitbare Funktionen, so ist deren **Quotient** f/g eine für alle x mit $g(x) \neq 0$ ableitbare Funktion und die erste Ableitung $(f/g)'$ von f/g ist für diese x festgelegt durch: $(f/g)'(x) = \big(f'(x)\,g(x) - g'(x)\,f(x)\big)/(g(x))^2$.

Klassiker für die Anwendung der Quotientenregel sind **gebrochen-rationale Funktionen** $f = p/q$, wobei p und q Polynomfunktionen sind.

Beispiel 6.14:
Ein Beispiel ist die Funktion $f(x) = (x^2 + x)/(x^4 + 1)$, also $p(x) = x^2 + x$ und $q(x) = x^4 + 1$. Für alle $x \in \mathbb{R}$ gilt $q(x) = x^4 + 1 \geq 1 > 0$ – der Nenner verschwindet somit in diesem Beispiel zum Glück nie. Zudem sind p und q ableitbar und die erste Ableitung dieser Polynomfunktionen sind wie folgt: $p'(x) = 2\,x + 1$ und $q'(x) = 4\,x^3$. Nach der Quotientenregel berechne ich damit

$$\begin{aligned} f'(x) &= \frac{p'(x)\,q(x) - q'(x)\,p(x)}{(q(x))^2} = \frac{(2\,x+1)\,(x^4+1) - 4\,x^3\,(x^2+x)}{(x^4+1)^2} \\ &= \frac{2\,x^5 + x^4 + 2\,x + 1 - 4\,x^5\; - 4x^4}{x^8 + 2\,x^4 + 1} = \frac{-2\,x^5 - 3\,x^4 + 2\,x + 1}{x^8 + 2\,x^4 + 1}. \end{aligned}$$ ■

Beispiel 6.15:
Spaß macht es meiner Ansicht nach, die Funktion $k(x) = (x^2 + x - 1)/e^x$ abzuleiten. Ich kann dies für alle x machen, denn der Nenner $g(x) = e^x$ ist überall ableitbar – $g'(x) = e^x$ – und zudem nirgends 0. Außerdem ist der Zähler $f(x) = x^2+x-1$ ebenfalls überall ableitbar: $f'(x) = 2x + 1$. Ich rechne die Ableitung $k' = (f' \cdot g - g' \cdot f)/g^2$ von k mit der Quotientenregel aus:

$$k'(x) = \frac{(2x+1)\,e^x - e^x\,(x^2+x-1)}{(e^x)^2} = \frac{(2x+1-x^2-x+1)\,e^x}{e^{2x}} = \frac{-x^2+x+2}{e^x}.$$

Hier habe ich e^x ausgeklammert und danach gekürzt: $e^x/e^{2x} = 1/e^x$. ■

Beispiel 6.16:
Ein etwas mehr exotisches Beispiel bildet die **Tangensfunktion** $\tan(x) = \sin(x)/\cos(x)$ für $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Bekanntlich ist $\cos(x) \neq 0$ für diese x. Zudem ist die Ableitung des Zählers $\sin'(x) = \cos(x)$, sowie die Ableitung des Nenners $\cos'(x) = -\sin(x)$. Dies verwende ich jetzt, um mit der Quotientenregel die Ableitung von $\tan(x)$ für $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ auszurechnen:

$$\begin{aligned} \tan'(x) &= \frac{\sin'(x)\cos(x) - \cos'(x)\sin(x)}{\cos^2(x)} = \frac{\cos(x)\cos(x) - (-\sin(x))\sin(x)}{\cos^2(x)} \\ &= \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{\cos^2(x)}{\cos^2(x)} + \frac{\sin^2(x)}{\cos^2(x)} = 1 + \tan^2(x). \end{aligned}$$ ■

Übungsaufgabe 6.9:
Berechnen Sie mit der Quotientenregel die Ableitung der folgenden Funktionen: a) $f(x) = (x-4)/(2\,x^2+1)$, b) $f(x) = (x+4)/(2\,x^2-1)$, c) $f(x) = \sin(x)/x$, d) $f(x) = (1-x^2)/(x^2+1)^2$ und e) $f(x) = (-x^2+x+2)/e^x$. Für welche x darf man jeweils überhaupt ableiten? Vergleichen Sie d) mit Aufgabe 6.8d) – fällt Ihnen etwas auf? ■

6.2 Anwendungen von Ableitungen und Kurvendiskussion

Ich möchte jetzt mit Ihnen über spezielle *Kurven diskutieren* – genauer gesagt möchte ich Ihnen zeigen, wie man den Graphen vorgegebener Funktionen bestimmen kann, um diese zu veranschaulichen. Bei der *Kurvendiskussion* kommt es insbesondere darauf an, charakteristische Punkte des Graphen zu identifizieren und dessen Verhalten zwischen und außerhalb von solchen Punkten zu beschreiben. Für eine solche Analyse von Funktionen wende ich Ableitungen an – diese sind hierfür ein mächtiges Hilfsmittel.

Wie ich Ihnen eingangs des letzten Abschnitts versprochen habe, komme ich hier zunächst noch einmal auf das Beispiel der Wanderung im „Pfälzer Wald“ zurück, die in Abbildung 6.1 als Graph einer Funktion f dargestellt ist. Betrachte ich diese Abbildung, so stelle ich fest, dass der Wert der Steigung für dieses f für alle $x \in [a, x_3) \cup (x_4, x_5)$ (das heißt $a \leq x < x_3$ und $x_4 < x < x_5$) positiv ist – dort geht es jeweils **monoton** aufwärts. Andererseits ist der Wert der Steigung negativ für alle $x \in (x_3, x_4) \cup (x_5, b]$ (das heißt $x_3 < x < x_4$ und $x_5 < x \leq b$) – dort geht es jeweils **monoton** abwärts. Der Wert der Steigung ist darüber hinaus 0 an den Gipfelpunkten $(x_3|f(x_3))$ und $(x_5|f(x_5))$, ebenso wie im Talpunkt $(x_4|f(x_4))$ – dort geht es jeweils weder auf- noch abwärts. Außerdem beobachte ich, dass *global betrachtet* der höchste Punkt der Wanderung an der Stelle x_3 erreicht ist: $f(x) \leq f(x_3)$ für alle $x \in [a, b]$ – der Höhenwert ist dort **global maximal**. *Global betrachtet* habe ich die geringste Höhe zu Beginn der Wanderung: $f(x) \geq f(a)$ für alle $x \in [a, b]$ – der Höhenwert ist dort **global minimal**. Insbesondere ist der Wert $f(x_4)$ im Talpunkt $(x_4|f(x_4))$ *global betrachtet* größer als der Wert an der Stelle a: $f(x_4) > f(a)$. Jedoch *lokal betrachtet* ist dieser Wert $f(x_4)$ am kleinsten: In einer kleinen Umgebung von x_4, sagen wir in einem Intervall der Form $(x_4 - c, x_4 + c)$, wobei $c > 0$ klein genug gewählt ist, gilt nämlich $f(x) \geq f(x_4)$ für alle $x \in (x_4 - c, x_4 + c)$ – dieser Höhenwert ist lediglich **lokal minimal**.

Sie wissen ja inzwischen, dass ich den Wert der Steigung von f an einer Stelle x auch erste Ableitung von f in x nenne und diese mit $f'(x)$ bezeichne. Im Folgenden werde ich Ihnen zeigen, dass ich für ableitbare Funktionen – grob gesprochen – dieselbe Diskussion wie eben führen kann. So werde ich Ihnen beispielsweise berichten, wie Sie die Stellen größter und kleinster Werte ableitbarer Funktion im globalen und lokalen

Sinne finden können. Zudem werde ich Ihnen den Zusammenhang zwischen der Art des Anstiegs und dem Vorzeichen der Ableitung darstellen.

Ich denke, dies ist eigentlich ein ganz guter Einstieg und deshalb beginne ich mit dem Zusammenhang des Vorzeichens der Ableitung zur **Monotonie** von Funktionen. Umgangssprachlich würde man hier möglicherweise denken, dass es hierbei um Funktionen f geht, die *monoton* – um nicht zu sagen besonders langweilig – sind. Ich werde Ihnen aber gleich zeigen, dass dies eigentlich nicht unbedingt auf solche Funktionen f – eher schon aber auf deren erste Ableitung f' – zutrifft, denn deren Vorzeichen ändert sich, wie Sie gleich sehen werden, für monotone Funktionen f nicht.

In Kapitel 2 wurde Ihnen erklärt, dass eine Funktion f **monoton steigend** ist, falls aus $x_1 < x_2$ stets folgt, dass $f(x_1) \leq f(x_2)$. ***Streng*** **monoton steigend** ist eine Funktion, falls aus $x_1 < x_2$ stets folgt, dass $f(x_1) < f(x_2)$. Analog wurden dort auch die Begriffe **monoton fallend** und ***streng*** **monoton fallend** erklärt. In Abbildung 6.9 finden Sie einige Graphen (streng) monotoner Funktionen.

Natürlich sind (streng) monotone Funktionen nicht immer ableitbar. Die (Quadrat-)Wurzelfunktion ist nämlich ein Beispiel einer Funktion die, wie ich Ihnen im letzten Abschnitt gezeigt habe, in $x = 0$ und damit insgesamt nicht ableitbar ist. Andererseits jedoch ist diese streng monoton steigend – um dies zu überprüfen, können Sie sich deren Graphen in Abbildung 6.5 und 6.9 (Mitte) intensiv anschauen oder nochmals in Kapitel 2 schmökern. Im Folgenden werde ich mich vor allem auf ableitbare Funktionen konzentrieren, denn für diese werde ich Ihnen einfache Kriterien zum Test auf (strenge) Monotonie darstellen können.

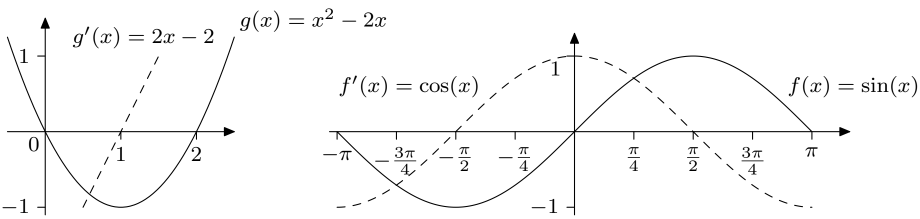

Abb. 6.7 Die Art der Monotonie hängt oftmals von den betrachteten Stellen x ab. Links: $g(x) = x^2 - 2x$ ist monoton fallend für $x \leq 1$ (dort: $g'(x) \leq 0$) und monoton steigend für $x \geq 1$ (dort: $g'(x) \geq 0$). Rechts: Die Sinusfunktion $f(x) = \sin(x)$ ist streng monoton steigend für $x \in [-\pi/4, \pi/4]$, denn dort gilt: $\sin'(x) = \cos(x) > 0$. Auch auf $[-\pi/2, \pi/2]$ betrachtet ist $\sin(x)$ streng monoton steigend, während diese Funktion auf $[-\pi, -\pi/2]$ streng monoton fallend ist – auf $[-\pi, \pi]$ ist $\sin(x)$ somit *nicht* monoton.

Um Ihnen ein Gefühl für den Zusammenhang des Vorzeichens der Ableitung und der Monotonie zu geben, betrachte ich beispielsweise eine monoton steigende, ableitbare Funktion f. Die Differenzenquotienten sind dann stets größer oder gleich 0:

$$\frac{f(x+h) - f(x)}{(x+h) - x} = \frac{f(x+h) - f(x)}{h} \geq 0.$$

Das sehe ich wie folgt: Für vorgegebenes x gilt $x < x + h$, *falls* ich $h > 0$ (klein genug) wähle. Setze ich im Geiste $x_1 = x$ und $x_2 = x + h$, so folgt, weil f monoton wachsend ist: $f(x) \leq f(x+h)$. Das kann ich auch, indem ich auf beiden Seiten der Ungleichung $f(x)$ abziehe, wie folgt hinschreiben: $f(x+h) - f(x) \geq 0$. In diesem Fall ist also der Zähler der Differenzenquotienten ≥ 0 und der Nenner $(x+h)-x = h > 0$. Die Differenzenquotienten selbst sind folglich ≥ 0. Betrachte ich andererseits den verbleibenden Fall $h < 0$, so drehen sich die betrachteten Ungleichheitszeichen um, denn nun ist $x+h < x$, und ich erhalte $f(x+h) - f(x) \leq 0$. Dann ist also der Zähler ≤ 0 und der Nenner $(x+h)-x = h < 0$ – die Differenzenquotienten selbst sind folglich wiederum ≥ 0. Die erste Ableitung $f'(x)$ von f in x hat nun keine andere Chance: Da diese als Grenzwert der Differenzenquotienten auftritt und diese stets größer oder gleich 0 sind, ist diese ebenfalls größer oder gleich 0:

$$f'(x) = \lim_{h \to 0} \overbrace{\left(\frac{f(x+h) - f(x)}{h} \right)}^{\geq 0} \geq 0.$$

Ich habe Ihnen somit gezeigt, dass die erste Ableitung monoton steigender, ableitbarer Funktionen stets größer oder gleich 0 ist. Glücklicherweise gilt auch die Umkehrung sowie eine Analogie für monoton fallende Funktionen.

Monotonie und Vorzeichen der ersten Ableitung

Eine ableitbare Funktion f ist *genau dann* monoton steigend, wenn die Ableitung f' von f für alle x nicht negativ ist: $f'(x) \geq 0$. Eine ableitbare Funktion f ist *genau dann* monoton fallend, wenn die Ableitung f' von f für alle x nicht positiv ist: $f'(x) \leq 0$.

Ich schaue mir Beispiele an:

Beispiel 6.17:
Wie wäre es mit $g(x) = x^2 - 2x$ für $x \geq 1$ (siehe Abbildung 6.7 (links))? Die Funktion nenne ich hier aus zwei Gründen g. Erstens zeige ich Ihnen, dass ich Funktionen nicht immer mit f bezeichnen muss. Zweitens werde ich diese Funktion g in einem späteren Beispiel noch einmal anschauen, und für dieses ist es, wie Sie dort sehen werden besser, dass ich diese Funktion nicht f nenne. Ich rechne nun flott die erste Ableitung von g aus: $g'(x) = 2\,x - 2$. Die Funktion g ist *für* $x \geq 1$ monoton steigend, denn für diese x gilt: $g'(x) = 2\,x - 2 \geq 2 \cdot 1 - 2 = 0$. Schaue ich mir dieselbe Funktionsvorschrift jetzt auch für $x \leq 1$ an, so stelle ich andererseits fest, dass dort $g'(x) = 2\,x - 2 \leq 0$ gilt – die Funktion g ist somit *für* $x \leq 1$ monoton fallend. ■

Die Art der Monotonie hängt also insbesondere davon ab, für welche Stellen x ich die Funktion g betrachte. Das ist nicht so außergewöhnlich – die Wanderung im „Pfälzer Wald“ hat Ihnen auch schon gezeigt, dass der Graph einer Funktion im Allgemeinen aus verschiedenen **Monotoniebereichen** besteht.

Beispiel 6.18:
Ein anderes Beispiel ist $f(x) = -\exp(x^3) = -e^{x^3}$ für $x \in \mathbb{R}$. Ich rechne die erste Ableitung mit der **Kettenregel** aus: $f'(x) = -3x^2\,e^{x^3}$. Weil $-3x^2 \leq 0$ und

$\exp(\bar{x}) = e^{\bar{x}} > 0$ gelten (unabhängig davon, was ich für $\bar{x}$ einsetze, also auch für $\bar{x} = x^3$), folgt $f'(x) \leq 0$ für alle $x \in \mathbb{R}$. Diese Funktion ist also monoton fallend. Hier muss ich erfreulicherweise keine Fallunterscheidung wie für g machen. ■

Sicher fragen Sie sich an dieser Stelle, ob es einen Zusammenhang der *strengen* Monotonie mit dem Vorzeichen der ersten Ableitung gibt. Hierzu kann ich Ihnen, ohne dass ich mich hier mit den etwas aufwendigeren Argumenten aufhalte, Folgendes berichten:

Strenge Monotonie und Vorzeichen der ersten Ableitung

Eine ableitbare Funktion f mit $f'(x) > 0$ für alle x ist streng monoton steigend. Eine ableitbare Funktion f mit $f'(x) < 0$ für alle x ist streng monoton fallend.

Natürlich ist die eben betrachtete Funktion $g(x) = x^2 - 2x$ *für* $x > 1$ *streng* monoton steigend, denn dort gilt $g'(x) = 2\,x - 2 > 0$. Ein mehr delikates Beispiel für eine streng monoton steigende Funktion ist $f(x) = \sin(x)$, für x aus $[-\pi/4, \pi/4]$ (siehe Abbildung 6.7 (rechts)). Um das zu sehen, bilde ich – wie im letzten Abschnitt gezeigt – die Ableitung $f'(x) = \cos(x)$ und stelle fest, dass $\cos(x) > 0$ für x aus $[-\pi/4, \pi/4]$ gilt. Sehen Sie, wie ungleich einfacher dies ist, als $\sin(x_1) < \sin(x_2)$ für alle $-\pi/4 \leq x_1 < x_2 \leq \pi/4$ auf eventuell anderem Wege nachzuweisen?

Ein mehr klassisches Beispiel ist die Exponentialfunktion $f(x) = \exp(x) = e^x$: Diese Funktion (siehe Abbildung 6.9(rechts)) ist streng monoton steigend, denn $f'(x) = e^x > 0$ für alle $x \in \mathbb{R}$.

Ich muss Sie an dieser Stelle noch darauf hinweisen, dass die Umkehrung der letzten Aussage im Allgemeinen nicht möglich ist: Es gibt nämlich *streng* monotone Funktionen, deren Ableitung lediglich $f'(x) \geq 0$ beziehungsweise $f'(x) \leq 0$ erfüllen.

Beispiel 6.19:
Das Polynom $f(x) = 1/3\,x^3$, $x \in \mathbb{R}$, ist ein solches Beispiel (siehe Abbildung 6.8). Ist $x_1 < x_2$, so folgt $f(x_1) = 1/3\,x_1^3 < 1/3\,x_2^3 = f(x_2)$ – diese Funktion f ist somit *streng* monoton steigend. Andererseits gilt für deren erste Ableitung: $f'(x) = x^2 \geq 0$. Hier ist zu beachten, dass „$\geq$“ *nicht* durch „$>$“ ersetzt werden kann, denn an der *speziellen Stelle* $x = 0$ berechne ich: $f'(0) = 0^2 = 0$. ■

Da ich mit Ihnen gerade so nett über *strenge* Monotonie plaudere, bietet es sich hier natürlich an, ein paar Worte über **umkehrbare Funktionen** zu verlieren, denn in Kapitel 2 wurde Ihnen erklärt, dass *streng monotone Funktionen stets umkehrbar* sind. Sie wissen also inzwischen auch wieder, dass eine Funktion f^{-1} **Umkehrfunktion** von f genannt wird, wenn die **Verkettung** $f^{-1} \circ f$ die **Identität** ergibt, also $f^{-1}(f(x)) = x$ für alle betrachteten x gilt. Der Einfachheit halber gehe ich jetzt davon aus, dass entweder $f'(x) > 0$ oder $f'(x) < 0$ für alle betrachteten x gilt. Nach dem, was im letzten Kasten steht, garantiert mir dies die strenge Monotonie von f und damit die Existenz der Umkehrfunktion f^{-1}. Nehme ich jetzt kurz einmal an, dass ich bereits die Ableitung von f kenne und hiermit die Ableitung $(f^{-1})'$ der Umkehrfunktion f^{-1}

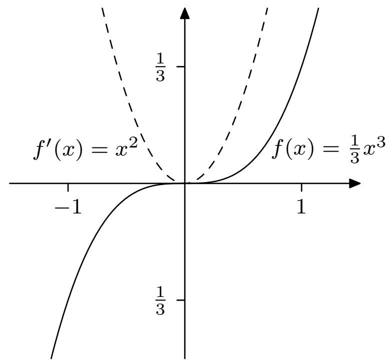

Abb. 6.8 Die Funktion $f(x) = 1/3\,x^3$ ist streng monoton – deren Ableitung $f'(x) = x^2$ (gestrichelt) verschwindet jedoch an der Stelle 0: $f'(0) = 0$. Die Stelle $x = 0$ ist trotzdem kein lokales Extremum von f, denn f' wechselt das Vorzeichen dort nicht.

von f bestimmen möchte, so kann ich dies durch Anwendung der **Kettenregel** tun. Ich leite hierzu in der Gleichung $f^{-1}(f(x)) = x$ die linke und die rechte Seite ab:

$$f'(x)\,(f^{-1})'(f(x)) = (x)'.$$

Schauen Sie im Zweifel noch einmal im vorherigen Abschnitt nach – hier habe ich jetzt $\bar{x} = f(x)$ gesetzt. Unter Beachtung von $f'(x) \neq 0$ kann ich nun beide Seiten durch $f'(x)$ teilen, und wenn ich mich dann noch an $(x)' = 1$ erinnere, so erhalte ich die folgende Aussage:

Erste Ableitung der Umkehrfunktion

Ist f eine ableitbare Funktion mit $f'(x) > 0$ oder $f'(x) < 0$ für alle x, so ist deren Umkehrfunktion f^{-1} ableitbar und die erste Ableitung $(f^{-1})'$ von f^{-1} ist festgelegt durch: $(f^{-1})'(\bar{x}) = 1/f'(x)$, wobei $\bar{x} = f(x)$ beziehungsweise $f^{-1}(\bar{x}) = x$.

Insbesondere erkenne ich, dass das Vorzeichen der Ableitung $(f^{-1})'$ der Umkehrfunktion f^{-1} mit dem Vorzeichen der Ableitung f' von f übereinstimmt. Ist somit f beispielsweise streng monoton steigend, so ist auch f^{-1} streng monoton steigend. Im Folgenden zeige ich Ihnen drei Beispiele von umkehrbaren Funktionen und die Berechnung der Ableitung der jeweils zugehörigen Umkehrfunktion. Die Graphen der Funktionen und deren Umkehrfunktion sehen Sie in Abbildung 6.9: Diese gehen bekanntermaßen durch Spiegelung an der **ersten Winkelhalbierenden** (dem Graphen der Identität, gestrichelte Linie) ineinander über.

Beispiel 6.20:
Einfache Beispiele umkehrbarer Funktionen sind lineare Funktionen $f(x) = c\,x+d$, die *nicht* konstant sind: $c \neq 0$. Ich gebe Ihnen die Umkehrfunktion direkt an: $f^{-1}(\bar{x}) = (1/c)\,\bar{x} - (d/c)$. Zur Sicherheit setze ich $\bar{x} = f(x)$ hier ein und berechne $f^{-1}(f(x)) = (1/c)\,f(x) - (d/c) = (1/c)\,(c\,x+d) - (d/c) = x$. Wenn ich nun die Ableitung der linearen Funktion f^{-1} ausrechnen möchte, so kann ich dies natürlich direkt tun: $(f^{-1})'(\bar{x}) =$

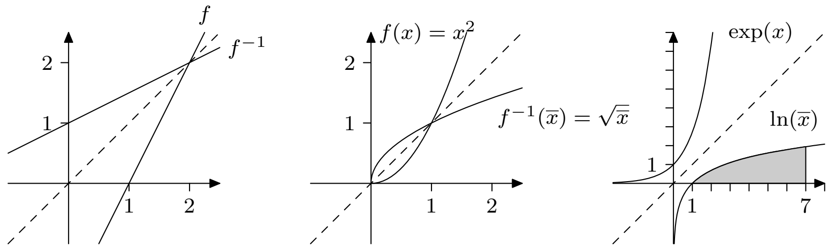

Abb. 6.9 Graphen von Funktionen und der zugehörigen Umkehrfunktion. Links: Die lineare Funktion $f(x) = 1/2\,x + 1$ besitzt die Umkehrfunktion $f^{-1}(\bar{x}) = 2\,\bar{x} - 2$. Mitte: Die Umkehrfunktion von $f(x) = x^2$ ist die Wurzelfunktion $f^{-1}(\bar{x}) = \sqrt{\bar{x}}$. Rechts: Die Exponentialfunktion $f(x) = \exp(x) = e^x$ besitzt den natürlichen Logarithmus $f^{-1}(\bar{x}) = \ln \bar{x}$ als Umkehrfunktion. Alle diese Funktionen sind streng monoton steigend.

$1/c$. Mit der obigen Aussage erhalte ich – $f'(x) = c$ verwendend – erwartungsgemäß dasselbe Ergebnis: $(f^{-1})'(\bar{x}) = 1/f'(x) = 1/c$. ■

Beispiel 6.21:
Ein weiterer Klassiker ist die quadratische Potenzfunktion $f(x) = x^2$ für $x \geq 0$ (siehe zum Beispiel Abbildung 6.4). Aus Kapitel 2 wissen Sie, dass deren Umkehrfunktion die (Quadrat-)Wurzelfunktion $f^{-1}(\bar{x}) = \sqrt{\bar{x}}$ ist: $\sqrt{x^2} = x^{2(1/2)} = x$ für alle $x \geq 0$. Die Wurzelfunktion ist allerdings nur für $\bar{x} > 0$ ableitbar. Möchte ich für $\bar{x} > 0$ deren Ableitung mithilfe der Ableitung $f'(x) = 2\,x$ von f ausrechnen, so kann ich das folgendermaßen tun:

$$(\sqrt{\bar{x}})' = (f^{-1})'(\bar{x}) = \frac{1}{f'(x)} = \frac{1}{2x} = \frac{1}{2\sqrt{\bar{x}}}\,,$$

und erhalte so auf neuem Weg ein Ergebnis, das ich schon aus dem vorherigen Abschnitt kenne. Hierbei habe ich beim letzten „="-Zeichen verwendet, dass $\bar{x} = x^2$, also $\sqrt{\bar{x}} = x$ ist. Beachten Sie, dass wegen $\bar{x} > 0$ auch $x > 0$ gilt. Für $x = 0$ klappt diese Vorgehensweise nicht, denn ich würde wegen $f'(0) = 0$ durch 0 teilen – keine allzu große Überraschung, dass hier etwas schief gehen muss, denn die Wurzelfunktion ist – wie ich Ihnen im letzten Abschnitt zeigte – in $\bar{x} = 0$ nicht ableitbar. ■

Beispiel 6.22:
Ein anderes Beispiel ist die Exponentialfunktion $\exp(x) = e^x$. Diese ist wegen $\exp'(x) = \exp(x) > 0$ auf ganz $\mathbb{R}$ umkehrbar. Die Umkehrfunktion von exp ist der **natürliche Logarithmus**: $\ln(\exp(x)) = x$ für alle x. Ich rechne jetzt mit meinen Kenntnissen über die Ableitung von exp (siehe vorheriger Abschnitt) die Ableitung von ln für $\bar{x} = \exp(x) > 0$ aus:

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

Dabei habe ich beim drittletzten „="-Zeichen benutzt, dass $x = \ln(\bar{x})$ gilt, und beim vorletzten „="-Zeichen davon profitiert, dass exp die Umkehrfunktion von ln ist, also auch $\exp(\ln(\bar{x})) = \bar{x}$ für alle $\bar{x} > 0$ gültig ist. ■

Soweit über den Zusammenhang von strenger Monotonie und der Bestimmung der Ableitung der Umkehrfunktion. Wie eingangs angekündigt, fahre ich nun fort, indem ich Ihnen zeige, wie Sie die Stellen größter und kleinster Werte ableitbarer Funktion im globalen und lokalen Sinne finden können:

Globale Extrema

Eine Stelle x_0 wird **globales Maximum von** f genannt, falls für alle x der Wert $f(x)$ kleiner oder gleich dem Wert von f in x_0 ist: $f(x) \leq f(x_0)$. Eine Stelle x_0 heißt **globales Minimum von** f, falls: $f(x) \geq f(x_0)$. Ist x_0 ein globales Maximum von f oder ein globales Minimum von f, so heißt x_0 **globales Extremum** von f.

Ein Vorteil (streng) monotoner Funktionen ist es, dass man deren globale Extrema leicht finden kann:

Globale Extrema monotoner Funktionen $f : [a, b] \mapsto \mathbb{R}$

Ist f eine monotone Funktion, die auf einem abgeschlossenen Intervall $[a, b]$ festgelegt ist: $f : [a, b] \mapsto \mathbb{R}$, so sind die Randstellen a und b globale Extrema von f.

Diese Aussage kann ich einfach einsehen. Ist f beispielsweise monoton steigend, so gilt $f(a) \leq f(x) \leq f(b)$ für alle $a < x < b$, dies bedeutet, dass die linke Randstelle a ein globales Minimum von f ist und die rechte Randstelle b ein globales Maximum von f ist. Ein solches Beispiel ist durch die Funktion $f(x) = 1/3\, x^3$, $x \in [-1, 1]$, aus Abbildung 6.8 gegeben. Diese ist sogar streng monoton steigend und es gilt: $f(-1) = -1/3 < f(x) < 1/3 = f(1)$ für alle $-1 < x < 1$: Damit ist $x_0 = -1$ das globale Minimum und $x_1 = 1$ das globale Maximum von diesem f. Ähnlich kann ich für $f(x) = -\exp((1/3)\, x^3) = -e^{(1/3)\, x^3}$, $x \in [0, 1]$ argumentieren. Oben hatte ich mit Ihnen gemeinsam geklärt, dass diese Funktion monoton fallend ist, und somit folgt, dass 0 deren globales Maximum ist und 1 deren globales Minimum ist: $f(0) = -1 > f(x) > -e^{1/3} = f(1)$. Ich habe somit gemeinsam mit Ihnen gesehen, dass das Auffinden globaler Extrema von (streng) monotonen Funktionen recht einfach ist, *solange* diese auf abgeschlossenen Intervallen $[a, b]$ betrachtet werden.

Schwieriger wird die Suche nach Stellen mit maximalen und minimalen Funktionswerten, falls die vorgegebene Funktion f insgesamt nicht monoton ist, sondern eher aus verschiedenen Monotoniebereichen besteht. Leider ist dies mit sehr vielen Funktionen so. Die Wanderung im „Pfälzer Wald“(siehe Abbildung 6.1) ist von dieser schwierigeren Art: Dort ist, wie ich eingangs erwähnte, die Stelle x_3 *im Innern* von $[a, b]$ ein globales Maximum, während die Randstelle a ein globales Minimum ist. Für solche Funktionen, bei denen – salopp ausgedrückt – die Art der Monotonie wechselt und Stellen im Innern von $[a, b]$ als Kandidaten für (globale) Extrema infrage kommen, muss ich im Allgemeinen feiner argumentieren und **lokale Extrema** mit in Betracht ziehen. Im „Pfälzer Wald“-Beispiel sind dies die Stellen x_3, x_4 und x_5.

Lokale Extrema

Eine Stelle x_0 wird **lokales Maximum von** f genannt, falls ich *ein* x_0 *umgebendes Intervall* $(x_0 - c, x_0 + c)$ finde ($c > 0$ ist eine feste Zahl), sodass dort für alle x der Wert $f(x)$ nicht größer als der Wert von f in x_0 ist: $f(x) \leq f(x_0)$, $x \in (x_0 - c, x_0 + c)$. Der Punkt $(x_0|f(x_0))$ heißt dann **Hochpunkt des Graphen von** f. Eine Stelle x_0 heißt **lokales Minimum von** f, falls: $f(x) \geq f(x_0)$ für alle x *aus einem* x_0 *umgebenden Intervall* $(x_0 - c, x_0 + c)$. Der Punkt $(x_0|f(x_0))$ heißt dann **Tiefpunkt des Graphen von** f. Ist x_0 ein lokales Maximum von f oder ein lokales Minimum von f, so heißt x_0 ein **lokales Extremum von** f. Der Punkt $(x_0|f(x_0))$ heißt dann **Extrempunkt des Graphen von** f.

Die Betragsfunktion $|x|$ besitzt ein lokales Minimum in $x = 0$. Das kann man einsehen, indem man deren Graphen in Abbildung 6.6 gründlich genug studiert. Wie ich Ihnen im vorherigen Abschnitt gezeigt habe, ist diese Funktion jedoch in $x = 0$ nicht ableitbar – der Wert der Ableitung existiert also dort nicht.

Einfacher wird die Bestimmung lokaler Extrema im Allgemeinen, falls ich ableitbare Funktionen betrachte, denn dann kann ich **analytisch** argumentieren. Unsere Wanderung im „Pfälzer Wald“ hat ja schon gezeigt, dass der Wert der Steigung an den Gipfel- und Talpunkten (die jetzt Hoch- und Tiefpunkte heißen) jeweils 0 ist. Das habe ich oben lediglich anschaulich klar gemacht. *Da der Wert der Steigung einer Funktion f an einer Stelle x gerade der Wert der ersten Ableitung $f'(x)$ an dieser Stelle x ist*, kann ich nun allgemeiner ein Kriterium zum Auffinden lokaler Extrema, beziehungsweise der zugehörigen Hoch- und Tiefpunkte, ableitbarer Funktionen f formulieren. Hierzu argumentiere ich folgendermaßen: Liegt beispielsweise ein lokales Maximum von f in x_0 vor, so ist f „links von x_0“, also in einem Intervall $(x_0 - c, x_0]$ monoton steigend, und „rechts von x_0“, also in einem Intervall $[x_0, x_0 + c)$ monoton fallend. Nach dem, was ich Ihnen oben über den Zusammenhang zwischen der Monotonie und dem Vorzeichen der ersten Ableitung erzählt habe, gelten somit: $f'(x) \geq 0$, $x \in (x_0 - c, x_0]$ und $f'(x) \leq 0$, $x \in [x_0, x_0 + c)$. Insbesondere $f'(x_0) \geq 0$ und $f'(x_0) \leq 0$, also $f'(x_0) = 0$. Analog kann ich argumentieren, falls x_0 ein lokales Minimum ist (siehe Abbildung 6.11). Ist x_0 ein lokales Extremum, so hat also die Tangente von f in x_0 die Steigung 0 und ist somit eine konstante Funktion – man spricht dann von einer **waagrechten Tangenten**.

Notwendiges Kriterium für lokale Extrema

Falls eine ableitbare Funktion f in x_0 ein lokales Extremum hat, so verschwindet dort die erste Ableitung von f, das heißt $f'(x_0) = 0$.

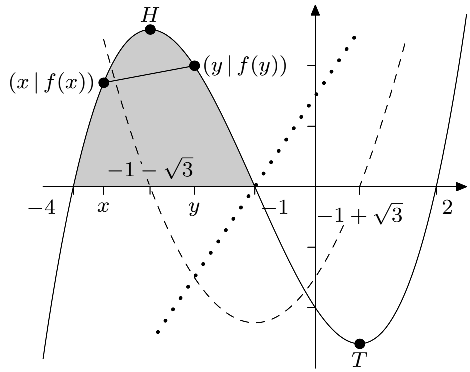

Abb. 6.10 Das Polynom $f(x) = 1/4\,x^3 + 3/4\,x^2 - 3/2\,x - 2$ besitzt ein lokales Maximum in $x_2 = -1 - \sqrt{3}$ und ein lokales Minimum in $x_1 = -1 + \sqrt{3}$. Die erste Ableitung $f'(x) = 3/4\,x^2 + 3/2\,x - 3/2$ (gestrichelt) verschwindet dort, $f'(x_1) = f'(x_2) = 0$, *und* wechselt dort jeweils das Vorzeichen. Im Bereich $x \leq -1$ ist der Graph dieser Funktion konkav, denn für die zweite Ableitung (gepunktete Gerade) gilt: $f''(x) = 3/2\,x + 3/2 \leq 0$. Im Bereich $x \geq -1$ ist dieser konvex, denn dort gilt: $f''(x) \geq 0$. In $x_0 = -1$ gelten $f''(-1) = 0$ und $f'''(-1) = 3/2$. Deshalb ist $x_0 = -1$ ein lokales Minimum von f' und somit eine Wendestelle von f. Am Wendepunkt $W = (-1|0)$ geht der Graph von f vom konkaven in den konvexen Bereich über.

Beispiel 6.23:
Ich schaue mir jetzt beispielsweise das **kubische Polynom**

$$f(x) = 1/4\,x^3 + 3/4\,x^2 - 3/2\,x - 2$$

an, dessen Graph ich in Abbildung 6.10 als durchgezogenen Bogen dargestellt habe und mache mich auf die Suche nach lokalen Extrema von f. Dazu rechne ich erst einmal – wie im letzten Abschnitt allgemeiner für **Polynomfunktionen** gezeigt wurde – die erste Ableitung aus:

$$f'(x) = 3/4\,x^2 + 3/2\,x - 3/2 = 3/4\,(x^2 + 2\,x - 2).$$

Um herauszufinden, an welchen Stellen lokale Extrema vorliegen könnten, setze ich die erste Ableitung gleich null: $3/4\,(x^2 + 2\,x - 2) = 0$. Ich erinnere mich jetzt an die Vorgehensweise zur Lösung **quadratischer Gleichungen**, insbesondere an die ***p*-*q*-Formel** aus Kapitel 3. Setze ich dort $p = 2$ und $q = -2$ ein, so kann ich die beiden Lösungen der quadratischen Gleichung $f'(x) = 0$ ausrechnen:

$$x_1 = -\frac{2}{2} + \sqrt{\frac{2^2}{4} - (-2)} = -1 + \sqrt{\frac{4}{4} + 2} = -1 + \sqrt{3} \text{ und } x_2 = -1 - \sqrt{3}.$$

Als *Kandidaten* für lokale Extrema von f habe ich also die beiden reellen Zahlen x_1 und x_2 erkannt. ■

Aber Sie sehen jetzt anhand dieses Beispiels, dass ich hier die Schwierigkeit habe, schnell festzustellen, um *welche Art* von lokalem Extremum (Maximum oder Minimum) es sich jeweils handelt. Irgendwie beschleicht mich an dieser Stelle auch ein ungutes Gefühl, denn obiges Kriterium besagt ja nur, dass an den lokalen Extrema die Ableitung verschwindet – ich habe hier aber sozusagen umgekehrt angefangen und die Stellen mit $f'(x) = 0$ bestimmt. Wer aber garantiert mir nun, dass an diesen Stellen nun auch tatsächlich lokale Extrema vorliegen? Ich werde Ihnen gleich zeigen, dass ich für diese Funktion f schnell checken kann, dass die Stellen x_1 und x_2 tatsächlich lokale Extrema sind. Zuvor aber muss ich zur Warnung das folgende Beispiel diskutieren, welches zeigt, dass das obige Kriterium *nicht hinreichend* für das Vorhandensein eines lokalen Extremums in x_0 ist: Leider garantiert mir die Bedingung $f'(x_0) = 0$ im Allgemeinen *nicht*, dass auch *tatsächlich* ein lokales Extremum in x_0 vorliegt.

Beispiel 6.24:
Um dies einzusehen, betrachte ich einen alten Bekannten: das Polynom $f(x) = 1/3\,x^3$ aus Abbildung 6.8. Weiter oben hatte ich Ihnen gezeigt, dass dessen erste Ableitung $f'(x) = x^2$ (nur) an der Stelle $x_0 = 0$ verschwindet, während diese Funktion zudem streng monoton steigend ist. Dies bedeutet, dass der Punkt $(x_0|f(x_0)) = (0|f(0)) = (0|0)$ weder ein Hoch- noch ein Tiefpunkt des Graphen von f ist, obgleich dort $f'(0) = 0$ gilt. Pech gehabt. ■

Das Verschwinden der ersten Ableitung ist eben lediglich ein *notwendiges* Kriterium für lokale Extrema: Es muss an den lokalen Extrema von ableitbaren Funktionen erfüllt sein. Umgekehrt garantiert mir das Verschwinden der ersten Ableitung an einer Stelle nicht, dass es sich bei dieser auch tatsächlich um ein lokales Extremum handelt.

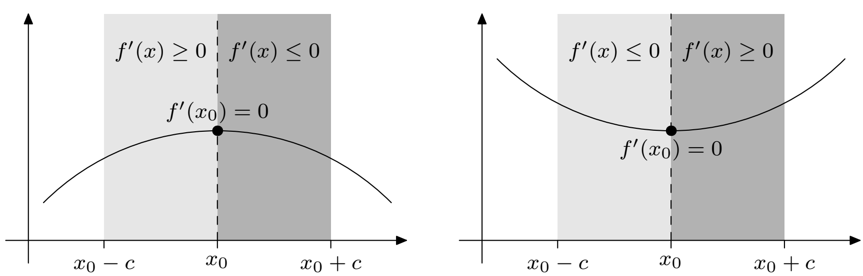

Abb. 6.11 Ist x_0 ein lokales Extremum, so verschwindet dort die Ableitung: $f'(x_0) = 0$. Hat die erste Ableitung f' dort einen Vorzeichenwechsel von $+$ nach $-$ (links), so handelt es sich um ein lokales Maximum von f. Liegt für diese dort ein Vorzeichenwechsel von $-$ nach $+$ vor (rechts), so handelt es sich um ein lokales Minimum von f.

Um ein **hinreichendes Kriterium** für lokale Extrema zu bekommen, muss ich dafür sorgen, dass mir so etwas wie in dem eben betrachteten Beispiel nicht passieren kann. Hierzu erinnere ich mich wiederum an meine Beschreibungen der Wanderung durch den „Pfälzer Wald“ eingangs dieses Abschnitts. Dort verschwindet die erste Ableitung an den Stellen x_3, x_4 und x_5, und ich erkenne, dass es sich dabei um lokale Extrema handelt. *Der Grund ist*, dass der Wert der Steigung hier in einer Umgebung dieser Stellen einen **Vorzeichenwechsel** vollzieht. Betrachte ich die Stelle

x_3, so wechselt die erste Ableitung f' ihr **Vorzeichen von + nach −**: in einem x_3 umgebenden Intervall $(x_3 - c, x_3 + c)$. Dort gelten nämlich: $f'(x) \geq 0$, $x \in (x_3 - c, x_3]$ und $f'(x) \leq 0$, $x \in [x_3, x_3 + c)$. An der Stelle x_3 liegt somit tatsächlich ein lokales Maximum vor. Betrachte ich die Stelle x_4, so wechselt die erste Ableitung f' ihr **Vorzeichen von − nach +** in einem x_4 umgebenden Intervall $(x_4 - c, x_4 + c)$. Dort gelten nämlich: $f'(x) \leq 0$, $x \in (x_4 - c, x_4]$ und $f'(x) \geq 0$, $x \in [x_4, x_4 + c)$. An der Stelle x_4 liegt somit tatsächlich ein lokales Minimum vor. Dieses **Vorzeichenwechsel-Kriterium** habe ich in Abbildung 6.11 veranschaulicht. In meinem letzten Beispiel $f(x) = (1/3)\,x^3$ ist dieses übrigens keinesfalls erfüllt: Die erste Ableitung $f'(x) = x^2$ ist hier stets ≥ 0 – insbesondere wechselt diese an der als lokales Extrema infrage kommenden Stelle $x_0 = 0$ das Vorzeichen nicht.

Ich betrachte jetzt – wie versprochen – noch einmal das obige kubische Polynom $f(x) = 1/4\,x^3 + 3/4\,x^2 - 3/2\,x - 2$ aus Beispiel 6.23. Als Kandidaten für die lokalen Extrema hatte ich – lediglich das notwendige Kriterium verwendend – $x_1 = -1+\sqrt{3} \approx 0{,}73$ und $x_2 = -1 - \sqrt{3} \approx -2{,}73$ ausgerechnet. Weil ich wissen möchte, ob es sich hierbei tatsächlich um lokale Extrema handelt, möchte ich nun checken, ob an diesen Stellen ein Vorzeichenwechsel von $f'(x) = 3/4\,(x^2 + 2\,x - 2)$ vorliegt. Hierzu betrachte ich die **Zerlegung in Linearfaktoren** (siehe Kapitel 3)

$$f'(x) = 3/4\,(x - x_1)\,(x - x_2) = 3/4\,(x + 1 - \sqrt{3})\,(x + 1 + \sqrt{3})$$

und ermittle das Vorzeichen der Ableitung in Abhängigkeit des gewählten x. Ist $x < x_2$, so gilt wegen $x_2 < x_1$ auch $x < x_1$. In diesem Fall ist also $x - x_2 < 0$ sowie $x - x_1 < 0$ und damit $f'(x) > 0$. Ist $x_2 < x < x_1$, so gelten $x - x_2 > 0$ sowie $x - x_1 < 0$ und damit $f'(x) < 0$. Im verbleibenden Fall $x_1 < x$ sehe ich genauso, dass $f'(x) > 0$ gilt. Ich erkenne somit: An der Stelle x_2 wechselt die erste Ableitung f' ihr Vorzeichen von + nach −, während in x_1 ein Vorzeichenwechsel von f' von − nach + vorliegt. Um dieses Vorzeichenverhalten von f' zu sehen, können Sie sich auch gerne noch einmal Abbildung 6.10 anschauen: Dort habe ich Ihnen als Extraservice den Graph des quadratischen Polynoms $f'(x)$ gestrichelt eingezeichnet. Jetzt ist nach dem obigen hinreichenden Kriterium klar, dass die Stelle x_2 tatsächlich ein lokales Maximum und x_1 tatsächlich ein lokales Minimum von f ist. Den zugehörigen Hoch- und Tiefpunkt habe ich in Abbildung 6.10 mit H beziehungsweise T gekennzeichnet.

Das war recht mühevoll, und es stellt sich die berechtigte Frage, ob es vielleicht nicht oftmals ein bequemeres, hinreichendes Kriterium als den eben am Beispiel gezeigten direkten Check auf Vorzeichenwechsel der ersten Ableitung gibt. Zum Glück ist dies der Fall – allerdings benötige ich hierzu **zweite Ableitungen**. Ich definiere das gleich ein bisschen allgemeiner:

Höhere Ableitungen

Eine ableitbare Funktion f heißt **zweimal ableitbar**, falls die Ableitung f' von f wiederum ableitbar ist. Die Ableitung $(f')'$ von f' nennt man dann **zweite Ableitung von** f und bezeichnet diese mit f'' (sprich: „f zwei Strich") oder $f^{(2)}$. Ist $i \in \mathbb{N}$, so heißt eine Funktion f ***i*-mal ableitbar**, falls die $(i-1)$-te Ableitung $f^{(i-1)}$ von f (existiert und) ableitbar ist. Die Ableitung $(f^{(i-1)})'$ von $f^{(i-1)}$ nennt man dann ***i*-te Ableitung von** f und bezeichnet diese mit $f^{(i)}$. Ist f ***i*-mal ableitbar** *für alle* $i \in \mathbb{N}$, so heißt f **unendlich oft ableitbar**.

Für das obige kubische Polynom $f(x) = 1/4\,x^3 + 3/4\,x^2 - 3/2\,x - 2$ aus Abbildung 6.10 habe ich bereits die erste Ableitung ausgerechnet: $f'(x) = 3/4\,x^2 + 3/2\,x - 3/2$. Ich erkenne, dass f' ein Polynom vom Grad zwei ist und kann dieses gemäß unseren Kenntnissen aus dem letzten Abschnitt wiederum ableiten. Die zweite Ableitung f'' von f ist somit das lineare Polynom $f''(x) = (f')'(x) = 3/2\,x + 3/2$ – dessen Graph habe ich Ihnen in Abbildung 6.10 als gepunktete Gerade veranschaulicht. Weil es mir Spaß macht, leite ich f noch einmal ab und erhalte die dritte Ableitung von f: $f'''(x) = (f'')'(x) = 3/2$. Jetzt sehe ich, dass die vierte Ableitung von f **verschwindet**: $f''''(x) = (f''')'(x) = 0$. Damit verschwindet natürlich auch die fünfte, sechste, ja alle nachfolgenden Ableitungen von f: $f^{(i)}(x) = 0$ für alle $i \geq 4$. Ich sehe also ein, dass dieses Polynom unendlich oft ableitbar ist. So etwas gilt allgemeiner:

Die i-te Ableitung $p^{(i)}$ von Polynomfunktionen p

Jede **Polynomfunktion p vom Grad n** ist unendlich oft ableitbar. Für $i \in \{1, \ldots, n\}$ ist die i-te Ableitung $p^{(i)}$ von p ein Polynom vom Grad $n-i$ und für $i \geq n+1$ verschwindet die i-te Ableitung $p^{(i)}$, das heißt $p^{(i)}(x) = 0$ für alle x.

Ohne dies weiter zu kommentieren, zeige ich Ihnen noch folgendes Beispiel, bei dem ich ein Polynom vom Grad fünf siebenmal abgeleitet habe:

$$p(x) = x^5 + 3\,x^3 - 2\,x,\; p'(x) = 5\,x^4 + 9\,x^2 - 2,\; p''(x) = 20\,x^3 + 18\,x,$$
$$p^{(3)}(x) = 60\,x^2 + 18,\; p^{(4)}(x) = 120\,x,\; p^{(5)}(x) = 120,\; p^{(6)}(x) = 0,\; p^{(7)}(x) = 0.$$

Übungsaufgabe 6.10:
Bestimmen Sie alle Ableitungen der folgenden Polynomfunktionen: a) $p(x) = 4\,x^4 - 6\,x^2 + e^8$, b) $f_t(x) = t^3\,x^3 - t\,x^2 + t\,x + t^3$. ■

Höhere Ableitungen kann ich auch für viele andere Funktionen, die keine Polynome sind, bilden.

Beispiel 6.25:
Ein Beispiel ist $f(x) = e^{x^2-x}$. Weil ich Übungsaufgabe 6.8b) gemacht habe, weiß ich, dass hier die erste Ableitung für alle $x \in \mathbb{R}$ wie folgt ist: $f'(x) = (2x-1)\,e^{x^2-x}$. Um

jetzt die zweite Ableitung von f auszurechnen wende ich zunächst die **Produktregel** an und verwende dann nochmals meine Kenntnisse über die erste Ableitung von f:

$$\begin{aligned} f''(x) &= (2x-1)'\,e^{x^2-x} + (e^{x^2-x})'\,(2x-1) = 2\,e^{x^2-x} + (2x-1)\,e^{x^2-x}\,(2x-1) \\ &= (2+(2x-1)^2)\,e^{x^2-x} = (2+4\,x^2-4\,x+1)\,e^{x^2-x} = (4\,x^2-4\,x+3)\,e^{x^2-x}. \end{aligned}$$

Wenn ich wollte, so könnte ich f'' jetzt noch einmal oder zweimal ableiten. ■

Übungsaufgabe 6.11:
Bestimmen Sie f''' und f'''' von $f(x) = e^{x^2-x}$. ■

Wenn man sich das so anschaut, so könnte man bei erster, oberflächlicher Betrachtung möglicherweise meinen, dass man höhere Ableitungen (einmal) ableitbarer Funktionen vielleicht immer bilden kann. Dies ist – Sie ahnen es – aber nicht so. Schauen Sie sich beispielsweise die Funktion $f(x) = x^{5/2}$ für $x \geq 0$ an. Diese Funktion ist ableitbar und ich erfülle ein im letzten Abschnitt gegebenes Versprechen, indem ich Ihnen deren Ableitung verrate: $f'(x) = 5/2\,x^{3/2}$, $x \geq 0$. Auch die zweite Ableitung f'' finde ich mühelos durch Anwendung der Grobregel zur Bildung der Ableitung von Potenzfunktionen aus dem letzten Abschnitt: $f''(x) = (5/2)(3/2)\,x^{1/2} = 15/4\,\sqrt{x}$, $x \geq 0$. Aber Achtung: f'' kann ich jetzt nicht noch einmal für alle $x \geq 0$ ableiten, denn hier taucht die Wurzelfunktion als Term auf, und diese ist bekanntlich in $x = 0$ nicht ableitbar. Dieses f ist somit eine Funktion, die zweimal, aber *nicht* dreimal ableitbar ist. In Übungsaufgabe 6.4b) finden Sie eine weitere solche Funktion.

Übungsaufgabe 6.12:
Zeigen Sie, dass $f(x) = \begin{cases} (-1/2)\,x^2, & \text{falls } x \geq 0 \\ 1/2\,x^2, & \text{falls } x < 0 \end{cases}$ einmal, aber *nicht* zweimal ableitbar ist. ■

Auch wenn es mir viel Spaß macht, über höhere Ableitung zu plaudern, so sollte ich doch nicht vergessen, warum ich diese betrachten wollte. Oben hatte ich Ihnen nämlich gesagt, dass durch die Verwendung der zweiten Ableitung der etwas mühselige direkte Check auf Vorzeichenwechsel der ersten Ableitung in der Umgebung der Stellen x_0, die als Kandidaten für lokale Extrema infrage kommen (also: $f'(x_0) = 0$), vermieden werden kann. Das kann ich mir wie folgt klar machen: Die zweite Ableitung f'' von f ist die (erste) Ableitung von f' – falls f somit in x_0 *zweimal ableitbar* ist, so existiert $f''(x_0)$ – der Grenzwert der Differenzenquotienten von f' in x_0 für h:

$$f''(x_0) = \lim_{h\to 0} \frac{f'(x_0+h) - f'(x_0)}{h} = \lim_{h\to 0} \frac{f'(x_0+h)}{h}.$$

Das letzte „="-Zeichen ist hier gültig, weil es sich bei x_0 um einen Kandidaten für ein lokales Extrema handelt, also $f'(x_0) = 0$ gilt. Ist nun beispielsweise $f''(x_0) < 0$, so haben die Differenzenquotienten $f'(x_0+h)/h$ für h nahe 0 keine andere Chance, als auch kleiner 0 zu sein. Betrachte ich jetzt Stellen „links von x_0", also $x_0 + h$, wobei $h < 0$, so ist dort $f'(x_0+h) > 0$. Schaue ich mir andererseits Stellen „rechts von x_0" an, also $x_0 + h$, wobei $h > 0$, so ist dort $f'(x_0+h) < 0$. Zusammenfassend wechselt die erste Ableitung f' in x_0 ihr Vorzeichen von $+$ nach $-$, denn in einem x_0 umgebenden Intervall $(x_0 - c, x_0 + c)$ gelten: $f'(x) > 0$, $x \in (x_0 - c, x_0)$ und $f'(x) < 0$, $x \in (x_0, x_0 + c)$. Somit liegt in x_0 nach dem obigen Vorzeichenwechsel-Kriterium ein lokales Maximum vor. Wenn ich mich nun noch einmal frage, was ich hier eigentlich verwendet habe, so erkenne ich, dass hier neben $f'(x_0) = 0$ die zentrale Annahme ist, dass das *Vorzeichen der zweiten Ableitung in* x_0 „negativ" ist: $f''(x_0) < 0$.

Hinreichendes Kriterium für lokale Extrema

Ist f eine zweimal ableitbare Funktion und x_0 eine Stelle, sodass $f'(x_0) = 0$, so besitzt f in x_0 ein lokales Extremum, *falls* zudem $f''(x_0) \neq 0$ gilt. Gilt dann $f''(x_0) < 0$, so ist x_0 ein lokales Maximum. Gilt dann $f''(x_0) > 0$, so ist x_0 ein lokales Minimum.

Ich schaue mir nochmals Beispiel 6.23 – also das kubische Polynom $f(x) = 1/4\,x^3 + 3/4\,x^2 - 3/2\,x - 2$ aus Abbildung 6.10 an. Oben habe ich dafür bereits die zweite Ableitung ausgerechnet: $f''(x) = 3/2\,(x + 1)$. Ich setze nun die Stellen $x_1 = -1 + \sqrt{3}$ und $x_2 = -1 - \sqrt{3}$ in die zweite Ableitung ein und ermittle deren Vorzeichen:

$$f''(x_1) = 3/2\,(-1 + \sqrt{3} + 1) = (3\sqrt{3})/2 > 0 \text{ und } f''(x_2) = (-3\sqrt{3})/2 < 0.$$

Aus dem eben genannten, hinreichenden Kriterium folgt, dass x_1 ein lokales Minimum von f und x_2 ein lokales Maximum von f ist. Dies bestätigt auf bequemere Art und Weise das Resultat, das ich Ihnen oben auf relativ mühseligem Weg durch direkten Check des Vorzeichens der ersten Ableitung gezeigt habe.

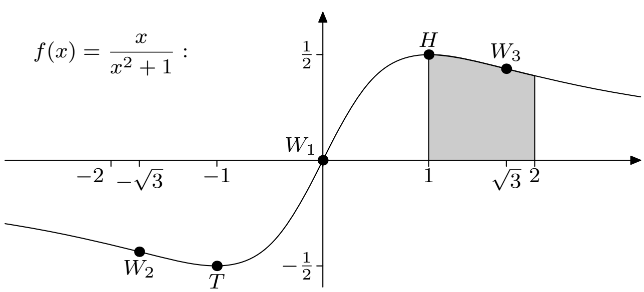

Abb. 6.12 Der Graph der Funktion $f(x) = x/(x^2 + 1)$. Dieser besitzt einen Hochpunkt $H = (1|1/2)$ und einen Tiefpunkt $T = (-1|-1/2)$ sowie drei Wendepunkte $W_1 = (0|0)$, $W_2 = (-\sqrt{3}|-\sqrt{3}/4), W_3 = (\sqrt{3}|\sqrt{3}/4)$.

Beispiel 6.26:

Etwas schwieriger ist das Finden der lokalen Extrema der gebrochen-rationalen Funktion $f(x) = x/(x^2 + 1)$, deren Graph ich Ihnen in Abbildung 6.12 skizziert habe. Um herauszufinden, welche Stellen als Kandidaten für lokale Extrema infrage kommen, sollte ich die erste Ableitung bilden. Dies tue ich mithilfe der **Quotientenregel** aus dem letzten Abschnitt:

$$f'(x) = \frac{(x)'\,(x^2+1) - (x^2+1)'\,x}{(x^2+1)^2} = \frac{(x^2+1) - 2x\cdot x}{(x^2+1)^2} = \frac{1-x^2}{(x^2+1)^2}.$$

Jetzt bestimme ich die Stellen, an denen f' verschwindet: $(1-x^2)/(x^2+1)^2 = 0$. Weil ich hier beide Seiten mit dem positiven Term $(x^2+1)^2$ multiplizieren kann, muss ich nur noch schauen, an welchen Stellen der Zähler verschwindet: $1-x^2 = 0$. Ich erkenne mit meinen Kenntnissen aus Kapitel 3, dass diese Gleichung die beiden Lösungen $x_0 = 1$ und $\tilde{x}_0 = -1$ hat – also: $f'(1) = 0$ und $f'(-1) = 0$. Jetzt möchte ich die zweite Ableitung f'' von f verwenden, um zu checken, ob diese Kandidaten auch tatsächlich lokale Extrema von f sind. Glücklicherweise kenne ich f'' bereits, denn Sie haben sicher Übungsaufgabe 6.8d) oder 6.9d) für mich gemacht:

$$f''(x) = (f')'(x) = \left(\frac{1-x^2}{(x^2+1)^2}\right)' = \frac{2x\,(x^2-3)}{(x^2+1)^3}.$$

Ich setze jetzt $x_0 = 1$ und $\tilde{x}_0 = -1$ hier ein und erhalte $f''(1) = (2(-2))/2^3 = -4/8 = -1/2 < 0$ sowie $f''(-1) = 1/2 > 0$. Das hinreichende Kriterium besagt nun, dass in $x_0 = 1$ ein lokales Maximum und in $\tilde{x}_0 = -1$ ein lokales Minimum von f vorliegt. Wenn ich möchte, so kann ich den zugehörigen Hochpunkt H beziehungsweise Tiefpunkt T des Graphen von f bestimmen, indem ich $x_0 = 1$ und $\tilde{x}_0 = -1$ in *die Funktion* f einsetze: $H = (1|f(1)) = (1|1/2)$, $T = (-1|f(-1)) = (-1|-1/2)$. Diese habe ich in Abbildung 6.12 als schwarze Punkte dargestellt. ■

Übungsaufgabe 6.13:
Zeigen Sie, dass $f(x) = e^{x^2-x}$ genau ein lokales Extremum besitzt und bestimmen Sie den zugehörigen Extrempunkt des Graphen von f. Beachten Sie, dass ich Ihnen weiter oben schon f' und f'' verraten habe. ■

Was macht man eigentlich, falls für einen Kandidaten x_0 mit $f'(x_0) = 0$ das hinreichende Kriterium nicht erfüllt ist – es kann ja sein, dass ich mit viel Pech $f''(x_0) = 0$ berechne? Natürlich kann ich immer – wie oben gezeigt – versuchen, über das Wechseln des Vorzeichens von f' zu argumentieren. Ich verrate Ihnen jetzt aber etwas, was man für höher ableitbare, sagen wir i-mal ableitbare, Funktionen auch manchmal mit Erfolg verwenden kann: Falls i eine *gerade Zahl* ist und

$$f'(x_0) = f''(x_0) = \ldots = f^{(i-1)}(x_0) = 0 \,\text{sowie}\, f^{(i)}(x_0) \neq 0$$

gelten, so besitzt f in x_0 ein lokales Extremum. Falls $f^{(i)}(x_0) < 0$ gilt, so ist x_0 dann ein lokales Maximum von f. Anderenfalls ist x_0 dann ein lokales Minimum von f.

Beispiel 6.27:
Ein Beispiel, für das ich dieses Kriterium anwenden kann, ist $f(x) = 1/12\,x^4$. Ich rechne die ersten vier Ableitungen aus:

$$f'(x) = 1/3\,x^3, f''(x) = x^2, f'''(x) = 2x, f''''(x) = 2\,.$$

Jetzt stelle ich fest, dass $f'(0) = f''(0) = f'''(0) = 0$ sowie $f''''(0) = 2 > 0$ gelten. Schön, $T = (0, 0)$ ist Tiefpunkt des Graphen dieses fs.

Aber Achtung: Ein vergleichbares Kriterium kann für *ungerade Zahlen i nicht* funktionieren. ■

Beispiel 6.28:
Das sehe ich ein, indem ich den uralten Bekannten $f(x) = 1/3\,x^3$ aus Abbildung 6.8 betrachte. Ich stelle für diesen unschwer fest, dass $f'(0) = f''(0) = 0$ sowie $f'''(0) = 2 \neq 0$ gelten. Das eben genannte hinreichende Kriterium scheitert hier, denn $i = 3$ ist ungerade. Kein Wunder, denn ich habe Ihnen oben ja schon gezeigt, dass $x = 0$ kein lokales Extremum von diesem f ist, wovon man sich auch durch intensives Betrachten von Abbildung 6.8 überzeugen kann. ■

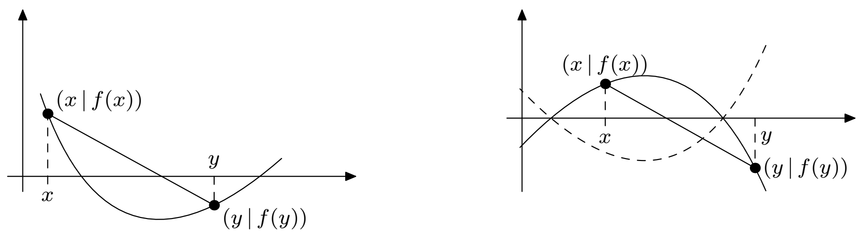

Abb. 6.13 Die durchgezogenen Bogen stellen konvexe (links) und konkave (rechts) Bereiche des Graphen von f dar. Spiegelt man an der x-Achse (Ersetzung von f durch $-f$), so wird ein konkaver Bereich konvex (gestrichelter Bogen) und umgekehrt.

Nun denke ich gründlich genug über die Suche nach lokalen und globalen Extrema von ableitbaren Funktionen berichtet zu haben und schlage jetzt vor, mit Ihnen noch ein bisschen über die **Konvexität** und **Konkavität** von Graphen sowie **Wendestellen** zu plaudern. **Konvexe** Bereiche des Graphen einer Funktion f sind dadurch charakterisiert, dass für jede Wahl von zwei Stellen $x < y$ das Kurvensegment des Graphen von f mit Endpunkten $(x|f(x))$ und $(y|f(y))$ dort stets unterhalb der (aus dem vorherigen Abschnitt bekannten) **Sekante** durch diese Punkte liegt – dies habe ich in Abbildung 6.13 (links) veranschaulicht. Analog liegen in den **konkaven** Bereichen alle solchen Kurvensegmente oberhalb der entsprechenden Sekanten (siehe durchgezogener Bogen in Abbildung 6.13 (rechts)). **Wendestellen** von f sind Stellen x_0, an denen der Graph von f von einem konvexen in einen konkaven Bereich wechselt oder umgekehrt – die zugehörigen Punkte $(x_0|f(x_0))$ nennt man **Wendepunkte**. In Abbildung 6.10 habe ich einen solchen Punkt mit W gekennzeichnet, während in Abbildung 6.12 drei Wendepunkte – W_1, W_2 und W_3 – zu sehen sind. Für *zweimal* ableitbare Funktionen besteht ein Zusammenhang zwischen diesen geometrischen Begriffen (Konvexität & Konkavität) und dem Vorzeichen der zweiten Ableitung.

> **Konvexität und Vorzeichen der zweiten Ableitung**
>
> Der Graph zweimal ableitbarer Funktionen f ist *genau dann* konvex, falls die zweite Ableitung f'' für alle x nicht negativ wird: $f''(x) \geq 0$.

Mit dem gestrichelten Bogen in Abbildung 6.13 versuche ich Ihnen klar zu machen, dass der Graph einer Funktion f genau dann konkav ist, wenn der Graph von $-f$ konvex ist. Dies impliziert somit, dass der Graph zweimal ableitbarer Funktionen f genau dann konkav ist, falls die zweite Ableitung f'' für alle x nicht positiv wird: $f''(x) \leq 0$. Um dies richtig zu verstehen, kann man sich gegebenenfalls klar machen, dass gemäß der **Faktorregel** aus dem letzten Abschnitt immer $(-f)'' = -f''$ gilt.

Konvexe Bereiche des Graphen einer Funktion f kann ich nun finden, indem ich beachte, dass in diesen die *erste Ableitung* f' von f monoton ist. Unsere allererste Aussage in diesem Abschnitt besagt ja, dass eine Funktion – ich nenne sie jetzt hier kurz g – genau dann monoton steigend ist, wenn deren erste Ableitung g' nicht negativ ist: $g'(x) \geq 0$. Wende ich diese Aussage nun *auf die erste Ableitung* $g = f'$ von f an, so folgt, dass diese genau dann monoton steigend ist, wenn deren erste Ableitung $(f')' = f''$ – also die *zweite Ableitung von* f – nicht negativ ist: $f''(x) \geq 0$. Die obige Aussage über die Konvexität und Vorzeichen der zweiten Ableitung besagt aber, dass *diese* Vorzeichenbedingung der zweiten Ableitung f'' die konvexen Bereiche von f charakterisiert. Zusammenfassend bilden die *Monotoniebereiche* der ersten Ableitung f' von f die konvexen und konkaven Bereiche des Graphen von f.

Ich denke, Beispiele zeigen jetzt mehr als tausend Worte. Ich betrachte das kubische Polynom $f(x) = 1/3\,x^3 - x^2 + 2/3$. Die erste Ableitung von f ist $f'(x) = x^2 - 2x$. Vielleicht kommt Ihnen f' irgendwie bekannt vor. Richtig – in Beispiel 6.17 habe ich diese Funktion g genannt: $g = f'$. Die Graphen von g und dessen erste Ableitung g' finden Sie immer noch in Abbildung 6.7 (links) – dies sind gerade die Graphen von f' beziehungsweise f''. Gemeinsam mit Ihnen habe ich oben geklärt, dass g für $x \leq 1$ monoton fallend und für $x \geq 1$ monoton steigend ist – dies bedeutet *für* f, dass $g'(x) = f''(x) \leq 0$ für $x \leq 1$ sowie $g'(x) = f''(x) \geq 0$ für $x \geq 1$ gelten. Für *die Funktion* f stelle ich somit fest, dass der Graph konkav für $x \leq 1$ sowie konvex für $x \geq 1$ ist.

Ein weiteres Beispiel ist die Funktion $F(x) = -\cos(x) + 4$ für $x \in [-\pi/4, \pi/4]$. Deren erste Ableitung F' rechne ich mit den Kenntnissen aus dem letzten Abschnitt wie folgt aus: $F'(x) = (-1)(\cos(x))' + (4)' = (-1)(-\sin(x)) + 0 = \sin(x)$. Dass ich hier ein großes „F“ als Bezeichnung wähle, hat – wie ich Ihnen noch erklären werde – zum einen mit den Inhalten des nächsten Abschnitts zu tun, und zum anderen möchte ich diese aber auch nicht mit der Funktion $f = F'$, deren Graph in Abbildung 6.7 (rechts) zu sehen ist, verwechseln, denn F und $F' = f$ unterscheiden sich offenbar. Weiter oben hatte ich Ihnen bereits gezeigt, dass $F'' = f'$ im betrachteten Intervall $[-\pi/4, \pi/4]$ stets größer als 0 und deshalb f dort streng monoton wachsend ist. Für F bedeutet dies, dass dessen Graph dort konvex ist.

Unser Interesse gilt jetzt den Stellen x_0, an denen der Graph von f von einem konvexen in einen konkaven Bereich wechselt und umgekehrt. Diese bilden ebenfalls charakteristische Punkte des Graphen von f, die sich von den Hoch- und Tiefpunkten im Allgemeinen unterscheiden.

Wendestellen

Eine Stelle x_0 wird **Wendestelle von** f genannt, falls f' in x_0 ein lokales Extremum hat. Der Punkt $(x_0|f(x_0))$ heißt dann **Wendepunkt des Graphen von** $\boldsymbol{f}$.

Das kubische Polynom $f(x) = 1/3\,x^3 - x^2 + 2/3$ von eben hat die Wendestelle $x_0 = 1$, denn $f'(x) = x^2 - 2x$ hat in x_0 ein lokales Extremum. Hierzu schaue ich mir $f''(x) = 2x - 2$ an und prüfe mühelos, dass das notwendige Kriterium $f''(1) = 0$ erfüllt ist. Um zu sehen, dass es sich hierbei auch tatsächlich um ein lokales Extremum von f' handelt, verwende ich unser bequemes, hinreichendes Kriterium. Ich bilde also die dritte Ableitung $f'''(x) = 2$ und stelle fest, dass $f'''(1) = 2 > 0$ gilt. Die Stelle $x_0 = 1$ ist somit ein lokales Minimum von f'. Es folgt, dass der Punkt $W = (1|f(1)) = (1|0)$ Wendepunkt von f ist, und ich erfahre sogar auch noch, dass der Graph von f in W von einem konkaven in einen konvexen Bereich übergeht. Ein ähnliches Beispiel ist in der Unterschrift zur Abbildung 6.10 diskutiert.

Beispiel 6.26 (Fortsetzung):
Die in Abbildung 6.12 veranschaulichte Funktion $f(x) = x/(1 + x^2)$ hat drei Wendestellen. Um dies zu sehen, erinnere ich mich zunächst an deren zweite Ableitung: $f''(x) = 2x\,(x^2 - 3)/(x^2 + 1)^3$. Die notwendige Bedingung $f''(x) = 0$ impliziert, dass $2x\,(x^2 - 3) = 0$ gelten muss. Als Kandidaten für die Wendestellen mache ich also $x_1 = 0, x_2 = -\sqrt{3}$ und $x_3 = \sqrt{3}$ aus. Ich muss jetzt checken, ob es sich tatsächlich um lokale Extrema von f' handelt. Ich mache das erst einmal, indem ich direkt f'' hinsichtlich Vorzeichenwechsel überprüfe. Mithilfe der **Binomischen Formel** schreibe ich den Zähler $z(x) = 2x\,(x^2 - 3)$ von f'' wie folgt hin: $z(x) = 2x\,(x^2 - 3) = 2x\,(x + \sqrt{3})(x - \sqrt{3})$. Dann erkenne ich, dass $z(x) < 0$ gilt, falls $x < -\sqrt{3}$ oder $0 < x < \sqrt{3}$ und sonst $z(x) \geq 0$ ist. Der Nenner von f'' ist zudem immer größer 0, und es folgt, das f'' in x_2 und x_3 einen Vorzeichenwechsel von $-$ nach $+$ hat sowie in x_1 einen Vorzeichenwechsel von $+$ nach $-$ besitzt. Somit sind x_2 und x_3 lokale Minima von f' und x_1 ist ein lokales Maximum von f'. Dies bedeutet, dass der Graph von f in $x_2 = -\sqrt{3}$ und $x_3 = \sqrt{3}$ jeweils von einem konkaven in einen konvexen Bereich übergeht, während dieser in $x_1 = 0$ von konvex nach konkav wechselt. Insbesondere sind $W_1 = (0|f(0)) = (0|0)$, $W_2 = (-\sqrt{3}|f(-\sqrt{3})) = (-\sqrt{3}| - \sqrt{3}/4)$ und $W_3 = (\sqrt{3}|f(\sqrt{3})) = (\sqrt{3}|\sqrt{3}/4)$ Wendepunkte von f. Anstatt des Checks des Vorzeichens von f'' kann ich hier natürlich auch unser „bequemes" hinreichendes Kriterium verwenden. Hierzu brauche ich die dritte Ableitung f''' von f. Okay – ich rechne diese aus. Erst einmal verwende ich die **Quotientenregel**:

$$f'''(x) = \left(\frac{2x\,(x^2 - 3)}{(x^2 + 1)^3}\right)' = \frac{(2x\,(x^2 - 3))'\,(x^2 + 1)^3 - ((x^2 + 1)^3)'\,(2x\,(x^2 - 3))}{(x^2 + 1)^6}.$$

Hier muss ich also die erste Ableitung von $2x\,(x^2-3)$ und $(x^2+1)^3$ finden. Für den ersten Term ist dies leicht: $(2x\,(x^2-3))' = (2x^3-6x)' = 6x^2-6 = 6\,(x^2-1)$. Für den zweiten Term benutze ich die **Kettenregel**: $((x^2+1)^3)' = 6x\,(x^2+1)^2$. Ich klammere jetzt den Term $6\,(x^2+1)^2$ im Zähler von f''' aus und vereinfache diesen durch Ausmultiplizieren der Restterme, wobei ich unter anderem die **Binomische Formel** $(x^2-1)\,(x^2+1) = x^4-1$ verwende:

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

Jetzt ergibt sich f''' durch Kürzen des Terms $(x^2+1)^2$:

$$f'''(x) \;=\; \frac{6\,(x^2+1)^2\,(-x^4+6x^2-1)}{(x^2+1)^6} = \frac{6\,(-x^4+6x^2-1)}{(x^2+1)^4}.$$

Ich setze nun die Wendestellen in f''' ein: $f'''(0) = -6 < 0, f'''(-\sqrt{3}) = f'''(\sqrt{3}) = 15/64 > 0$ und bestätige somit auf alternativem Weg, dass diese tatsächlich lokale Extrema von f' – also Wendestellen von f – sind. Außerdem erkenne ich so auch, dass der Graph von f in $x_2 = -\sqrt{3}$ und $x_3 = \sqrt{3}$ jeweils von einem konkaven in einen konvexen Bereich übergeht, während dieser in $x_1 = 0$ von konvex nach konkav wechselt. Ich habe Ihnen diese alternative Vorgehensweise auch deshalb gezeigt, weil ich darauf hinweisen möchte, dass die Verwendung des „bequemeren" hinreichenden Kriteriums sich manchmal als aufwendiger als der direkte Check auf Vorzeichenwechsel herausstellen kann. Dies hängt damit zusammen, dass die Bestimmung der dritten Ableitung zuweilen recht komplizierte Rechnungen erfordert. ■

Die Beispiele und meine Ausführungen auf den letzten drei Seiten zeigen Ihnen, dass ich die notwendigen und hinreichenden Kriterien, die ich oben zunächst zur Bestimmung von lokalen Extrema von ableitbaren Funktionen f formuliert habe, nun im neuen Kontext benutzen darf. *Da die Wendestellen von f lokale Extrema der ersten Ableitung f' von f sind*, kann ich diese bestimmen, indem ich die obigen Kriterien auf die Funktion f' anwende. Ich fasse zusammen, was ich bereits entwickelt habe:

Notwendiges und hinreichendes Kriterium für Wendestellen

Notwendiges Kriterium: Falls eine zweimal ableitbare Funktion f in x_0 eine Wendestelle hat, so verschwindet dort die zweite Ableitung von f, das heißt $f''(x_0) = 0$.
Hinreichendes Kriterium: Ist f eine dreimal ableitbare Funktion und x_0 eine Stelle, so dass $f''(x_0) = 0$, so besitzt f in x_0 eine Wendestelle, *falls* zudem $f'''(x_0) \neq 0$ gilt.

An dieser Stelle komme ich nochmals auf unseren uralten Bekannten $f(x) = 1/3\,x^3$ zurück, der in Abbildung 6.8 veranschaulicht ist. Weiter oben hatte ich Ihnen gezeigt, dass die Stelle $x_0 = 0$ *kein* lokales Extremum von diesem f ist. Insbesondere habe ich Ihnen erklärt, dass $f'(0) = f''(0) = 0$ sowie $f'''(0) = 2 \neq 0$ gelten. Hiermit erkenne ich *nun* sofort, dass der Punkt $W = (0|f(0)) = (0|0)$ ein Wendepunkt von f ist. Solche Wendepunkte, an denen auch die erste Ableitung verschwindet, nennt man, geleitet durch das dortige Aussehen des Graphens, **Sattelpunkte des Graphen** von f.

Ich fasse kurz zusammen: Falls die zweite Ableitung an einer Stelle verschwindet, so kann ich im Allgemeinen noch nicht davon ausgehen, dass dort auch tatsächlich eine Wendestelle vorliegt. Überprüfen kann ich dies schnell, indem ich schaue, ob die dritte Ableitung an dieser Stelle ungleich 0 ist. Achtung: Manchmal hat man hierbei Pech.

Übungsaufgabe 6.14:
a) Stellen Sie fest, ob $f(x) = (1/12)\,x^4$ eine Wendestelle besitzt. b) Zeigen Sie, dass $f(x) = e^{x^2-x}$ keine Wendestellen hat. ■

Ich denke, dass Ihnen inzwischen klar geworden ist, wie man Ableitungen zur Kurvendiskussion verwenden kann. Insbesondere haben Sie gesehen, wie man mit Ableitungen charakteristische Punkte des Graphen, also Hoch-, Tief- und Wendepunkte, findet. Außerdem zeigte ich Ihnen, wie man durch Ableitungen das Verhalten des Graphen zwischen und außerhalb von solchen Punkten beschreibt. Verwendet man nur diese Kenntnisse, so kann man schon den Graphen recht vieler Funktionen erfolgreich ins **Koordinatensystem** einzeichnen. An dieser Stelle möchte ich Sie aber noch darauf hinweisen, dass es neben der Anwendung von Ableitungen noch eine Reihe von anderen Möglichkeiten gibt, sich zusätzliche Informationen über den Verlauf von Graphen zu verschaffen. Solche Quellen für Zusatzinformationen beschreibe ich im Folgenden stichwortartig, wobei ich mich an Beispielen oben bereits behandelter Funktionen orientiere:

• **Nullstellen von** f: Dies sind Stellen x_0, an der die Funktion den Wert 0 besitzt: $f(x_0) = 0$. In Kapitel 3 wurde Ihnen die Lösung von Gleichungen beschrieben – diese Methoden helfen hier oftmals weiter. Ich illustriere dies an ein paar in diesem Abschnitt aufgetretenen Beispielen. Für die in Abbildung 6.7 (links) dargestellte Funktion $g(x) = x^2 - 2x$ finde ich diese wie folgt. Ich klammere zunächst x aus: $x^2 - 2x = x(x-2)$. Dann setze ich den Funktionsterm „gleich 0“: $x(x-2) = 0$. Jetzt kann ich die Nullstellen $x_0 = 0$ und $x_1 = 2$ ablesen. Für die Sinusfunktion in Abbildung 6.7 (rechts) kenne ich die (unendlich vielen) Nullstellen noch aus alten Schülerzeiten: $\sin(k\pi) = 0$, $k \in \mathbb{Z}$. Für das Polynom $f(x) = 1/4\,x^3 + 3/4\,x^2 - 3/2\,x - 2$ aus Abbildung 6.10 habe ich die erste Nullstelle $x_0 = -1$ geraten: $f(-1) = 0$. Danach habe ich – weil ich weiß, dass ich jetzt den **Linearfaktor** $(x+1)$ abspalten kann – den folgenden Ansatz gemacht:

$$1/4\,x^3 + 3/4\,x^2 - 3/2\,x - 2 = (a\,x^2 + b\,x + c)\,(x+1).$$

Rechne ich hier die rechte Seite aus, so erhalte ich für diese

$$a\,x^3 + (a+b)\,x^2 + (b+c)\,x + c,$$

und ich kann durch **Koeffizientenvergleich** sofort ablesen, dass $a = 1/4$ und $c = -2$ gelten. Zudem müssen $a + b = 3/4$ und $b + c = -3/2$ erfüllt sein und beides klappt glücklicherweise für die Wahl $b = 1/2$. Die restlichen beiden Nullstellen von

$$f(x) = 1/4\,x^3 + 3/4\,x^2 - 3/2\,x - 2 = (1/4\,x^2 + 1/2\,x - 2)(x+1) = 1/4\,(x^2 + 2\,x - 8)(x+1)$$

kann ich jetzt mit der **p-q-Formel** aus Kapitel 3 ausrechnen, indem ich $x^2 + 2\,x - 8 = 0$ (mit $p = 2$ und $q = -8$) löse: $f(-4) = f(2) = 0$. Für die Funktion $f(x) = e^{x^2 - x}$ muss ich mir keine großen Gedanken machen: Sie besitzt keine Nullstelle, denn $exp(\bar{x}) = e^{\bar{x}}$ ist – unabhängig davon was für $\bar{x}$ eingesetzt wird – stets größer als 0. Schließlich schaue ich mir noch die gebrochen-rationale Funktion $f(x) = x/(x^2 + 1)$ aus Abbildung 6.12 an. Um zu sehen, an welchen Stellen diese Funktion verschwindet, setze ich den Funktionsterm gleich null: $x/(x^2 + 1) = 0$. Weil ich hier beide Seiten mit dem positiven Term $x^2 + 1$ multiplizieren kann, muss ich nur noch schauen, an welchen Stellen der Zähler – x – verschwindet. Ich erkenne jetzt unschwer, dass $x_1 = 0$ die einzige Nullstelle von diesem f ist.

• **Symmetrie des Graphen**: Die Graphen mancher Funktionen f sind **achsensymmetrisch** zu einer Geraden $x = a$ parallel zur **y-Achse** mit Abstand a. Falls dies so ist, so gilt für alle x:

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

Anschaulich bedeutet dies, dass die Achse $x = a$ Spiegelachse des Graphen von f ist. Ein Beispiel ist die Funktion $f(x) = x^2$ (siehe Abbildung 6.4). Der Graph dieser Funktion ist achsensymmetrisch zur y-Achse $x = 0$. Nachrechnen kann ich das wie folgt:

$$f(0 + x) = f(x) = x^2 = (-x)^2 = f(-x) = f(0 - x).$$

Andere Beispiele von Funktionen mit zur y-Achse symmetrischen Graphen sind die Cosinusfunktion, die in Abbildung 6.7 (rechts) gestrichelt dargestellt ist: $\cos(x) = \cos(-x)$, und die Betragsfunktion $|x|$ aus Abbildung 6.6. Abbildung 6.7 (links) veranschaulicht den Graphen der Funktion $g(x) = x^2 - 2x$. Dieser ist symmetrisch zur Achse $x = 1$. Das sehe ich ein, indem ich

$$g(1 + x) = (1 + x)^2 - 2(1 + x) = 1 + 2x + x^2 - 2 - 2x = x^2 - 1$$

und

$$g(1 - x) = (1 - x)^2 - 2(1 - x) = 1 - 2x + x^2 - 2 + 2x = x^2 - 1$$

berechne: Es kommt nämlich – wie Sie sehen können – das Gleiche heraus. Anschaulich erkenne ich dies durch intensive Betrachtung von Abbildung 6.7 (links): $x = 1$ ist Spiegelachse des Graphen von g.

Die Graphen mancher Funktionen f sind **punktsymmetrisch** zu einem Punkt $(a|b)$ im Koordinatensystem. Falls dies so ist, so gilt für alle x:

$$f(a + x) = -f(a - x) + 2b.$$

Anschaulich bedeutet dies, dass der Punkt $(a|b)$ Spiegelpunkt des Graphen von f ist. Ich belasse es hier bei drei Beispielen, in denen die Graphen **punktsymmetrisch zum Ursprung** sind, das heißt $(a|b) = (0|0)$. Der in Abbildung 6.8 dargestellte Graph der Funktion $f(x) = 1/3\,x^3$ besitzt diese Eigenschaft:

$$f(x) = 1/3\,x^3 = -(1/3\,(-x)^3) = -f(-x).$$

Die in Abbildung 6.7 (rechts) dargestellte Sinusfunktion besitzt diese Eigenschaft ebenfalls: $\sin(x) = -\sin(-x)$. Für die gebrochen-rationale Funktion $f(x) = x/(x^2+1)$ aus Abbildung 6.12 weise ich die Punktsymmetrie zum Ursprung wie folgt nach. Ich berechne $f(-x)$, indem ich x im Funktionsterm überall durch $-x$ ersetze:

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

und erkenne nun, dass dies mit $-f(x)$ übereinstimmt. Anschaulich sehe ich dies durch intensive Betrachtung von Abbildung 6.12: $(0|0)$ ist Spiegelpunkt des Graphen von f.

• **Verhalten für große x**: Ist eine Funktion f für große x definiert, so kann ich mir Gedanken darüber machen, wie sich die Funktionswerte $f(x)$ verhalten, wenn x immer größer wird. Man kann somit untersuchen, wie sich die Werte der Funktion im Grenzübergang

$$\lim_{x\to\infty} f(x)$$

für x **gegen unendlich** (in Zeichen: $x \to \infty$) verhalten. Bevor Sie weiterlesen, schauen Sie bitte gegebenenfalls noch einmal nach, was ich Ihnen im letzten Abschnitt über Grenzwerte und deren Existenz erzählt habe. Dies ist prinzipiell auch hier gültig.

Grob gesprochen kann es wiederum zu zwei Situationen kommen. Erstens: Ein eindeutiger Grenzwert existiert, das heißt es gibt eine reelle Zahl S, sodass

$$\lim_{x\to\infty} f(x) = S.$$

Dies bedeutet, dass sich die Funktion f für große x immer mehr der durch S festgelegten **konstanten Funktion** $a_f(x) = S$ annähert. Man nennt dann S eine **Schranke von** f und den Graph von a_f **horizontale Asymptote von** f. Zweitens: Der obige Grenzwert existiert nicht – ist also entweder ∞ oder aber nicht eindeutig festlegbar. In diesem Fall gibt es keine horizontale Asymptote.

Die Funktion $f(x) = 1/x+5$ besitzt die horizontale Asymptote $a_f(x) = 5$. Ich sehe nämlich leicht ein, dass

$$\lim_{x\to\infty} (1/x+5) = 0+5 = 5$$

gilt. Grund: Umso grösser x gewählt wird, umso näher wird der Term $1/x$ bei 0 liegen, also geht $1/x$ gegen 0 für $x \to \infty$. Ein anderes Beispiel ist die gebrochen-rationale Funktion $f(x) = x/(x^2+1)$ aus Abbildung 6.12. Um die Grenzwertberechnung hier durchführen zu können, klammere ich zunächst im Zähler und Nenner jeweils den Term mit dem grössten Exponenten aus und kürze dann:

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

Beachte ich, dass $1/x$ und $1/x^2$ gegen 0 für $x \to \infty$ gehen, so erhalte ich daraus

$$\lim_{x\to\infty} \frac{x}{x^2+1} = \lim_{x\to\infty} \frac{1}{x}\,\frac{1}{1+\frac{1}{x^2}} = 0 \cdot \frac{1}{1+0} = 0.$$

Ich erkenne nun, dass sich der Graph dieses f für große x immer mehr dem Graphen der **Nullfunktion** $a_f(x) = 0$ annähert. Abbildung 6.12 zeigt uns dies: Die Funktionswerte liegen immer näher an der x-Achse, umso größer x gewählt wird. Betrachte ich nun die Funktion $f(x) = x^2/(x^2+1)$, so kann ich genauso vorgehen und erhalte

$$\lim_{x\to\infty} \frac{x^2}{x^2+1} = \lim_{x\to\infty} \frac{1}{1+\frac{1}{x^2}} = \frac{1}{1+0} = 1.$$

Sie sehen: Falls der Term mit dem größten Exponenten im Zähler und Nenner der gleiche ist (im Beispiel ist dies jeweils x^2), so kommt bei der Grenzwertbildung ebenfalls eine reelle Zahl heraus – $a_f(x) = 1$ ist in diesem Beispiel die horizontale Asymptote. Betrachte ich nun aber die Funktion $f(x) = x^3/(x^2+1)$ und gehe genauso vor wie eben, so erhalte ich

$$\lim_{x\to\infty} \frac{x^3}{x^2+1} = \lim_{x\to\infty} x\,\frac{1}{1+\frac{1}{x^2}} = \infty,$$

denn x geht offensichtlich gegen unendlich für $x \to \infty$. Für dieses Beispiel existiert keine horizontale Asymptote, denn die Funktion wächst für große x über jede vorgeschriebene Schranke hinaus. Genauso ist das auch mit den Funktionen deren Graphen in den Abbildungen 6.7 (links), 6.8, 6.9 und 6.10 dargestellt sind. Ein Beispiel der delikaten Art ist die Sinusfunktion aus Abbildung 6.7 (rechts). Diese wächst zwar *nicht* über alle Schranken, denn es gilt $|\sin(x)| \leq 1$ für alle x (vergleiche erster Abschnitt in Kapitel 4), dennoch kann ich den Grenzwert von $\sin(x)$ für $x \to \infty$ nicht eindeutig festlegen, denn die Werte der Sinusfunktion schwingen auch für große x stets zwischen -1 und 1 hin und her. Ich habe sozusagen unendlich viele Kandidaten für den Grenzwert (alle Werte aus $[-1, 1]$) – kann mich aber für keinen entscheiden.

Zum Abschluss dieses Abschnitts betrachte ich noch zwei Beispiele zur Kurvendiskussion:

Beispiel 6.29:
Das erste Beispiel behandelt die Funktionenschar

$$f_t(x) = \frac{1}{t^2}\,x^3 - x,$$

wobei $t > 0$ ein fester Parameter des kubischen Polynoms f_t in der reellen Variablen x ist. Ich berechne zunächst die Nullstellen von f_t. Hierzu klammere ich x aus dem Funktionsterm aus:

$$(1/t^2)\,x^3 - x = x\,((1/t^2)\,x^2 - 1).$$

Jetzt sehe ich, dass $x_0 = 0$ eine Nullstelle von f_t ist: $f_t(0) = 0$. Nun begebe ich mich auf die Suche nach eventuellen weiteren Nullstellen. Hierzu betrachte ich die Gleichung $(1/t^2)\,x^2 - 1 = 0$, denn *ein Produkt von Termen ist genau dann* 0, *wenn einer dieser Terme verschwindet.* Addiere ich hier die Zahl 1 auf beiden Seiten und multipliziere ich danach beide Seiten mit t^2, so sehe ich, dass diese Gleichung sich auch als $x^2 = t^2$ schreiben lässt. Die Lösungen sind somit $x_1 = t$ und $x_2 = -t$. Ich kann jetzt f_t folgendermaßen als Produkt von Linearfaktoren hinschreiben:

$$f_t(x) = (1/t^2)\,x\,(x-t)\,(x+t).$$

Ich stelle nebenbei fest, dass der Graph von f_t punktsymmetrisch zum Ursprung ist. Um Sie zu überzeugen, rechne ich das nach, indem ich $f_t(-x)$ ermittle:

$$f_t(-x) = (1/t^2)\,(-x)^3 - (-x) = (-1/t^2)\,x^3 + x = -((1/t^2)\,x^3 - x)$$

und erkenne, dass dies mit $-f(x)$ übereinstimmt. Nun interessieren mich lokale Extrema von f_t. Hierzu bilde ich die erste Ableitung (nach x):

$$f_t'(x) = (3/t^2)\,x^2 - 1.$$

Um Kandidaten für lokale Extrema von f_t zu finden, muss ich die Nullstellen von f_t' finden: $(3/t^2)\,x^2 - 1 = 0$. Ich addiere auch hier die Zahl 1 auf beiden Seiten, multipliziere danach beide Seiten mit $t^2/3$ und habe die Gleichung in die folgende Form gebracht: $x^2 = t^2/3$. Jetzt sehe ich, dass $x_3 = \sqrt{t^2/3} = (\sqrt{3}/3)\,t$ und $x_4 = -(\sqrt{3}/3)\,t$ die Kandidaten für lokale Extrema von f_t sind. Ich muss jetzt checken, ob es sich auch tatsächlich um lokale Extrema handelt. Hierzu bilde ich die zweite Ableitung von f_t (nach x):

$$f_t''(x) = (6/t^2)\,x.$$

Einsetzen ergibt: $f_t''(x_3) = (6/t^2)\,(\sqrt{3}/3)\,t = (2\sqrt{3}t)/t^2 = (2\sqrt{3})/t > 0$ und $f_t''(x_4) = (-2\sqrt{3})/t < 0$. Somit ist $T = (x_3|f_t(x_3)) = ((\sqrt{3}/3)t\,|\,(-2\sqrt{3}t)/9)$ der Tiefpunkt des Graphen von f_t und $H = (x_4|f_t(x_4)) = (-(\sqrt{3}/3)t\,|\,(2\sqrt{3}t)/9)$ der Hochpunkt des Graphen von f_t. Übrigens: Wenn ich den Parameter t hier durch die x-Koordinaten der Tiefpunkte ausdrücke, also $t = (3/\sqrt{3})x_3$, und diesen dann in die y-Koordinaten der Tiefpunkte einsetze $y_3 = (-2\sqrt{3}t)/9 = (-2\sqrt{3}\,(3/\sqrt{3})x_3)/9 = (-2/3)\,x_3$, so erhalte ich eine Funktion $g(\bar{x}) = (-2/3)\,\bar{x}$, $\bar{x} > 0$, deren Graph die Lage aller Tiefpunkte beschreibt – dieser wird manchmal **Ortskurve** der Tiefpunkte genannt. Sehen Sie, dass $g(\bar{x}) = (-2/3)\,\bar{x}$ für $\bar{x} < 0$ auch die **Ortskurve** der Hochpunkte ist? Jetzt interessieren mich noch die Wendestellen von f_t. Hierzu setze ich die zweite Ableitung f_t'' gleich 0: $(6/t^2)\,x = 0$, und ich erkenne, dass die Stelle $x_0 = 0$ ein Kandidat für eine Wendestelle ist. Ich muss jetzt checken, ob es sich auch tatsächlich um eine Wendestelle handelt. Hierzu bilde ich die dritte Ableitung $f_t'''(x) = 6/t^2$. Dies ist offenbar eine konstante Funktion – keine große Überraschung, denn f_t ist ein kubisches Polynom der Variablen x. Ich erkenne, dass f_t''' stets – also auch für die Stelle x_0 – ungleich 0 ist. Somit ist der Punkt $W = (0|f(0)) = (0|0)$ der Wendepunkt von f_t.

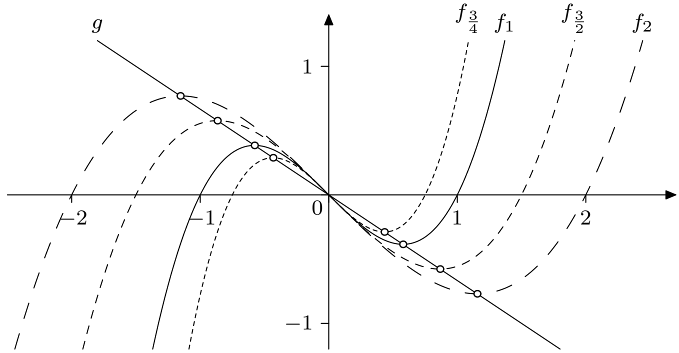

Abb. 6.14 Der Graph des kubischen Polynoms $f_t(x) = \frac{1}{t^2}\, x^3 - x$ für verschiedene Wahlen des Parameters t. Die Extrempunkte von f_t liegen auf der Geraden g.

In Abbildung 6.14 habe ich Ihnen den Verlauf des Graphens von f_t für verschiedene Wahlen von Parametern $t > 0$ zusammen mit der Ortskurve der Hoch- und Tiefpunkte illustriert. ■

Übungsaufgabe 6.15:
Führen Sie eine möglichst vollständige Kurvendiskussion für die Funktionenschar $f_t(x) = t^3/x^2 - t$, wobei $t > 0$ ein fester Parameter ist. ■

Beispiel 6.30:
Das zweite Beispiel ist von eher praktischer Natur. Stellen Sie sich vor, Sie sind der Leiter der Forschungsabteilung einer Nahrungsmittelfabrik, die unter anderem Ravioli-Gerichte in Dosen vertreibt. Das Füllvolumen einer zylinderförmigen Dose sollte – so wollen das Ihre langjährigen Kunden – ein Liter sein. Sie stellen sich aus Kostengründen die Frage, welche Höhe h und welchen Radius r (siehe Abbildung 6.15 (links)) die Dosen haben sollten, sodass der Blechverbrauch minimal wird. Um diese Frage zu beantworten, können Sie die hier entwickelten Methoden verwenden. Zunächst sollte man sich hierzu an die Formel für das Volumen V eines Zylinders erinnern: $V = \pi r^2 h$. Durch die Forderung, dass ein Liter in die Dose passen sollte ($V = 1$), kann ich nun die Höhe h durch den Radius r ausdrücken: $h = 1/(\pi r^2)$. Der Verbrauch an Blech wird durch die Fläche des Mantels O bestimmt. Diese lässt sich für einen Zylinder bekanntlich wie folgt bestimmen: $O = 2\pi r^2 + 2\pi r h$. Setze ich $h = 1/(\pi r^2)$ hier ein, so beschreibt die Funktion

$$O(r) = 2\,(\pi r^2 + \frac{1}{r}),\; r > 0$$

den Flächeninhalt des Mantels in Abhängigkeit vom Radius r unter der Bedingung $V = 1$. Mein Ziel ist es, diese Funktion zu minimieren. Ich suche hierzu erst einmal nach lokalen Minima. Ich berechne somit die erste Ableitung von O nach der Variablen r:

$$O'(r) = 2\,(2\pi r - 1/r^2),$$

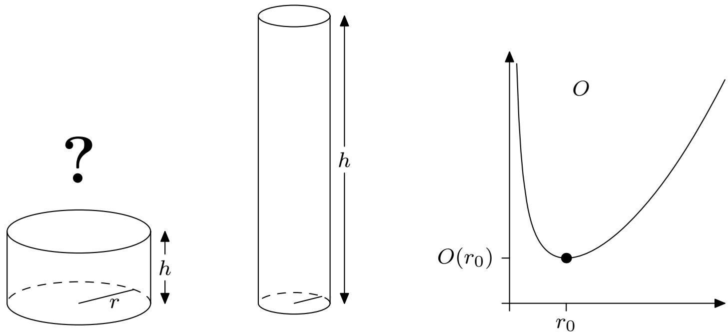

Abb. 6.15 Der Forschungsleiter einer Nahrungsmittelfabrik fragt sich: Für welche Wahl des Radius r ist die Mantelfläche $O(r)$ eines Zylinders mit festem Volumen minimal?

und versuche herauszufinden, an welchen Stellen diese verschwindet: $2\pi r - 1/r^2 = 0$. Ich erweitere jetzt den ersten Term mit r^2 und erhalte $(2\pi r^3 - 1)/r^2 = 0$, das heißt $2\pi r^3 - 1 = 0$. Auflösen nach r ergibt einen Kandidaten für ein lokales Extremum: $r_0 = \sqrt[3]{1/(2\pi)}$. Jetzt brauche ich die zweite Ableitung von O:

$$O''(r) = 4\,(\pi + 1/r^3)$$

und dort setze ich zum Checken meinen Kandidaten r_0 ein:

$$O''(r_0) = 4\,(\pi + 1/r_0^3) = 4\,(\pi + 2\pi) = 12\,\pi > 0.$$

Glück gehabt: $r_0 = \sqrt[3]{1/(2\pi)} = (2\pi)^{-1/3} \approx 0.54$ ist tatsächlich ein lokales Minimum. Der entsprechende Tiefpunkt T ist $T = (r_0|O(r_0)) = (\sqrt[3]{1/(2\pi)}|3\sqrt[3]{2\pi})$, denn

$$O(r_0) = 2\,(\pi r_0^2 + \frac{1}{r_0}) = 2\pi\,(2\pi)^{-\frac{2}{3}} + 2\,(2\pi)^{\frac{1}{3}} = 3\,(2\pi)^{\frac{1}{3}} \approx 5.54.$$

Die zu r_0 gehörige Höhe h_0 ist $h_0 = 1/(\pi r_0^2) = 2^{2/3}\pi^{-1/3} \approx 1.08$. Für diese Funktion O ist nun klar, dass T sogar das einzige *globale Minimum* ist, denn sonst müsste mir eben – ähnlich wie dies im Eingangsbeispiel „Pfälzer Wald“ ist – noch irgendwo ein lokales Maximum begegnet sein. Um aber ganz sicher zu gehen, überprüfe ich hier noch, wie sich die Funktion O für $r \to \infty$ und $r \to 0$ – also zum Rand ihres Definitionsbereichs hin – verhält. Da der Term r^2 für $r \to \infty$ über alle Schranken wächst – also unendlich groß wird –, folgt

$$\lim_{r\to\infty} O(r) = \lim_{r\to\infty} 2\,\left(\pi r^2 + \frac{1}{r}\right) = \infty.$$

Da der Term $1/r$ für $r \to 0$ ebenfalls über alle Schranken wächst, folgt zudem

$$\lim_{\substack{r\to 0\\ r>0}} O(r) = \lim_{\substack{r\to 0\\ r>0}} 2\,\left(\pi r^2 + \frac{1}{r}\right) = \infty.$$

Dies bedeutet insbesondere, dass zum Rand des Definitionsbereichs von O hin ganz sicher keine Stellen mit kleineren Werten von O wie in r_0 auftreten. Den ungefähren

Verlauf des Graphen von O habe ich in Abbildung 6.15 (rechts) für Sie dargestellt. Zusammenfassend lautet meine Empfehlung somit, Dosen mit einem Radius von etwa 5,4 cm und der Höhe von etwa 10,8 cm zu verwenden. Die Fläche des verwendeten Blechs pro Dose ist dann ungefähr 55,4 cm^2. ■

Vermutlich sind Sie nicht in der Forschungsabteilung einer Nahrungsmittelfabrik tätig – aber vielleicht tragen Sie sich mit dem Gedanken, ein Haus zu bauen:

Übungsaufgabe 6.16:
Für Ihre neue Villa benötigen Sie einen rechteckigen Bauplatz mit 2 Quadratkilometern Fläche. Da die Kosten für die Villa immens sind, machen Sie sich Gedanken darüber, an welcher Stelle Sie Einsparungen vornehmen könnten. Nach längeren Überlegungen kommen Sie auf die Idee, die Länge der Umzäunung des auszuwählenden Grundstücks – also dessen Umfang – zu minimieren. Welche Längen müssten die Seiten dieses rechteckigen Grundstücks idealerweise haben? ■

6.3 Integration von Funktionen

Nachdem Sie die Anwendungen von Ableitungen kennen gelernt haben, zeige ich Ihnen jetzt ein weiteres zentrales, mathematisches Werkzeug zur Analyse von reellwertigen Funktionen: Die **Integration**. Diese ermöglicht es, für viele Funktionen den Flächeninhalt unter deren Graphen zu bestimmen. Interessanterweise sind die Integration und die Ableitung von Funktionen in gewisser Weise Umkehrungen voneinander. Nicht nur ich bin der Meinung, dass dies sicherlich eine der grundlegenden Entdeckung (ich sage bewusst *nicht* „Erfindung“) in der Mathematik war. Diese haben wir im Wesentlichen zwei Mathematikern, die im 17. Jahrhundert lebten, zu verdanken: Gottfried Wilhelm Leibniz und Sir Isaac Newton.

Ich beginne mit zwei sehr überschaubaren Beispielen. In Abbildung 6.16 (links) sehen Sie den Graphen der **konstanten** Funktion $f(x) = 1$. Wenn x ein Intervall $[a, b]$ durchläuft, umschließt der entsprechende Teil des Graphen von diesem f mit der x-Achse eine Rechtecksfläche, deren Inhalt $F_a(b)$ offenbar gleich $b - a$ ist. Da das gestrichelt umrandete Rechteck den Flächeninhalt $F(b) = b$ hat und der Inhalt der Fläche des weißen Rechtecks mit Grundseite $[0, a]$ offenbar $F(a) = a$ ist, kann man $F_a(b)$ als Differenz dieser Flächeninhalte ausdrücken: $F_a(b) = F(b) - F(a)$. Etwas schwieriger – aber nicht wirklich schwer – ist die Berechnung solcher umschließender Flächen für die **Identität** $f(x) = x$. Wenn x das Intervall $[a, b]$ durchläuft, umschließt der entsprechende Teil des Graphen von diesem f mit der x-Achse eine Fläche $F_a(b)$, deren Inhalt $1/2\, b^2 - 1/2\, a^2$ ist. In Abbildung 6.16 (rechts) erkennen Sie nämlich, dass sich diese Fläche aus der Differenz des Inhalts der Fläche $F(b)$ des gestrichelt umrandeten Dreiecks mit Grundseite $[0, b]$ mit dem Inhalt der Fläche $F(a)$ des weißen Dreiecks mit Grundseite $[0, a]$ ergibt. Da $F(b)$ gerade die Hälfte des Flächeninhalts eines Quadrats mit Seitenlänge b ist, gilt $F(b) = 1/2\, b^2$. Genauso sehe ich ein, dass $F(a) = 1/2\, a^2$ gilt.

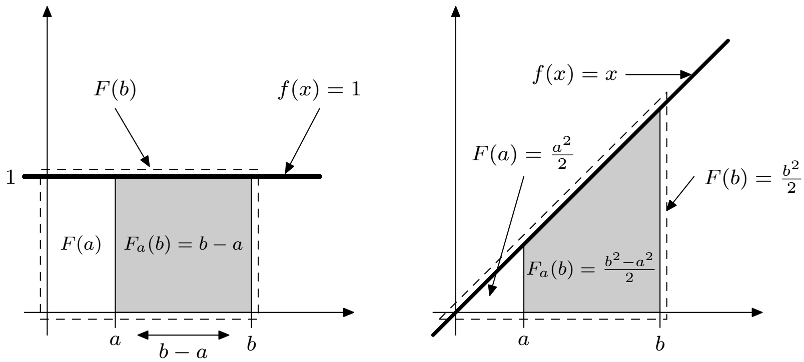

Abb. 6.16 Berechnung bestimmter Intergrale für $f(x) = 1$ (links) und $f(x) = x$ (rechts). Der Graph von f umschließt bei Durchlauf von $[a, b]$ mit der x-Achse die graue Fläche mit Inhalt $F_a(b)$. Dieser ist gerade die Differenz der Inhalte der umstrichelten und der weißen Fläche: $F(b) - F(a)$.

Meine Wahl der bei erster Betrachtung vielleicht etwas seltsam erscheinenden Bezeichnungen $F_a(b), F(b)$ und $F(a)$ hat einen guten Grund, der Ihnen im Verlauf der folgenden Seiten klar werden wird. Genauso wird es sicher mit der Verwendung des **Integralzeichens** sein, mit dessen Hilfe ich hier schon unsere ersten beiden, eben erzielten Ergebnisse zur Berechnung **bestimmter Integrale** formuliere:

$$F_a(b) = \int_a^b 1\,dx = F(b) - F(a) = b - a \text{ und } F_a(b) = \int_a^b x\,dx = F(b) - F(a) = \frac{b^2}{2} - \frac{a^2}{2}.$$

Interessanter – aber auch schwieriger – wird die Bestimmung des Flächeninhalts für Funktionen, deren Graph eine gekrümmte Kurve ist. Das erste Beispiel, das mir hierzu einfällt, ist die **quadratische Potenzfunktion** $f(x) = x^2$. Abbildung 6.17 (rechts) zeigt, dass sich der Inhalt der umschließenden Fläche $F_a(b)$ des Graphen von diesem f auf $[a, b]$ mit der x-Achse (grau) aus der Differenz der Inhalte der umschließenden Fläche $F(b)$ von diesem Graphen hinsichtlich $[0, b]$ (gestrichelt umrandet) und der umschließenden Fläche $F(a)$ von diesem Graphen hinsichtlich $[0, a]$ (umrandete weiße Fläche) ergibt. Da der Graph von f gekrümmt ist, kann ich diese Flächeninhalte $F(b)$ und $F(a)$ nicht so einfach wie in den obigen Beispielen direkt berechnen. Ich versuche deshalb zunächst, einen dieser beiden Flächeninhalte – sagen wir $F(b)$ – *näherungsweise* zu berechnen. Dies tue ich, indem ich f durch **stückweise konstante Funktionen** approximiere (siehe Abbildung 6.17 (links)), denn deren Flächeninhalte kann ich leicht wie oben ausrechnen. Hierzu unterteile ich das Intervall $[0, b]$ in n Teilintervalle $I_{j,n}$, wobei $j \in \{0, \ldots, n-1\}$. Testweise wähle ich $I_{j,n} = [j\,(b/n), (j+1)\,(b/n)]$. Die Länge von $I_{j,n}$ ist $(j+1)\,(b/n) - j\,(b/n) = b/n$. Je größer also n ist, umso näher ist diese Länge bei 0. In jedem $I_{j,n}$ wähle ich nun eine Stelle aus und betrachte den dortigen Funktionswert von f. Testweise wähle ich den Wert am rechten Intervallrand von $I_{j,n}$: $f((j+1)\,b/n)$. Der Inhalt der Fläche unter dem Graphen von f hinsichtlich

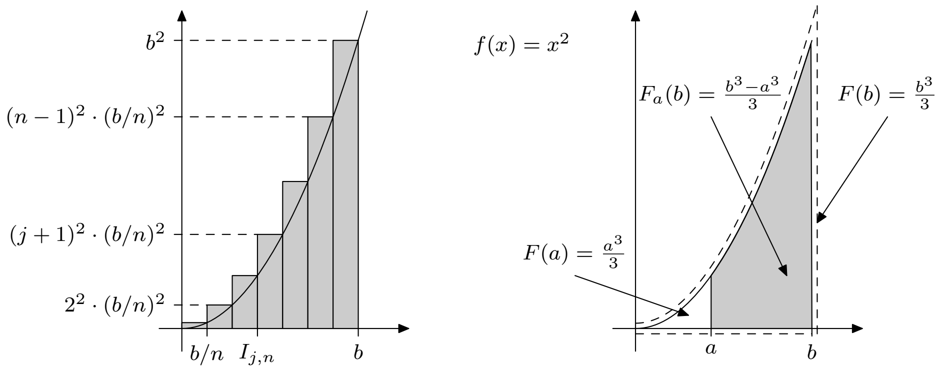

Abb. 6.17 Näherung des Flächeninhalts $F(b)$ für $f(x) = x^2$ durch Verwendung einer stückweise konstanten Funktion (links). Der Graph von diesem f umschließt bei Durchlauf von $[a, b]$ mit der x-Achse die graue Fläche mit Inhalt $F_a(b) = F(b) - F(a)$ – dies ist wiederum die Differenz der Inhalte der umstrichelten und der weißen Fläche (rechts).

$I_{j,n}$ ist dort *ungefähr* der Inhalt der Fläche des Graphen der konstanten Funktion mit Wert $f((j+1)\,b/n) = (j+1)^2\,(b/n)^2$. Der Inhalt der zugehörigen Rechtecksfläche ist offenbar $(j+1)^2\,(b/n)^3$, denn die Seitenlängen dieses Rechtecks sind b/n und $(j+1)^2\,(b/n)^2$. Die Summe der Flächeninhalte $\mathcal{F}_n$ von *allen* diesen Rechtecken ergibt eine Näherung an die gesuchte Fläche $F(b)$. Ich kann $\mathcal{F}_n$ wie folgt hinschreiben:

$$\begin{aligned}\mathcal{F}_n &= 1^2 \left(\frac{b}{n}\right)^3 + 2^2 \left(\frac{b}{n}\right)^3 + \ldots + (j+1)^2 \left(\frac{b}{n}\right)^3 + \ldots + n^2 \left(\frac{b}{n}\right)^3 \\ &= (1 + 4 + \ldots + (j+1)^2 + \ldots + n^2) \left(\frac{b}{n}\right)^3 = \sum_{j=0}^{n-1} (j+1)^2 \left(\frac{b}{n}\right)^3\end{aligned}$$

Für kleine n wie in Abbildung 6.17 (links) – dort zeige ich den Fall $n = 6$ – ist diese Näherung an $F(b)$ noch sehr grob. Die Hoffnung ist aber, dass ich so den Inhalt $F(b)$ *beliebig genau annähern* kann, indem ich sehr große n wähle. Sie sollten mir nun glauben, dass die Summe der ersten n Quadratzahlen $1+4+\ldots+(j+1)^2+\ldots+n^2$ sich durch die folgende Formel vereinfacht schreiben lässt: $(n(n+1)(2n+1))/6$. Um dies nachzuweisen, könnte ich **vollständige Induktion** verwenden. Das lasse ich heute, denn vermutlich werden Sie sich erst später mit dieser Beweistechnik beschäftigen müssen. Die vereinfachende Formel verwendend ergibt sich nun

$$\mathcal{F}_n = \frac{n(n+1)(2n+1)}{6} \left(\frac{b}{n}\right)^3 = \frac{b^3}{6} \left(1 + \frac{1}{n}\right) \left(2 + \frac{1}{n}\right),$$

denn n lässt sich einmal wegkürzen und $(n+1)/n = 1+1/n$ sowie $(2n+1)/n = 2+1/n$. Wähle ich nun immer größere n, also immer feinere Unterteilungen in Teilintervalle, so hoffe ich, dass die Näherungsinhalte $\mathcal{F}_n$ den gesuchten Flächeninhalt $F(b)$

immer besser approximieren. Meine Hoffnung wird hier glücklicherweise erfüllt: Der **Grenzübergang** für $n \to \infty$ führt auf einen *eindeutigen* **Grenzwert**

$$\lim_{n\to\infty} \mathcal{F}_n = \lim_{n\to\infty} \frac{b^3}{6}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right) = \frac{b^3}{6}\lim_{n\to\infty}\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right) = \frac{b^3}{6}\,2 = \frac{b^3}{3},$$

denn $1/n$ geht gegen 0 für $n \to \infty$. $F(b)$ ist hier in der Tat gleich $b^3/3$. Indem ich gedanklich b durch a ersetze, kann ich auf genau dieselbe Art erkennen, dass $F(a)$ den Wert $a^3/3$ hat. Somit folgt aus der Anfangsüberlegung, dass der Inhalt der umschließenden Fläche $F_a(b)$ des Graphen von $f(x) = x^2$ auf $[a, b]$ mit der x-Achse den Wert $F(b) - F(a) = b^3/3 - a^3/3$ hat. Sie haben sicherlich nichts dagegen, dass ich dieses Ergebnis zur Berechnung des bestimmten Integrals von $f(x) = x^2$ in der obigen Form formuliere:

$$F_a(b) = \int_a^b x^2\,dx = F(b) - F(a) = \frac{b^3}{3} - \frac{a^3}{3}.$$

Ein bisschen Übung kann jetzt bestimmt nicht schaden:

Übungsaufgabe 6.17:
a) Bestimmen Sie den Inhalt der Fläche, die der Graph von $f(x) = x^2$ bei Durchlauf von $[1, 3]$ mit der x-Achse umschließt. b) Bestimmen Sie die Stelle $t > 1$, sodass der Graph von f aus a) bei Durchlauf von $[1, t]$ mit der x-Achse eine Fläche mit Inhalt $7/3$ umschließt. ■

Sicher ist Ihnen bei der Bildung der bisherigen bestimmten Integrale schon eine gewisse Regelmäßigkeit aufgefallen. Sie sehen, worauf ich hinauswill, wenn ich verrate, dass $\int_a^b x^3\,dx = b^4/4 - a^4/4$ und $\int_a^b x^4\,dx = b^5/5 - a^5/5$ gelten. Die *Grobregel* zur Ermittlung des bestimmten Integrals einer allgemeinen **Potenzfunktion** $p_i(x) = x^i$ hinsichtlich $[a, b]$ ist also *gewissermaßen umgekehrt* zur Grobregel zum Ableiten (siehe erster Abschnitt in diesem Kapitel) solcher Funktionen. Sie lautet: Erhöhe den Exponenten um eins und teile durch diesen neuen Exponenten $x^{i+1}/(i + 1)$ – dann setze b und a in den so entstandenen Term jeweils ein und bilde die Differenz.

Integration der Potenzfunktion $p_i(x) = x^i$, wobei $i \in \mathbb{N} \cup \{0\}$

Das bestimmte Integral der **Potenzfunktion** $p_i(x) = x^i$ hinsichtlich $[a, b]$ ist gegeben durch

$$\int_a^b x^i\,dx = \frac{b^{i+1}}{i+1} - \frac{a^{i+1}}{i+1} = \frac{b^{i+1} - a^{i+1}}{i+1}.$$

Natürlich ist die Berechnung von bestimmten Integralen nicht nur auf Potenzfunktionen beschränkt. Ich denke, es ist nun schon an der Zeit, Ihnen zu verraten, was man allgemein unter der **Integrierbarkeit** von Funktionen versteht und was ein **bestimmtes Integral** einer reellwertigen Funktion ist:

Integrierbarkeit und bestimmtes Integral

Eine Funktion f heißt auf $[a, b]$ **integrierbar**, falls für jede Unterteilung von $[a, b]$ in n Teilintervalle $I_{j,n} = [x_{j,n}, x_{j+1,n}]$ deren Längen $x_{j+1,n} - x_{j,n}$ gegen 0 für $n \to \infty$ gehen und für jede Wahl von Punkten $\tilde{x}_{j,n} \in I_{j,n}$, wobei $j \in \{0, \ldots, n-1\}$, der Grenzwert der Folge

$$\mathcal{F}_n = f(\tilde{x}_{0,n})(x_{1,n} - x_{0,n}) + \ldots + f(\tilde{x}_{j,n})(x_{j+1,n} - x_{j,n}) + \ldots + f(\tilde{x}_{n-1,n})(x_{n,n} - x_{n-1,n})$$

für $n \to \infty$ existiert und jeweils gleich ist. In diesem Fall heißt der Grenzwert

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \mathcal{F}_n = \lim_{n\to\infty} \sum_{j=0}^{n-1} f(\tilde{x}_{j,n})(x_{j+1,n} - x_{j,n})$$

(sprich: „Integral von a bis b über f nach dx") das **bestimmte Integral von f hinsichtlich** $[a, b]$.

Ich vermute, dass diese Definition für den Normalbürger bei erstmaligem Überfliegen nicht unbedingt sofort sonnenklar und intuitiv vollkommen einsichtig ist. Auch wenn Sie mit großer Wahrscheinlichkeit meine Sicht der Dinge teilen, dass es sich hierbei um eine klare Darstellung des Begriffs der Integrierbarkeit von Funktionen handelt, so sehe ich mich jetzt doch veranlasst, einige nähere Erläuterungen abzugeben.

Das Integralzeichen tritt hier lediglich bei der Bezeichnung eines speziellen Grenzwerts auf. Somit muss ich – wie immer, wenn es um Grenzwerte geht – darauf achten, dass dieser Grenzwert *existiert* (eine *reelle* Zahl ist) und *eindeutig festlegbar* ist (keine Auswahl von mehreren Kandidaten besteht). Bei näherer Betrachtung sieht man dann, um welche Art des Grenzübergangs es sich handelt. Ähnlich wie in meinem (etwas vereinfachenden) Beispiel mit $f(x) = x^2$ sind die $\mathcal{F}_n$ Summen von Flächeninhalten von Rechtecken der Seitenlängen $x_{j+1,n} - x_{j,n}$ (die Länge des Intervalls $I_{j,n}$) und $f(\tilde{x}_{j,n})$ (ein Funktionswert von f an einer Stelle $\tilde{x}_{j,n}$ aus $I_{j,n}$). Diese Flächeninhalte sind somit $f(\tilde{x}_{j,n})\,(x_{j+1,n} - x_{j,n})$. Die Summen $\mathcal{F}_n$ sind Näherungen an den Inhalt der Fläche, den der Graph von f mit der x-Achse beim Durchlauf von $[a, b]$ umschließt. Ich betrachte hier aber jeweils *alle* Unterteilungen von $[a, b]$ in n Teilintervalle $I_{j,n}$ der Länge $x_{j+1,n} - x_{j,n}$, die für größer werdende n immer näher bei 0 liegen. Mehr noch: Gleichzeitig schaue ich mir für jede solche Unterteilung *alle* Möglichkeiten an, einen Funktionswert von f an einer Stelle $\tilde{x}_{j,n}$ aus $I_{j,n}$ zu wählen. Falls nun für alle diese Wahlen von Unterteilungen und für alle Wahlen von Stellen die so entstehenden Näherungsinhalte $\mathcal{F}_n$ im Grenzübergang für $n \to \infty$ auf einen *eindeutigen, reellen* Grenzwert $\lim\limits_{n\to\infty} \mathcal{F}_n$ führen, so wird f auf $[a, b]$ integrierbar genannt, und den Grenzwert selbst nennt man das bestimmte Integral von f hinsichtlich $[a, b]$. Die Hoffnung ist also, dass dieses für alle so gebildeten Näherungsinhalte $\mathcal{F}_n$ beliebig genau angenähert wird.

Leibniz stand vor der die (mathematische) Welt bewegenden Frage, wie er diesen Grenzwert dann am besten bezeichnen könnte. Der Grenzwert entsteht aus Summen. Deshalb liegt es nah, für den Grenzwert ein ähnliches Zeichen wie das Summenzeichen $\sum$ zu verwenden: $\int$. Jetzt kennen Sie die *Bedeutung des* **Integralzeichens** $\int$. Die einzelnen Terme $f(\tilde{x}_{j,n})\,(x_{j+1,n} - x_{j,n})$ der Summen $\mathcal{F}_n$ sind Flächeninhalte von Rechtecken. Wenn n immer größer wird, so sollte die *Differenz* $x_{j+1,n} - x_{j,n}$ immer kleiner werden, und man wird in diesem Prozess *fast alle* x als ein $\tilde{x}_{j,n}$ gewählt haben. Deshalb erscheint es sinnvoll, in der Grenzwertbezeichnung hierfür Symbole zu verwenden, die diesen Prozess intuitiv widerspiegeln: $f(x)\,dx$. Schließlich setzt man a unten und b oben an das $\int$-Zeichen, damit man nicht vergisst, über welchem Intervall man f integrieren möchte. So kommt man insgesamt auf die natürliche Bezeichnung $\int_a^b f(x)\,dx$ für den Grenzwert von $\mathcal{F}_n$ für $n \to \infty$. Übrigens nennt man dann f den **Integranden** (auch: **Integrandenfunktion**), x die **Integrationsvariable** und a, b heißen die **Integrationsgrenzen**.

Möglicherweise haben Sie immer noch das leicht beunruhigende Gefühl, dass es sich bei der Integrierbarkeit von Funktionen um etwas Schwieriges handelt. Man müsste schließlich *alle* Unterteilungen von Teilintervallen $I_{j,n}$ mit immer kleineren Längen und gleichzeitig *jede* Auswahl von Werten von $f(\tilde{x}_{j,n})$ an Stellen $\tilde{x}_{j,n} \in I_{j,n}$ betrachten und dann prüfen, ob der Grenzwert $\int_a^b f(x)\,dx$ von $\mathcal{F}_n$ stets existiert und dann auch noch eindeutig ist. In der Tat: So vorzugehen ist oftmals viel zu aufwendig. Jetzt sehen Sie, warum sich die Herren Newton und Leibniz vermutlich – ich kann sie ja leider nicht mehr persönlich fragen – so ihre berechtigten Gedanken über die Integration von Funktionen gemacht haben. Die eben beschriebene Vorgehensweise war in etwa Stand der Wissenschaft *vor* deren eingangs erwähnte, grundlegende Entdeckung. Diese führte auf eine allgemeine Maschinerie, die – wie ich Ihnen unten zeigen werde – die Integration von Funktionen wesentlich vereinfacht.

Darüber hinaus habe ich eine weitere, sehr positive Nachricht für Sie: Fast alle Funktionen, die Ihnen normalerweise begegnen, sind integrierbar. So sind beispielsweise **monotone Funktionen** *stets* integrierbar. Auch **ableitbare Funktionen** – ja sogar eine allgemeinere Klasse von Funktionen, die man **stetige Funktionen** nennt – sind *immer* integrierbar. Sehr grob gesprochen sind dies Funktionen, deren Graph man zeichnen kann, ohne den Schreiber abzusetzen. Etwas genauer bedeutet dies, dass (Definitions-)Lücken, Sprünge in den Funktionswerten und Stellen, an denen die Funktionswerte in gewisser Weise übermäßig stark oszillieren, nicht auftreten dürfen. Sehr viele Funktionen sind stetig und die *Klasse aller integrierbarer Funktionen* ist sogar erfreulicherweise *noch größer*. Man muss schon etwas basteln, um eine Funktion zu finden, die nicht integrierbar ist.

Beispiel 6.31:
Ich zeige Ihnen lediglich das folgende Beispiel einer auf $[0, 1]$ *nicht* integrierbaren Funktion f:

$$f(x) = \begin{cases} 0, & \text{falls} \quad x \in [0,1] \cap \mathbb{Q} \\ 1, & \text{falls} \quad x \in [0,1] \cap (\mathbb{R} \setminus \mathbb{Q}) \end{cases}.$$

$\mathbb{Q}$ bezeichnet hier, wie in Abschnitt 1.2, die rationalen Zahlen.

Diese Funktion hat natürlich insbesondere keine Chance, monoton oder gar stetig zu sein, denn um deren Graph zu zeichnen, müsste ich den Schreiber unendlich oft zwischen 0 und 1 hin- und herbewegen. ■

Das kann ich nicht und deshalb bewege ich mich nun selbst in eine positive Richtung, indem ich von jetzt an – auch wenn ich dies nicht jedes Mal explizit erwähne – *immer davon ausgehe*, dass die betrachtete(n) Funktion(en) auf den entsprechenden Intervallen integrierbar sind. Ich nenne jetzt erst einmal natürliche Eigenschaften des bestimmten Integrals solcher Funktionen:

Eigenschaften des bestimmten Integrals

a) $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$ **(Intervall-Additivität)**

b) $\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$ **(Vertauschung der Integrationsgrenzen)**

c) $\int_a^a f(x)\,dx = 0$ **(Leeres Integral)**

d) Falls $f(x) \geq 0$ für alle $x \in [a,b]$, so gilt: $\int_a^b f(x)\,dx \geq 0$ **(Positivität)**

Um diese Eigenschaften einzusehen, sollte ich mich daran erinnern, dass das bestimmte Integral von f hinsichtlich $[a,b]$ der Grenzwert von Summen $\mathcal{F}_n$ ist, deren Terme von der Bauart $f(\tilde{x}_j)(x_{j+1,n} - x_{j,n})$ sind. Eigenschaft a) besagt, dass das bestimmte Integral hinsichtlich $[a,b]$ die Summe der bestimmten Integrale hinsichtlich der beiden Teilintervalle $[a,c]$ und $[c,b]$ ist. Diese wird einsichtig, indem ich mir die Summen der Bauart $\mathcal{F}_n$ für $[a,c]$ und $[c,b]$ anschaue und klar mache, dass deren Summe von der Form $\mathcal{F}_n$ für $[a,b]$ ist. Um b) einzusehen, stellt man zunächst fest, dass beim Durchlauf von $[a,b]$ in umgekehrter Reihenfolge die Terme $f(\tilde{x}_j)\,(x_{j,n} - x_{j+1,n}) = -f(\tilde{x}_j)\,(x_{j+1,n} - x_{j,n})$ bei der Bildung der Summe $\mathcal{F}_n$ auftreten. Das Minuszeichen bleibt hier bei der Grenzwertbildung erhalten und deshalb tritt es bei der Vertauschung der Integrationsgrenzen in b) auf. Aus a) und b) folgt aber c):

$$\int_a^a f(x)\,dx \overset{\text{a)}}{=} \int_a^b f(x)\,dx + \int_b^a f(x)\,dx \overset{\text{b)}}{=} \int_a^b f(x)\,dx - \int_a^b f(x)\,dx \overset{\text{c)}}{=} 0.$$

Weiter sehe ich d) wie folgt ein: Aus $f(x) \geq 0$ folgt, dass alle Terme $f(\tilde{x}_j)\,(x_{j+1,n} - x_{j,n})$ ebenfalls ≥ 0 sind, denn es ist ja auch $x_{j+1,n} - x_{j,n} > 0$. Also $\mathcal{F}_n \geq 0$ – und somit ist der Grenzwert für $n \to \infty$ auch ≥ 0.

Jetzt fällt mir auf, dass d) Folgendes impliziert: Falls $f(x) \leq 0$ für alle $x \in [a,b]$, so gilt: $\int_a^b f(x)\,dx \leq 0$. Dies bedeutet, dass das bestimmte Integral manchmal auch *negativ* werden kann. Beispielsweise sind sämtliche Funktionswerte von $f(x) = x$

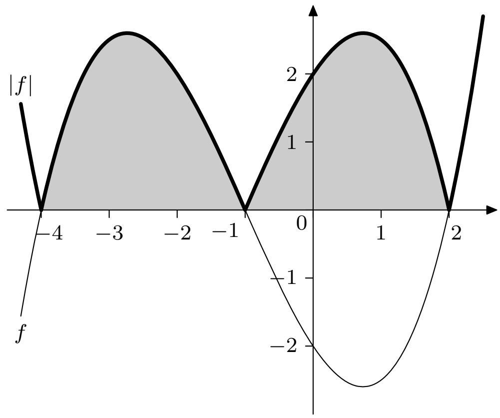

Abb. 6.18 Das unbestimmte Integral von f (dünne Kurve, siehe auch Abbildung 6.10) ergibt im Allgemeinen lediglich den orientierten Flächeninhalt. Um den absoluten Flächeninhalt (grau) zu bestimmen, muss man die Funktion $|f|$ (breite Kurve) integrieren: An allen Stellen, an denen $f(x)$ nicht positiv ist (im Bild: für $x \in [-1, 2]$), wird anstatt f die Funktion $-f$ integriert.

für $x \in [-1, -3]$ nicht positiv und ich berechne nach dem, was ich Ihnen eingangs zeigte,

$$\int_{-3}^{-1} x\,dx = ((-1)^2 - (-3)^2)/2 = -4.$$

Falls eine Funktion im betrachteten Intervall das Vorzeichen wechselt, so kann sogar 0 der Wert des bestimmten Integrals sein:

$$\int_{-1}^{1} x\,dx = (1^2 - (-1)^2)/2 = 0.$$

Sie sehen, dass das unbestimmte Integral im Allgemeinen *nicht* den **absoluten Flächeninhalt** der vom Graphen von f mit der x-Achse umschlossenen Fläche ergibt, sondern lediglich deren **orientierten Flächeninhalt**. Das Vorzeichen des Integranden f bestimmt in den verschiedenen Teilintervallen auch das Vorzeichen des Integrals. Das bestimmte Integral stellt also die Summierung der entsprechenden, *mit Vorzeichen belegten*, Flächeninhalte dar. Wegen

$$\int_{-1}^{0} x\,dx = -1/2 \text{ und } \int_{0}^{1} x\,dx = 1/2$$

verwundert es mich somit nicht mehr, dass ich oben $\int_{-1}^{1} x\,dx = 0$ berechnet habe, denn das muss wegen Eigenschaft a) so sein. Möchte ich den absoluten Flächeninhalt der vom Graphen von f mit der x-Achse umschlossenen Fläche wissen, so sollte ich das bestimmte Integral von der (verketteten) Funktion $|f| = |.| \circ f$ bilden:

$$\int_{a}^{b} |f(x)|\,dx.$$

Praktisch bedeutet das wegen der Festlegung des Betrags $|.|$ (siehe erster Abschnitt in diesem Kapitel), dass ich anstatt f an allen Stellen in $[a, b]$, an denen f nicht positiv ist, die Funktion $-f$ integriere. In Abbildung 6.18 ist der Graph von $|f|$ für ein vorgegebenes f als breitere Kurve dargestellt. Ich werde weiter unten das zu dieser Abbildung gehörige konkrete Beispiel noch ein bisschen detaillierter anschauen. Hier bemerke ich erst einmal noch, dass sich der absolute Flächeninhalt in obigem Beispiel folgendermaßen berechnet:

$$\int_{-1}^{1} |x|\, dx = \int_{-1}^{0} -x\, dx + \int_{0}^{1} x\, dx = -\int_{-1}^{0} x\, dx + \int_{0}^{1} x\, dx = -(-1/2) + 1/2 = 1.$$

Hier habe ich beim zweiten „="-Zeichen den **Faktor** -1 einfach einmal locker vor das Integral gezogen.

Diese „lockere" Rechenoperation sollte ich jetzt aber genauer erklären: Aus dem ersten Abschnitt dieses Kapitels wissen Sie, dass uns die Ableitungsregeln für das **Faktorprodukt** $\lambda \cdot f$, festgelegt durch $(\lambda \cdot f)(x) = \lambda\, f(x)$, und die **Summe von Funktionen** $f + g$, festgelegt durch $(f + g)(x) = f(x) + g(x)$, die Berechnung von Ableitungen oft vereinfacht hat. Ich werde nun ähnliche *Regeln zur Vereinfachung der Berechnung von Integralen* formulieren.

Das bestimmte Integral von $\lambda \cdot f$ hinsichtlich $[a, b]$ ist der Grenzwert von Summen $\mathcal{F}_n$, bei denen die einzelnen Terme von der Bauart $\lambda f(\tilde{x}_j)\,(x_{j+1,n} - x_{j,n})$ sind. Ich kann also λ aus $\mathcal{F}_n$ und bei der Grenzwertbildung für $n \to \infty$ ausklammern. Das bestimmte Integral von $f + g$ hinsichtlich $[a, b]$ ist der Grenzwert von Summen $\mathcal{F}_n$, bei denen die einzelnen Terme von der Bauart

$$(f(\tilde{x}_j) + g(\tilde{x}_j))\,(x_{j+1,n} - x_{j,n}) = f(\tilde{x}_j)\,(x_{j+1,n} - x_{j,n}) + g(\tilde{x}_j)\,(x_{j+1,n} - x_{j,n})$$

sind. Ich kann also zwei Summen der Form $\mathcal{F}_n$ betrachten und diese bei der Grenzwertbildung separat betrachten. Auf diese Art ergeben sich die folgenden beiden einfachen Integrationsregeln.

Integrationsregeln für das Faktorprodukt $\lambda \cdot f$ und die Summe $f + g$

Das bestimmte Integral des Faktorprodukts $\lambda \cdot f$ hinsichtlich $[a, b]$ ist das Produkt von λ mit dem bestimmten Integral von f hinsichtlich $[a, b]$:

$$\int_{a}^{b} (\lambda \cdot f)(x)\, dx = \lambda \int_{a}^{b} f(x)\, dx.$$

Das bestimmte Integral der Summe $f + g$ von Funktionen f und g hinsichtlich $[a, b]$ ist gleich der Summe der bestimmten Integrale von f und g hinsichtlich $[a, b]$:

$$\int_{a}^{b} (f + g)(x)\, dx = \int_{a}^{b} f(x)\, dx + \int_{a}^{b} g(x)\, dx.$$

Endlich kann ich jetzt das bestimmte Integral auch für etwas schwierigere Funktionen als die Potenzfunktionen ausrechnen:

Beispiel 6.32:
Falls ich mich beispielsweise für das bestimmte Integral der Funktion $h(x) = 2x - x^2$ hinsichtlich $[0, 2]$ interessiere, so kann ich die Summen- und Faktorregel anwenden und erhalte zunächst

$$\int_0^2 2x - x^2\, dx = \int_0^2 2x\, dx + \int_0^2 -x^2\, dx = 2 \int_0^2 x\, dx - \int_0^2 x^2\, dx.$$

Somit habe ich das bestimmte Integral von h durch bestimmte Integrale von Funktionen – $f(x) = x$ und $g(x) = x^2$ – ausgedrückt, für die ich diese schon kenne:

$$\int_0^2 x\, dx = 2^2/2 - 0^2/2 = 2 \text{ und } \int_0^2 x^2\, dx = 2^3/3 - 0^3/3 = 8/3.$$

Das Ergebnis ist somit $\int_0^2 h(x)\, dx = \int_0^2 2x - x^2\, dx = 2 \cdot 2 - 8/3 = 4/3$. Eine Interpretation dieses Ergebnis ist, dass die Fläche, die für $x \in [0, 2]$ von der durch $2x$ festgelegten Gerade und der **Normalparabel** umschlossen wird, den Inhalt $4/3$ hat.

■

Beispiel 6.33:
Als nächstes Beispiel berechne ich das bestimmte Integral des kubischen Polynoms $f(x) = 1/4\, x^3 + 3/4\, x^2 - 3/2\, x - 2$ (siehe Abbildung 6.10) hinsichtlich $[-4, -1]$:

$$\int_{-4}^{-1} f(x)\, dx = \frac{1}{4} \int_{-4}^{-1} x^3\, dx + \frac{3}{4} \int_{-4}^{-1} x^2\, dx - \frac{3}{2} \int_{-4}^{-1} x\, dx - 2 \int_{-4}^{-1} 1\, dx$$
$$= \tfrac{1}{4}\,(\tfrac{(-1)^4-(-4)^4}{4}) + \tfrac{3}{4}\,(\tfrac{(-1)^3-(-4)^3}{3}) - \tfrac{3}{2}\,(\tfrac{(-1)^2-(-4)^2}{2}) - 2\,((-1) - (-4)) = \tfrac{81}{16}.$$

Hierbei habe ich wiederum zunächst die Summen- und Faktorregel und dann die obigen Kenntnisse über die bestimmten Integrale der Potenzfunktionen $1, x, x^2$ und x^3 verwendet. Es ist schon recht erfreulich und vielleicht sogar ein wenig erstaunlich, dass ich so den Inhalt der relativ kompliziert berandeten, grau hervorgehobenen Fläche in Abbildung 6.10 ermitteln kann. Die Zahl $81/16$ ist der absolute Inhalt dieser Fläche, denn es gilt $f(x) \geq 0$ für alle $x \in [-4, -1]$. Anders sieht das mit dem bestimmten Integral von diesem f hinsichtlich $[-4, 2]$ aus – hierfür berechne ich durch analoge Vorgehensweise wie eben: $\int_{-4}^2 f(x)\, dx = 0$. Der hinsichtlich $[-4, 2]$ berechnete orientierte Flächeninhalt hat also den Wert 0. Der absolute Flächeninhalt über $[-4, 2]$ hingegen ist $81/8$, denn ich ermittle die graue Fläche in Abbildung 6.18 wie folgt:

$$\int_{-4}^2 |f(x)|\, dx = \int_{-4}^{-1} f(x)\, dx - \int_{-1}^2 f(x)\, dx = \tfrac{81}{16} - (-\tfrac{81}{16}) = \tfrac{81}{8}.$$ ■

Übungsaufgabe 6.18:
a) Berechnen Sie den orientierten und absoluten Flächeninhalt der vom Graphen von $f(x) = x^3 - x$ über $[-1, 1]$ mit der x-Achse umschlossenen Fläche. b) Ermitteln Sie

das bestimmte Integral $\int_{-1}^{1} f(x)\,dx$ für die beiden Funktionen $f = p$ und $f = f_t$ aus Übungsaufgabe 6.10 . ■

Übungsaufgabe 6.19:
Finden Sie für $t > 0$ eine explizite Darstellung des Integrals $\int_0^t -x^2 + 2x\,dx$. Fassen Sie das explizite Ergebnis als *Funktion* F_0 *in* t auf und leiten Sie F_0 spaßeshalber nach t ab. Fällt Ihnen etwas auf? ■

Die Verwendung der Faktor- und Summenregel erlaubt mir, bestimmte Integrale nun für eine größere Klasse von Funktionen als die Potenzfunktionen zu berechnen:

Integration von Polynomen

Das bestimmte Integral einer **Polynomfunktion p vom Grad n**, $p(x) = a_n\,x^n + \ldots + a_i\,x^i + \ldots + a_1\,x + a_0$ hinsichtlich $[a, b]$ ist gegeben durch

$$\int_a^b p(x)\,dx = a_n\left(\tfrac{b^{n+1}-a^{n+1}}{n+1}\right) + \ldots + a_i\left(\tfrac{b^{i+1}-a^{i+1}}{i+1}\right) + \ldots + a_1\left(\tfrac{b^2-a^2}{2}\right) + a_0(b-a).$$

Prima – aber natürlich gebe ich mich *nicht* damit zufrieden, lediglich Polynome zu integrieren. Ich habe Ihnen ja bereits oben verraten, dass die Klasse aller integrierbaren Funktionen recht umfangreich ist. Besonders wünschenswert wäre es deshalb, eine *allgemeine Maschinerie zur Integration* zu besitzen, die für sehr viele dieser Funktionen sozusagen simultan anwendbar ist. Ein mathematischer Traum würde hierbei in Erfüllung gehen, wenn diese Maschinerie so wäre, dass bestimmte Integrale auf ähnlich einfache Art wie für Polynome zu berechnen wären. Ich darf es Ihnen hier schon vorab verraten: Ein solcher Traum wurde vor langer Zeit Realität.

Um mich gemeinsam mit Ihnen an die Vorgehensweise heranzutasten, betrachte ich noch einmal ein Beispiel vom Beginn dieses Kapitels. Oben habe ich Ihnen erklärt, dass

$$F_a(b) = \int_a^b x^2\,dx = \frac{b^3}{3} - \frac{a^3}{3}$$

das bestimmte Integral von $f(x) = x^2$ hinsichtlich $[a, b]$ ist. Ich verwende hier *bewusst wieder* die Bezeichnung $F_a(b)$ vom Anfang dieses Abschnitts, denn ich werde gleich die obere Integrationsgrenze als *Variable der Funktion* F_a auffassen. Die feste Stelle b ersetze ich nun also durch eine Variable – sagen wir t –, damit keine Verwechslung mit x passieren kann. Was soll's, werden Sie vielleicht sagen — im Prinzip steht da doch noch dasselbe:

$$F_a(t) = \int_a^t x^2\,dx = \frac{t^3}{3} - \frac{a^3}{3}.$$

Stimmt. *Jedoch* meine Interpretation ist nun anders: F_a betrachte ich jetzt als eine Funktion in der Variablen t: $F_a(t) = t^3/3 - a^3/3$. Mit dieser Funktion F_a kann ich alles machen, was ich üblicherweise mit Funktionen anstelle: Nullstellen bestimmen, Graphen zeichnen, ..., und insbesondere kann ich das F_a hier spaßeshalber nach *t ableiten.* Da $a^3/3$ eine Konstante ist, berechne ich die Ableitung F_a' von F_a wie folgt:

$$F_a'(t) = \frac{1}{3}\,(t^3)' - \left(\frac{a^3}{3}\right)' = \frac{1}{3}\,3t^2 = t^2.$$

Die entscheidende *Beobachtung* ist hier, dass die Ableitung F_a' von F_a mit der Funktion f, für die ich das Integral bilde (der Integrandenfunktion), übereinstimmt: $F_a'(t) = f(t)$. Da Sie Übungsaufgabe 6.19 gemacht haben, kennen Sie dieses Phänomen schon für eine weitere Funktion. „Zufall" mögen Sie vielleicht sagen – lesen Sie bitte entspannt weiter.

Dieses Phänomen tritt offenbar für jede Potenzfunktion $f(x) = x^i$ auf. Indem ich $F_a(t) = \int_a^t x^i\,dx$ setze und dann nach der Variablen t ableite, sehe ich dies ein:

$$F_a'(t) = \left(\int_a^t x^i\,dx\right)' = \left(\frac{t^{i+1}}{i+1}\right)' - \left(\frac{a^{i+1}}{i+1}\right)' = (i+1)\,\frac{t^i}{i+1} = t^i = f(t).$$

Jetzt kann ich für jede Polynomfunktion $p(x) = a_n\,x^n + \ldots + a_i\,x^i + \ldots + a_1\,x + a_0$ dieses Phänomen beobachten – leite ich $\int_a^t p(x)\,dx$ ab, so erhalte ich p selbst:

$$a_n \overbrace{\left(\frac{t^{n+1} - a^{n+1}}{n+1}\right)'}^{t^n} + \ldots + a_i \overbrace{\left(\frac{t^{i+1} - a^{i+1}}{i+1}\right)'}^{t^i} + \ldots + a_1 \overbrace{\left(\frac{t^2 - a^2}{2}\right)'}^{t} + a_0 \overbrace{(t-a)'}^{1} = p(t).$$

Ich *beobachte* also für die bisher betrachteten f, dass die **Integralfunktion von** f

$$F_a(t) = \int_a^t f(x)\,dx$$

ableitbar ist und deren erste Ableitung F_a' mit der Integrandenfunktion f übereinstimmt: $F_a'(t) = f(t)$. Diese *Beobachtung* war der *Ausgangspunkt* der Entwicklung einer *allgemeinen Maschinerie zur Integration.* Die Funktion F_a wird manchmal auch **unbestimmtes Integral von** f genannt, denn die obere Integrationsgrenze t ist hier variabel und deshalb *nicht näher bestimmt.* Manchmal – insbesondere wenn die untere Grenze *nicht mehr* allzu sehr interessiert – bezeichnet man dieses je nach Fortgeschrittenheitsgrad der persönlichen Schreibfaulheit mit $\int^t f(x)\,dx$ oder $\int^t f$.

Da es sehr viele integrierbare Funktionen gibt, kann man überlegen, für welche Funktionen f möglichst allgemeinerer Art die Integralfunktion $F_a(t) = \int_a^t f(x)\,dx$ ableitbar sein könnte, sodass zusätzlich auch noch $F_a' = f$ gilt. Da f idealerweise von recht allgemeiner Art sein sollte, muss ich mich jetzt an die Vorgehensweise aus dem ersten Abschnitt dieses Kapitels zum Nachweis der **Ableitbarkeit** durch Bildung der **Differenzenquotienten** und des **Differenzialquotienten** erinnern. Ich gehe also wie bei den dortigen Untersuchungen vor: Somit betrachte ich für eine vorgegebene (integrierbare) Funktion f zunächst die Differenzenquotienten der Integralfunktion

$$\frac{F_a(t+h) - F_a(t)}{h} = \frac{\int_a^{t+h} f(x)\,dx - \int_a^t f(x)\,dx}{h}.$$

Indem ich die Eigenschaften b) und a) des bestimmten Integrals für den hier auftretenden Zähler verwende, erhalte ich für diesen

$$\int_a^{t+h} f(x)\,dx + \int_t^a f(x)\,dx = \int_t^a f(x)\,dx + \int_a^{t+h} f(x)\,dx = \int_t^{t+h} f(x)\,dx,$$

und somit sind die Differenzenquotienten der Integralfunktion als $\int_t^{t+h} f(x)\,dx/h$ gegeben. Das hier auftretende Integral $\int_t^{t+h} f(x)\,dx$ kann ich für Funktionen f sehr allgemeiner Art leider nicht mehr weiter vereinfachen. Für jede nicht gar zu verrückte Funktionen f – nämlich für f, deren Graph ich zeichnen kann, ohne den Schreiber abzusetzen – gelingt mir dies jedoch zum Glück. Diese fs habe ich oben *stetige* Funktionen genannt. Für stetige f ist die Darstellung als Rechtecksflächeninhalt

$$\int_t^{t+h} f(x)\,dx = f(\xi)\,h$$

möglich. Hier ist ξ (sprich: „xi") eine Stelle in $[t, t+h]$. Dieses ξ ist im Allgemeinen leider nicht näher spezifiziert – ich weiß lediglich, dass es eine solche Stelle gibt. In Abbildung 6.19 habe ich versucht, diesen **Mittelwertsatz der Integralrechnung** darzustellen. Die Differenzenquotienten der Integralfunktion sind somit:

$$\frac{F_a(t+h) - F_a(t)}{h} = \frac{\int_t^{t+h} f(x)\,dx}{h} = \frac{f(\xi)\,h}{h} = f(\xi).$$

Die Stelle ξ hängt im Allgemeinen von f und der Wahl von t und h ab. Wenn nun h gegen 0 geht, $h \to 0$, so nähert sich $t+h$ immer mehr t an, und $\xi \in [t, t+h]$ hat keine andere Chance, als für $h \to 0$ gegen t zu gehen: $\lim_{h\to 0} \xi = t$. Somit *erscheint es*

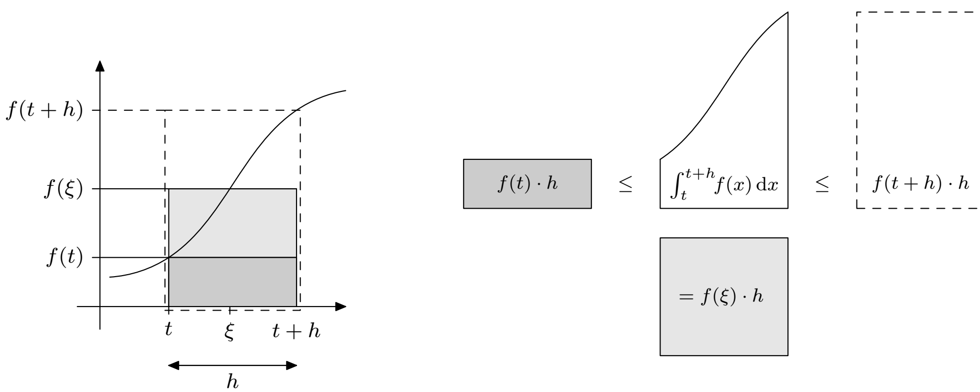

Abb. 6.19 Der Wert des Integrals $\int_t^{t+h} f(x)\,dx$ lässt sich durch den Inhalt der Rechtecksflächen mit Seitenlänge h (grau und umstrichelt) einschließen (rechts oben). Falls f eine *stetige* Funktion ist, so existiert eine Rechtecksfläche mit Seitenlängen h und $f(\xi)$ (rechts unten), sodass deren Inhalt *exakt* der Wert des Integrals $\int_t^{t+h} f(x)\,dx$ ist. Dies ist die Aussage des **Mittelwertsatzes der Integralrechnung**. Der Mittelwert $f(\xi) \in [f(t), f(t+h)]$ ist hier der Funktionswert von f an einer (nicht näher spezifizierten) Stelle $\xi \in [t, t+h]$.

einleuchtend, dass der Wert von f in ξ, $f(\xi)$, für $h \to 0$ gegen den Wert von f in t, $f(t)$, geht: $\lim_{h\to 0} f(\xi) = f(t)$. Das stimmt hier auch. Nur für gar zu verrückte Funktionen könnte der letztgenannte Grenzwert eventuell nicht existieren – es entsteht hier keine Schwierigkeit, denn ich habe mich bereits oben entschlossen, *stetige* Funktionen f zu betrachten. Genau genommen ist die Gültigkeit von $\lim_{h\to 0} f(\xi) = f(t)$ unter der Voraussetzung $\lim_{h\to 0} \xi = t$ sogar gerade eine Form der mathematischen Beschreibung der *Stetigkeit von Funktionen* – aber das führt heute noch zu weit. Jetzt kann ich den Differenzialquotienten der Integralfunktion von f bestimmen:

$$\lim_{h\to 0} \frac{F_a(t+h) - F_a(t)}{h} = \lim_{h\to 0} f(\xi) = f(t),$$

und ich erkenne, dass die Integralfunktion F_a ableitbar ist. *Entscheidend ist*, dass das Phänomen, das ich oben lediglich für Polynomfunktionen *beobachtet* habe, für die große Klasse der stetigen Funktionen f ebenfalls gültig ist: Die erste Ableitung F_a' der Integralfunktion F_a stimmt dann mit f überein: $F_a' = f$.

Ableitbarkeit der Integralfunktion stetiger Funktionen

Für jede stetige Funktion f auf $[a, b]$ ist $F_a(t) = \int_a^t f(x)\,dx$, $t \in [a, b]$ eine ableitbare Funktion, deren erste Ableitung mit f übereinstimmt: $F_a'(t) = f(t)$, $t \in [a, b]$.

Leite ich die Integralfunktion F_a einer stetigen Funktion f (nach t) ab, so erhalte ich die Funktion f selbst:

$$F_a'(t) = (\int_a^t f(x)\,dx)' = f(t).$$

Ableiten hebt also das Integrieren auf. Mehr noch: Betrachte ich eine Integralfunktion F_a, so ist F_a' die zugehörige Integrandenfunktion (falls diese stetig ist):

$$\int_a^t F_a'(x)\,dx = F_a(t).$$

Integrieren hebt also das Ableiten auf. Jetzt sehen Sie, warum Integrieren und Ableiten als *Umkehrung voneinander* aufgefasst werden können.

Das ist zwar – Sie stimmen mir sicher zu – hochinteressant, aber ich sollte nicht aus dem Blickwinkel verlieren, dass ich Ihnen eine allgemeine Maschinerie zur Integration stetiger Funktionen vorstellen wollte. Außerdem habe ich Ihnen versprochen, dass die Berechnung solcher unbestimmter Integrale oft ähnlich einfach wie für Polynome gehen sollte. Hierzu mache ich mir erst noch einmal kurz klar, dass das bestimmte Integral von stetigen Funktionen f hinsichtlich $[a, b]$ mit der Integralfunktion F_a ausgerechnet werden kann, indem ich *diese lediglich an der Stelle b auswerte*:

$$\int_a^b f(x)\,dx = F_a(b).$$

Ein Problemchen ist hier noch, dass die Integralfunktion F_a leider nicht sehr explizit beschrieben ist. $F_a(t) = \int_a^t f(x)\,dx$ ist eher formal als ein Integral mit Variablen t in der oberen Integrationsgrenze dargestellt. Man könnte sich an dieser Stelle enttäuscht abwenden – wäre da nicht unsere *Beobachtung*, dass $F_a' = f$ gilt. *Diese hilft uns jetzt entscheidend weiter*. Um von der formalen Darstellung von F_a als Integral wegzukommen, könnte ich zur Berechnung des bestimmten Integrals von f – testweise – versuchen, *irgendeine Funktion* F zu finden, deren Ableitung f ergibt: $F' = f$. Ein solches F würde ich somit gewissermaßen durch das *Umkehren des Ableitungsprozesses* finden. Meine stille Hoffnung ist, dass sich ein solches F nicht allzu sehr von F_a unterscheiden würde, und dass dieses F sich eher explizit darstellen ließe, sodass ich es bequem zur Berechnung des bestimmten Integrals von f verwenden kann. Eine ableitbare Funktion F mit dieser Eigenschaft – $F' = f$ – nennt man **Stammfunktion von** f. Ich denke, dass ich jetzt nicht weiterschreiben sollte, ohne ein paar Beispiele für Stammfunktionen aufzulisten:

Beispiel 6.34:
Die Funktion $F(x) = x^3/3$ ist sicherlich eine Stammfunktion von $f(x) = x^2$, denn $F'(x) = 1/3\,(x^3)' = (1/3)\,3x^2 = x^2 = f(x)$. Die Funktion $G(x) = x^3/3 + 15$ ist ebenfalls Stammfunktion von $f(x) = x^2$, denn **Konstanten** *verschwinden beim Ableiten* und somit: $G'(x) = (1/3\,x^3)' + (15)' = x^2 = f(x)$. Die Funktion $P_i(x) = x^{i+1}/(i+1)$ ist eine Stammfunktion von $p_i(x) = x^i$, denn $P_i'(x) = 1/(i+1)\,(x^{i+1})' = 1/(i+1)\,(i+1)\,x^i = x^i = p_i(x)$. Die Funktion

$$F(x) = 1/16\,x^4 + 1/4\,x^3 - 3/4\,x^2 - 2\,x + \pi$$

ist eine Stammfunktion von dem oben schon betrachteten kubischen Polynom

$$f(x) = 1/4\,x^3 + 3/4\,x^2 - 3/2\,x - 2,$$

denn

$$F'(x) = 1/16\,(x^4)' + 1/4\,(x^3)' - 3/4\,(x^2)' - 2\,(x)' + (\pi)' = 4/16\,x^3 + 3/4\,x^2 - 3/2\,x - 2.$$

■

Das sind alles Funktionen, für die ich Ihnen oben schon verraten habe, wie man deren bestimmte Integrale berechnet. Um zu dokumentieren, dass ich es wirklich ernst meine, muss ich wohl noch ein paar Beispiele von Stammfunktionen nicht polynomialer Funktionen diskutieren:

Beispiel 6.35:
Die Funktion $F(x) = \exp(x) = e^x$ ist eine Stammfunktion von $f(x) = e^x$, denn $F'(x) = (e^x)' = e^x = f(x)$. Der natürliche Logarithmus $\ln(x)$ ist – wie ich Ihnen im letzten Abschnitt gezeigt habe – eine Stammfunktion von $f(x) = 1/x$, denn es gilt $\ln'(x) = 1/x$ für alle $x > 0$. Wenn Sie zu Beispiel 6.15 zurückblättern, erkennen Sie, dass $F(x) = (x^2 + x - 1)/e^x$ eine Stammfunktion von $f(x) = (-x^2 + x + 2)/e^x$ ist – dort habe ich F „k" genannt und $F' = f$ gezeigt. Die Funktion $F(x) = -\cos(x) + 4$ ist eine Stammfunktion von $f(x) = \sin(x)$, denn im letzten Abschnitt habe ich Ihnen im Rahmen der Diskussion des Zusammenhangs der Konvexität und dem Vorzeichen der zweiten Ableitung $F'(x) = \sin(x) = f(x)$ gezeigt. Ein Beispiel der delikaten Art ist die

Funktion $g(x) = (1-x^2)/(x^2+1)^2$, für die ich die Stammfunktion $G(x) = x/(x^2+1)$ kenne. Der Graph von G ist in Abbildung 6.12 dargestellt. In Beispiel 6.26 habe ich dieses G mit f bezeichnet und bereits dort habe ich Ihnen $G' = g$ gezeigt. ∎

Sie sehen, dass ich es mir in diesen Beispielen recht einfach gemacht habe. Schließlich habe ich hier vor allem Funktionen f betrachtet, für die ich bereits zuvor gesehen habe, dass diese als Ableitung einer anderen Funktion F auftreten: $F' = f$.

Übungsaufgabe 6.20:
a) Zeigen Sie, dass $H(x) = -1/3\,x^3 + x^2 - 12$ eine Stammfunktion von $h(x) = -x^2 + 2x$ ist. b) Mit den Informationen zur Lösung von Übungsaufgabe 6.8b) bestimme man eine Stammfunktion F von $f(x) = (2x-1)\,e^{x^2-x}$. Finden Sie noch weitere, sich von F unterscheidende Stammfunktionen von f? ∎

Sie haben es eben bestimmt gemerkt: Falls F eine Stammfunktion einer vorgegebenen Funktion f ist, also $F' = f$, so kann ich eine beliebige Konstante c zu F addieren und erhalte so weitere Stammfunktionen von f: $(F+c)' = F' + (c)' = F' + 0 = F' = f$. Vielleicht etwas überraschend ist, dass sich auf diese Art sogar *alle* Stammfunktionen von f ergeben. Um das zu sehen, schaue ich mir zwei Stammfunktionen – nennen wir sie F und G – derselben Funktion f an. Die erste Ableitung von deren Differenz $F - G$ verschwindet für alle x:

$$(F-G)'(x) = F'(x) - G'(x) = f(x) - f(x) = 0$$

und somit ist $F - G$ konstant.

Stammfunktionen unterscheiden sich nur um additive Konstanten

Zwei Stammfunktionen F und G derselben Funktion f unterscheiden sich nur um eine additive Konstante c: $F(x) = G(x) + c$ für alle x.

Stammfunktionen stoßen das Tor für eine allgemeine Maschinerie zur Integration stetiger Funktionen weit auf. Ich *wähle* – mit dem Ziel der Berechnung des bestimmten Integrals $\int_a^b f(x)\,dx$ – *irgendeine* Stammfunktion F von f. Die etwas unhandliche Darstellung der Integralfunktion $F_a(t) = \int_a^t f(x)\,dx$ sollte mich nicht vergessen lassen, dass F_a ebenfalls eine Stammfunktion der stetigen Funktion f ist, denn es gilt schließlich: $F_a' = f$. F_a und F unterscheiden sich folglich nur um eine Konstante c: $F_a(t) - F(t) = c$ für alle t. Ich habe mich entschieden, die Variable wiederum t zu nennen – und nicht etwa x –, weil ich das für F_a schon die ganze Zeit so gemacht habe. In diesem Szenario kann ich die Konstante c sogar leicht berechnen, indem ich $t = a$ einsetze: $c = F_a(a) - F(a) = \int_a^a f(x)\,dx - F(a) = 0 - F(a) = -F(a)$. Ich schreibe jetzt der Übersicht halber F auf die andere Seite: $F_a(t) = F(t) - F(a)$ für alle t. An der speziellen Stelle $t = b$ ergibt sich somit $F_a(b) = F(b) - F(a)$. Das bestimmte Integral $\int_a^b f(x)\,dx = F_a(b)$ kann ich somit mithilfe *der von mir gewählten Stammfunktion F* einfach ausrechnen, indem ich *lediglich* den Funktionswert von F in b und a bestimme und die Differenz bilde:

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Die grundlegende Entdeckung war es somit, dass ich für *jede* stetige Funktion f mithilfe einer ihrer Stammfunktionen F das bestimmte Integral im Wesentlichen genauso wie in meinen drei einfachen Eingangsbeispielen (Abbildungen 6.16 und 6.17 (rechts)) berechnen kann! Damit Sie diese Aussage nicht mehr vergessen werden, formuliere ich Ihnen das Ergebnis im folgenden **Hauptsatz der Differenzial- und Integralrechnung**, der den *zentralen* Zusammenhang zwischen Ableiten ($F' = f$) und Integrieren ($\int f = F$) beschreibt:

Hauptsatz der Differenzial- und Integralrechnung

Ist f eine auf $[a, b]$ stetige Funktion und F eine beliebige Stammfunktion von f, so lässt sich das bestimmte Integral von f hinsichtlich $[a, b]$ wie folgt berechnen: $\int_a^b f(x)\,dx = F(b) - F(a)$.

Zusammenfassend ist die nunmehr entwickelte *allgemeine Maschinerie zur Integration stetiger Funktionen* wie folgt:

- Finde *irgendeine* Stammfunktion F von f.
- Berechne $F(b)$ und $F(a)$ durch „Einsetzen" der Stellen b und a in F.
- Die Differenz $F(b) - F(a)$ ist das bestimmte Integral von f hinsichtlich $[a, b]$.

Für die letztgenannte Differenz verwendet man oft zwei abkürzende *Schreibweisen*:

$$\Big[F(x)\Big]_a^b = F(x)\Big|_a^b = F(b) - F(a).$$

Jetzt könnte ich schon bestimmte Integrale für viele stetige Funktionen f explizit ausrechnen.

Beispiel 6.36:
So hatte ich Ihnen beispielsweise oben gezeigt, dass $F(x) = x^3/3$ und $G(x) = x^3/3+15$ Stammfunktion von $f(x) = x^2$ sind. Somit berechne ich nun

$$\int_1^{10} x^2\,dx = F(10) - F(1) = \tfrac{10^3}{3} - \tfrac{1^3}{3} = \tfrac{999}{3} = 333$$

und

$$\int_{-1}^{1} x^2\,dx = \big[F(x)\big]_{-1}^{1} = \tfrac{1^3}{3} - \tfrac{(-1)^3}{3} = \tfrac{1}{3} + \tfrac{1}{3} = \tfrac{2}{3},$$

aber auch

$$\int_1^{10} x^2\,dx = \frac{x^3}{3} + 15\Big|_1^{10} = (\tfrac{10^3}{3} + 15) - (\tfrac{1^3}{3} + 15) = \tfrac{999}{3} = 333$$

und

$$\int_{-1}^{1} x^2\,dx = G(1) - G(-1) = (\tfrac{1^3}{3} + 15) - (\tfrac{(-1)^3}{3} + 15) = \tfrac{1}{3} + \tfrac{1}{3} = \tfrac{2}{3}. \qquad \blacksquare$$

Sie sehen, es ist *vollkommen unerheblich, welche* Stammfunktion ich zur Berechnung verwende, denn die Konstante – im Beispiel ist dies die Zahl 15 – tritt mit positivem und negativem Vorzeichen auf und fällt deshalb bei der Berechnung heraus. Das ist immer so.

Beispiel 6.37:
$F(x) = 1/16\,x^4 + 1/4\,x^3 - 3/4\,x^2 - 2\,x + \pi$ ist eine Stammfunktion von $f(x) = 1/4\,x^3 + 3/4\,x^2 - 3/2\,x - 2$ und ich berechne das obige bestimmte Integral $\int_{-4}^{-1} f(x)\,dx = \big[F(x)\big]_{-4}^{-1} = F(-1) - F(-4)$ jetzt als

$$\tfrac{1}{16}(-1)^4 + \tfrac{1}{4}(-1)^3 - \tfrac{3}{4}(-1)^2 - 2(-1) - (\tfrac{1}{16}(-4)^4 + \tfrac{1}{4}(-4)^3 - \tfrac{3}{4}(-4)^2 - 2(-4)) = \tfrac{81}{16}.$$

Da $(e^x)' = e^x$ gilt, berechne ich

$$\int_{1}^{3} e^x\,dx = \big[e^x\big]_1^3 = e^3 - e^1 = e\,(e^2 - 1)$$

oder aber auch

$$\int_{-1}^{1} e^x\,dx = \big[e^x\big]_{-1}^1 = e^1 - e^{-1} = e - \tfrac{1}{e}.$$

Wegen $\ln'(x) = 1/x$ für alle $x > 0$, finde ich

$$\int_1^2 \frac{1}{x}\,dx = \ln(x)\Big|_1^2 = \ln(2).$$

Da $(-\cos(x))' = \sin(x)$, berechne ich

$$\int_0^{\pi} \sin(x)\,dx = \big[-\cos(x)\big]_0^{\pi} = (-\cos(\pi)) - (-\cos(0)) = 2$$

oder aber auch

$$\int_{-\pi}^{\pi} \sin(x)\,dx = \big[-\cos(x)\big]_{-\pi}^{\pi} = -\cos(\pi) - (-\cos(-\pi)) = 1 - 1 = 0. \qquad \blacksquare$$

Übungsaufgabe 6.21:
a) Berechnen Sie $\int_0^2 -x^2 + 2x\,dx$.
b) Berechnen Sie $\int_1^2 (2x-1)\,e^{x^2-x}$.
c) Berechnen Sie $\int_1^2 (1-x^2)/(x^2+1)^2\,dx$.
d) Bestimmen Sie $\int_0^{\pi} (\cos(x) - \sin(x))\,e^x\,dx$, indem Sie Ihre in Übungsaufgabe 6.7b) gewonnenen Kenntnisse verwenden.
e) Bestimmen Sie für $x > 0$ eine Stammfunktion F von $f(x) = 1/x^2$ und ermitteln Sie dann $\int_1^2 1/x^2\,dx$. Hinweis: Lesen Sie noch einmal unterhalb von Übungsaufgabe 6.7 nach. $\blacksquare$

Sie sehen also: Sobald ich eine Stammfunktion F von f *explizit* kenne, ist die Berechnung des bestimmten Integrals von f hinsichtlich $[a,b]$ relativ leicht – ich muss lediglich $\big[F(x)\big]_a^b = F(b) - F(a)$ ausrechnen. Allerdings bin ich nicht immer in dieser glücklichen Situation. Leider ist das für uns derzeit zum Beispiel mit den Integralen

$$\int_a^b x\,e^x\,dx,\ \int_1^7 \ln(x)\,dx,\ \int^t \sin^2(x)\,dx,\ \int_a^b 6x\,(3x^2+1)^{71}\,dx,\ \int_1^2 \frac{\ln(x)}{x}\,dx$$

der Fall. Natürlich kann ich hier formal jeweils eine Stammfunktion hinschreiben, denn die Integralfunktionen – zum Beispiel $\int_1^t \ln(x)\,dx$ – sind solche Funktionen. Nur hilft das herzlich wenig, denn ich kenne dann die Funktion F, in die ich gerne die Integrationsgrenzen einsetzen würde, leider noch *nicht explizit.* Zur Information sage ich Ihnen, dass es tatsächlich stetige Integranden f gibt, für die es nicht möglich ist, einen **geschlossenen Ausdruck** durch elementare Funktionen für eine ihrer Stammfunktionen anzugeben. So kann ich beispielsweise keine der Stammfunktionen $\int_a^t e^{-x^2}\,dx + c$ von $f(x) = e^{-x^2}$ explizit als geschlossenen Ausdruck hinschreiben. Dasselbe passiert, wenn ich die Länge eines **Ellipsenbogens** ausrechnen möchte (siehe Ende Kapitel 4) – dies führt auf sogenannte **elliptische Integrale**, die nicht durch elementare Funktionen ausdrückbar sind, was sehr ärgerlich ist. Aber so ist das eben.

Glücklicherweise gibt es aber neben den oben bereits formulierten Regeln (Integration des Faktorprodukts und der Summe von Funktionen) auch die folgenden beiden Integrationsregel, mit denen ich manchmal geschlossene Ausdrücke für Stammfunktionen finden kann. Grob gesprochen erhalte ich diese durch Integration der Ableitungsregeln, die im ersten Abschnitt dieses Kapitels **Produkt- und Kettenregel** genannt wurden.

Betrachte ich zwei ableitbare Funktionen f und g, so besagt die Produktregel, dass $(f \cdot g)' = f' \cdot g + g' \cdot f$ gilt. Die Idee ist es nun, die Funktionen beider Seiten zu integrieren. Für $(f \cdot g)'$ kenne ich eine explizite Stammfunktion $f \cdot g$, während ich vielleicht für $f' \cdot g$ eine solche finden möchte. Da $f' \cdot g = (f \cdot g)' - g' \cdot f$ gilt, genügt es, hierzu eine explizite Stammfunktion für $g' \cdot f$ zu finden,

$$\int^t f' \cdot g = \int^t (f \cdot g)' - \int^t g' \cdot f = (f \cdot g)(t) - \int^t g' \cdot f,$$

und die Hoffnung ist, dass dies zumindest manchmal relativ leicht gelingt.

Produktintegration – Partielle Integration

Sind f und g auf $[a,b]$ ableitbare Funktionen, so gilt

$$\int_a^b f'(x)\,g(x)\,dx = f(x)\,g(x)\Big|_a^b - \int_a^b f(x)\,g'(x)\,dx.$$

Beispiel 6.38:
Ein typisches Beispiel ist die Suche nach einer Stammfunktion von $h(x) = x\,e^x$. Ich setze $f(x) = e^x$ und $g(x) = x$ und erkenne $f'(x) = e^x$ und $g'(x) = 1$. Somit gilt hier $h(x) = f'(x)g(x)$. Wende ich nun partielle Integration an, so erhalte ich

$$\int^t \overbrace{e^x\,x}^{=f'(x)\,g(x)} \;dx = \overbrace{e^t\,t}^{=f(t)\,g(t)} - \int^t \overbrace{e^x\,1}^{=f(x)\,g'(x)} \;dx.$$

Das unbestimmte Integral auf der rechten Seite lässt sich explizit hinschreiben $\int^t e^x\,dx = e^t$, und somit folgt, dass $F(t) = t\,e^t - e^t = (t-1)\,e^t$ eine Stammfunktion von f ist. Ich kann dies jetzt checken, indem ich die Ableitung von F ausrechne. Die Produktregel ergibt tatsächlich:

$$F'(t) = (t-1)'\,e^t + (t-1)(e^t)' = 1\,e^t + (t-1)\,e^t = t\,e^t = h(t).$$

Möchte ich nun ein bestimmtes Integral von h ausrechnen, so geht dies wie immer:

$$\int_a^b x\,e^x\,dx = (x-1)\,e^x\Big|_a^b = (b-1)\,e^b - (a-1)\,e^a.$$

Beispielsweise berechne ich

$$\int_1^2 x\,e^x\,dx = e^2.$$

■

Beispiel 6.39:
Ein anderes Beispiel ist die Berechnung des bestimmten Integrals $\int_1^7 \ln(x)\,dx$. Hier erscheint es zunächst schwierig, etwas Erfolgversprechendes zu tun – den **natürlichen Logarithmus** ln kann ich anscheinend nicht weiter vereinfachen oder trickreich zerlegen. An dieser Stelle beginnen Sie zu verstehen, weshalb manche Leute *Ableiten als Technik, Integrieren aber als Kunst* bezeichnen. Nach einigem Herumprobieren kommt man vielleicht auf die Idee, $g(x) = \ln(x)$ *künstlich* als Produkt zweier Funktionen aufzufassen $g(x) = 1 \cdot \ln(x)$. Wende ich nun partielle Integration an, so erhalte ich für $t > 0$:

$$\int^t \overbrace{1 \cdot \ln(x)}^{=(x)'\,g(x)} \,dx = \overbrace{t\,\ln(t)}^{=t\,g(t)} - \int^t \overbrace{x\,\frac{1}{x}}^{=x\,\ln'(x)} \;dx.$$

Das unbestimmte Integral auf der rechten Seite lässt sich explizit hinschreiben $\int^t x\,1/x\,dx = \int^t 1\,dx = t$, und somit folgt, dass $F(t) = t\,\ln(t) - t = t(\ln(t) - 1)$ eine Stammfunktion von $\ln(t)$ ist. Anwenden der Produktregel bestätigt das:

$$F'(t) = (t\,\ln(t))' - (t)' = (t)'\ln(t) + t(\ln(t))' - 1 = 1\,\ln(t) + t\,1/t - 1 = \ln(t).$$

Mit der Stammfunktion F kann ich nun das oben erwähnte bestimmte Integral berechnen:

$$\int_1^7 \ln(x)\,dx = \big[x\,\ln(x) - x\big]_1^7 = 7 \cdot \ln(7) - 7 - (1 \cdot \ln(1) - 1) = 7 \cdot \ln(7) - 6.$$

Die entsprechende Fläche habe ich in Abbildung 6.9 (rechts) grau dargestellt. ■

Beispiel 6.40:
Ein ebenfalls nettes Beispiel ist die Suche nach einer Stammfunktion von $f(x) = \sin^2(x) = \sin(x)\,\sin(x)$. Mit partieller Integration erhalte ich zunächst

$$\int^t \overbrace{\sin(x)\,\sin(x)}^{=(-\cos(x))'\,\sin(x)}\,dx = -\cos(t)\,\sin(t) - \int^t \overbrace{-\cos^2(x)}^{=-\cos(x)\,(\sin(x))'}\,dx.$$

Das gefällt mir auf den ersten Blick gar nicht, denn rechts taucht der Integrand $-\cos^2(x)$ auf, der genauso kompliziert ist wie der Integrand, von dem ich ausgegangen bin: $\sin^2(x)$. In der Tat landet man an dieser Stelle manchmal in der Sackgasse und muss sich etwas anderes überlegen. Hier ist dies aber zum Glück nicht so. Ich erinnere mich an $\cos^2(x) = 1 - \sin^2(x)$ und berechne so

$$\int^t \sin^2(x)\,dx = -\cos(t)\,\sin(t) + \int^t 1\,dx - \int^t \sin^2(x)\,dx.$$

Damit ist $F(x) = 1/2\,(-\cos(x)\,\sin(x) + x)$ eine Stammfunktion von $f(x) = \sin^2(x)$.
■

Übungsaufgabe 6.22:
Bestimmen Sie mit partieller Integration alle Stammfunktionen von $f(x) = \sin(x)\,x$ und $g(x) = x^2\,e^x$. ■

Die zweite Integrationsregel, die ebenfalls manchmal für die Bestimmung von Stammfunktionen verwendet werden kann, heißt **Substitutionsregel**. Hier wird – zumindest in der Standardanwendung, für die ich Ihnen zunächst Beispiele zeige – nach Möglichkeit ein komplizierter Integrand durch einen einfacheren Integranden *ersetzt*. Ich gehe hier davon aus, dass ich für eine einfache Funktion f eine Stammfunktion F kenne und der komplizierte Integrand von der Form $g' \cdot (f \circ g)$ ist – f ist hier also mit g **verkettet** und dies wird mit g' multipliziert. Das sieht zwar sehr speziell aus – kommt aber in der Tat öfter vor, wie man gemeinhin denkt. Nach der **Kettenregel** aus dem ersten Abschnitt gilt $(F \circ g)'(x) = g'(x)\,F'(g(x))$ – also folgt wegen $F' = f$: $(F \circ g)'(x) = g'(x)\,f(g(x))$. Damit ist $(F \circ g)$ eine Stammfunktion von $g' \cdot (f \circ g)$ und ich berechne das Integral hinsichtlich des komplizierteren Integranden $g' \cdot (f \circ g)$ wie folgt:

$$\int^t g' \cdot (f \circ g) = (F \circ g)(t) = F(g(t)) = \int^{g(t)} f.$$

Ich sehe, dass sich dieses Integral durch Integration des einfacheren Integranden f ergibt, indem ich dort lediglich die Integrationsgrenze(n) anpasse.

Substitutionsregel

Ist f stetig und g ableitbar auf $[a, b]$, so gilt

$$\int_a^b g'(x)\,f(g(x))\,dx = \int_{g(a)}^{g(b)} f(\bar{x})\,d\bar{x}.$$

Ein Beispiel sagt vielleicht mehr als tausend Worte:

Beispiel 6.41:
Ich betrachte die Polynomfunktion $p(x) = 6x\,(3x^2+1)^{71}$ (vom Grad 143). Ich erkenne, dass hier für $g(x) = 3x^2 + 1$ gerade $g'(x) = 6x$ als multiplikativer Term in p auftritt. Weiter sehe ich, dass ich mithilfe der Potenzfunktion $f(\bar{x}) = \bar{x}^{71}$ das Polynom p als $g'\,(f \circ g)$ hinschreiben kann. Wenn ich nämlich die Variable $\bar{x}$ durch $g(x)$ ersetze, so folgt $\bar{x} = 3x^2 + 1$. Weiter entsinne ich mich, dass ich für die Potenzfunktion f keine Probleme habe, eine Stammfunktion zu finden: $F(\bar{x}) = 1/72\,\bar{x}^{72}$. Also berechne ich

$$\int^t 6x\,(3x^2+1)^{71}\,dx = \int^{g(t)} \bar{x}^{71} d\bar{x} = 1/72\,\bar{x}^{72}\Big|^{3t^2+1} = 1/72\,(3t^2+1)^{72}.$$

In der Tat ist $P(x) = 1/72\,(3x^2+1)^{72}$ eine Stammfunktion von p, was Sie durch Ableiten von P mit der Kettenregel verifizieren können. ■

Beispiel 6.42:
Da ich vermeiden möchte, Sie mit einem der weiter oben genannten Integrale alleine zu lassen, wende ich mich nun dem bestimmten Integral $\int_1^2 \ln(x)/x\,dx$ zu. Ich versuche – unsere allgemeine Maschinerie verwendend – erst einmal eine Stammfunktion von $\ln(x)/x$ zu finden. Hierzu erkenne ich zunächst, dass $g'(x) = 1/x$ die Ableitung von $g(x) = \ln(x)$ ist. Das sollte man ausnutzen. Hierzu definiere ich $f(\bar{x}) = \bar{x}$, denn der obige Integrand kann als $g' \cdot (f \circ g)$ geschrieben werden – ersetze ich nämlich $\bar{x} = g(x) = \ln(x)$ in der Festlegung von f und multipliziere ich gleichzeitig mit $g'(x) = 1/x$, so erhalte ich den Integranden $1/x \cdot \ln(x)$. Für f habe ich keine Probleme, eine Stammfunktion zu finden, denn mein zweites Beispiel in diesem Abschnitt zeigte ja bereits prinzipiell, dass $F(\bar{x}) = 1/2\,\bar{x}^2$ die Eigenschaft $F' = f$ besitzt. Also berechne ich jetzt eine Stammfunktion des Integranden $\ln(x)/x$ für $t > 0$:

$$\int^t \frac{\ln(x)}{x}\,dx = \int^{\ln(t)} \bar{x}\,d\bar{x} = \left[\frac{\bar{x}^2}{2}\right]^{\ln(t)} = \frac{\ln^2(t)}{2}.$$

Schließlich berechne ich

$$\int_1^2 \ln(x)/x\,dx = [\ln^2(x)/2]_1^2 = \ln^2(2)/2 - \ln^2(1)/2 = \ln^2(2)/2.$$ ■

Beispiel 6.43:
Am Ende von Kapitel 4 fiel es mir schwer, mit den dortigen Mitteln den **Flächeninhalt einer Ellipse** um den Mittelpunkt $(0|0)$ mit Hauptachsenradius r und Nebenachsenradius b zu berechnen. Dies möchte ich nun mithilfe der Integralrechnung nachholen. Eine solche Ellipse besitzt, wie dort gesehen, die Darstellung $x^2/r^2 + y^2/b^2 = 1$. Löse ich dies nach y auf, so erhalte ich für die Punkte auf der Ellipse mit nicht negativen y-Werten $y = b\sqrt{1 - x^2/r^2}$. Um den Flächeninhalt F der Ellipse zu berechnen muss ich also das Integral

$$I = \int_{-r}^{r} b\sqrt{1 - \frac{x^2}{r^2}}\,dx = 2b\int_0^r \sqrt{1 - \frac{x^2}{r^2}}\,dx$$

ausrechnen, welches die Hälfte von F darstellt. Betrachte ich die Parameterdarstellung der Ellipse aus Kapitel 4, das heißt insbesondere $x = r\sin(t)$, so kann ich x im Integral dadurch ersetzen, indem ich den so entstehenden Integranden mit der Ableitung $x' = r\cos(t)$ multipliziere und die Integrationsgrenzen anpasse ($0 = r\sin(0)$ und $r = r\sin(\pi/2)$):

$$\begin{aligned} I &= 2b\int_0^{\frac{\pi}{2}} r\cos(t)\sqrt{1-\frac{(r\sin(t))^2}{r^2}}\,dt = 2rb\int_0^{\frac{\pi}{2}}\cos(t)\sqrt{1-\sin^2(t)}\,dt \\ &= 2rb\int_0^{\frac{\pi}{2}}(1-\sin^2(t))\,dt = 2rb\left(\tfrac{\pi}{2}-\int_0^{\frac{\pi}{2}}\sin^2(t)\,dt\right). \end{aligned}$$

Hier habe ich im vorletzten Schritt verwendet, dass $\cos(t) = \sqrt{1-\sin^2(t)}$ und $\sqrt{z}\cdot\sqrt{z} = z$ gelten. Jetzt kann ich meine Kenntnisse aus Beispiel 6.40 anwenden und berechne

$$\int_0^{\frac{\pi}{2}}\sin^2(t)\,dt = \tfrac{1}{2}(-\cos(\tfrac{\pi}{2})\sin(\tfrac{\pi}{2})+\tfrac{\pi}{2}) - \tfrac{1}{2}(-\cos(0)\sin(0)+0) = \tfrac{\pi}{4},$$

also ingesamt $I = 2rb\pi/4 = rb\pi/2$. Als Flächeninhalt der Ellipse F mit Hauptachsenradius r und Nebenachsenradius b erhalte ich somit

$$F = 2\cdot I = \pi rb.$$ ■

Dies war auch ein ganz schönes Beispiel um zu zeigen, dass ich die Regel im letzten Kasten auch von rechts nach links lesen kann und dann ebenfalls anwenden darf. Ich muss dann – wie ich Ihnen an Beispiel 6.43 gezeigt habe – die Ableitung des zu ersetzenden Terms an den Integranden heranmultiplizieren und die Integrationsgrenzen anpassen. Es erfordert schon etwas Erfahrung zu erkennen, welchen Term man in diesem Sinne ersetzen sollte, sodass der resultierende Integrand einfach auflösbar ist. Man nennt diese Vorgehensweise – im Unterschied zu der Methode in den anderen obigen Beispielen – manchmal auch die **2. Substitutionsregel**.

Ich denke es ist nun wirklich an der Zeit, dass ich mich nach diesen mühevollen Beschreibungen ein wenig entspanne und Ihnen ein letztes Mal in diesem Kapitel das Zepter in die Hand gebe:

Übungsaufgabe 6.23:
Bestimmen Sie mit der obigen Substitutionsregel alle Stammfunktionen von $f(x) = 3x^2\,(x^3+4)^{88}$ und $g(x) = (2x-1)\,e^{x^2-x}$. ■

Übungsaufgabe 6.24:
Bestimmen Sie den Inhalt der in Abbildung 6.12 grau dargestellten Fläche. ■

7 Grundlagen der Wahrscheinlichkeitsrechnung

Die Wahrscheinlichkeitstheorie oder, etwas pragmatischer bezeichnet, Wahrscheinlichkeitsrechnung ist für mich schon immer der geheimnisvollste, geradezu mystischste Teil der Mathematik. Das soll Sie jetzt keinesfalls er- oder gar abschrecken, im Gegenteil will ich Ihre Neugierde wecken. Die Bezeichnung „geheimnisvoll“ bezieht sich nämlich in keiner Weise auf die harten mathematischen Fakten, mit denen man es in der Wahrscheinlichkeitsrechnung zu tun hat, sondern vielmehr auf die verblüffende Tatsache, dass die solchermaßen errechneten Ergebnisse mit der realen Welt wirklich etwas zu tun haben. Oder sind Sie nicht verblüfft, wenn Sie am Wahlabend, nur wenige Minuten nach Schließung der Wahllokale und nach Auszählung von nur ganz wenigen Wahlkreisen bereits auf ein Prozent genau das Gesamtergebnis erfahren können? (Nur Ketzer vermuten, dass die restlichen Stimmen meist gar nicht mehr ausgezählt werden und der Wahlleiter das Endergebnis nach ein paar Stunden auf Basis der ersten Hochrechnung festlegt.)

Ein anderes Beispiel für das Mysterienhafte der Wahrscheinlichkeitsrechnung: Beim Wurf einer Münze, die auf einer Seite „Kopf“ und auf der anderen Seite „Wappen“ trägt, geht man mit Fug und Recht davon aus, dass das Erscheinen beider Seiten gleich wahrscheinlich ist; da man in der mathematischen Wahrscheinlichkeitsrechnung nicht wie in der Umgangssprache mit Prozentzahlen, sondern mit Zahlen zwischen 0 und 1 rechnet, haben also die beiden Würfe „Kopf“ und „Wappen“ jeweils die Wahrscheinlichkeit 1/2. Stillschweigend nimmt man hierbei an, dass die Münze niemals auf dem Rand stehen bleibt; ich verkneife mir zu sagen, dass dies auch ziemlich „unwahrscheinlich“ wäre.

Nehmen Sie nun einmal an, Sie hätten die Münze bereits 999-mal geworfen, und *jedes Mal* wäre „Kopf“ erschienen; wie wahrscheinlich ist es, dass beim tausendsten Wurf noch einmal „Kopf“ erscheint? Intuitiv will man sicherlich antworten: „Ziemlich unwahrscheinlich!“, aber die Wahrscheinlichkeitsrechnung sagt, dass wiederum mit Wahrscheinlichkeit 1/2 „Kopf“ erscheinen wird, denn jeder Wurf ist ein von den anderen Würfen *unabhängiges Ereignis*, die Münze hat sozusagen „kein Gedächtnis“. Wenn Sie das (verständlicherweise) nicht glauben wollen, dann stellen Sie sich vor, dass just in dem Moment, in dem Sie zum tausendsten Wurf ansetzen, Ihr Freund (sofern Sie nach einem Mathematikstudium noch Freunde haben) den Raum betritt und Sie ihn nach der Wahrscheinlichkeit dieses Wurfs fragen. Da er ebenso wie die

G. Walz et al., *Brückenkurs Mathematik*

Münze selbst von den ersten 999 Würfen nichts weiß, wird er sicherlich „1/2“ antworten, und er hätte recht!

Um Sie nun zum Ende dieser einleitenden Bemerkungen noch vollständig zu verwirren, darf ich Ihnen das Folgende sagen: Wenn Sie *vor* dem ersten Wurf nach der Wahrscheinlichkeit dafür gefragt werden, 1 000-mal hintereinander „Kopf“ zu werfen, so lautet die korrekte Antwort: „$2^{-1\,000}$“; das ist eine Zahl, die nach dem Komma erst mal etwa 300 Nullen hat, bevor eine gültige Ziffer erscheint.

Wenn Sie jetzt drauf und dran sind, dieses Buch in die Ecke zu feuern oder zumindest zuzuklappen, dann warten Sie noch einen Moment! Ich glaube, ich kann Ihnen auf den folgenden Seiten erklären, warum das mit den Wahrscheinlichkeiten so ist wie es nun einmal ist und Sie in die Lage versetzen, mit einfachen bis mittelschweren wahrscheinlichkeitsrechnerischen Problemen fertig zu werden; und viel mehr wird man auch im Studium nicht von Ihnen verlangen.

7.1 Kombinatorik

Eine der übersichtlichsten Situationen in der Wahrscheinlichkeitsrechnung liegt vor, wenn es bei einem Versuch (wobei in dieser Sprachregelung auch ein Münzwurf ein Versuch ist) nur endlich viele, sagen wir n, Versuchsausgänge möglich und diese auch noch alle gleich wahrscheinlich sind. Da nämlich die Gesamtwahrscheinlichkeit, also die Wahrscheinlichkeit dafür, dass *überhaupt irgendein* Ausgang stattfindet (was klar ist), per Definition gleich 1 ist, die Summe aller n Einzelwahrscheinlichkeiten also 1 ergeben muss, muss jede dieser Einzelwahrscheinlichkeiten gleich $1/n$ sein.

Versuch mit endlich vielen gleich wahrscheinlichen Ausgängen

Hat ein Versuch n verschiedene Ausgänge, die alle gleich wahrscheinlich sind, so tritt jeder Ausgang mit der Wahrscheinlichkeit $1/n$ ein.

Ein erstes Beispiel hatte ich oben schon erwähnt: Beim Münzwurf gibt es nur zwei mögliche Ausgänge, folglich hat jeder dieser beiden die Wahrscheinlichkeit 1/2. Ein anderes Standardbeispiel der elementaren Wahrscheinlichkeitsrechnung ist das Würfeln mit einem Standardwürfel; da es hier sechs verschiedene Ausgänge gibt (nämlich die sechs Augenzahlen), tritt jede einzelne dieser sechs möglichen Augenzahlen mit der Wahrscheinlichkeit 1/6 auf.

Jetzt kommt eine wichtige Überlegung: Auch beim zufälligen Herausgreifen von sechs Kugeln aus einer Trommel mit 49 Kugeln, also beim Zahlenlotto oder Lotto am Samstag, ist jede einzelne Auswahl von sechs Kugeln gleich wahrscheinlich, Sie können also die Wahrscheinlichkeit eines bestimmten Ergebnisses ganz leicht als Kehrwert der Gesamtzahl aller Möglichkeiten berechnen, wenn, ja, wenn Sie wissen, wie viele Möglichkeiten es insgesamt gibt, 6 aus 49 Kugeln auszuwählen.

Genau darum geht es in der Kombinatorik, also der Lehre von der Anzahl der Kombinationen: Man will herausfinden, wie viele Möglichkeiten es gibt, aus einer gewissen Anzahl von Dingen eine kleinere Anzahl nach gewissen Regeln auszuwählen; der Kehrwert dieser Möglichkeitenzahl gibt dann gerade die Wahrscheinlichkeit dafür an, eine bestimmte Möglichkeit anzutreffen.

Ach du liebe Güte! Ich habe gerade den letzten Abschnitt nochmal gelesen und habe ihn selbst kaum verstanden, wie mag es da erst Ihnen gehen? So geht das nicht weiter, ich muss jetzt mal konkret werden; also: Das gerade geschilderte Denkmodell mit der Lottotrommel ist tatsächlich ein sehr klassisches, allerdings hat man in früheren Zeiten die Kugeln nicht in Glastrommeln, sondern in Urnen gelegt, weshalb man auch von Urnenmodellen spricht. Außerdem muss man noch unterscheiden, nach welchen Regeln man die Kugeln entnehmen darf. All das wird im folgenden, demnach etwas ausführlicheren Textkasten geschildert:

Urnenmodelle

Unter einem **Urnenmodell** versteht man folgendes Gedankenexperiment: In einem Gefäß (der Urne) seien n unterscheidbare Kugeln enthalten, von denen nach und nach k Stück zufällig (also ohne hinzusehen) ausgewählt werden. Man fragt nun nach der Anzahl der Möglichkeiten, die diese Auswahl hervorbringen kann, wobei man noch folgende Auswahlregeln unterscheidet: Wird jede Kugel nach ihrer Entnahme wieder in die Urne zurückgelegt, so spricht man von **Auswahl mit Zurücklegen**, wird sie nicht zurückgelegt, so nennt man das **Auswahl ohne Zurücklegen**.

In jedem der beiden Fälle wird außerdem noch unterschieden, ob es auf die Reihenfolge der Entnahme ankommt (**Auswahl mit Beachtung der Reihenfolge**) oder eben nicht (**Auswahl ohne Beachtung der Reihenfolge**).

Insgesamt sind also die folgenden vier Urnenmodelle zu unterscheiden:

- **Auswahl mit Zurücklegen mit Beachtung der Reihenfolge**
- **Auswahl mit Zurücklegen ohne Beachtung der Reihenfolge**
- **Auswahl ohne Zurücklegen mit Beachtung der Reihenfolge**
- **Auswahl ohne Zurücklegen ohne Beachtung der Reihenfolge**

In jedem der vier Fälle kann man genau sagen, wie viele Möglichkeiten es insgesamt gibt, und somit die Wahrscheinlichkeit für eine bestimmte Auswahl als Kehrwert dieser Zahl berechnen. Ich werde Ihnen im Folgenden den jeweiligen Formelausdruck, so weit es im Rahmen dieses Buches möglich und nötig ist, ableiten, auf jeden Fall aber plausibel machen. In welcher Art und Weise die Kugeln unterscheidbar gemacht werden, ist natürlich völlig egal, beispielsweise kann man sie unterschiedlich einfärben, mit Buchstaben oder mit Zahlen versehen; da ich als Mathematiker von Haus aus am besten mit Zahlen umgehen kann (was meine Frau allerdings manchmal auch bezweifelt), will ich für die folgenden Beispiele annehmen, dass die n Kugeln in der Urne durchnummeriert sind.

Ich beginne mit dem meiner Meinung nach einfachsten Fall, nämlich der *Auswahl mit Zurücklegen mit Beachtung der Reihenfolge*: Für die Auswahl der ersten Kugel habe ich offenbar genau n Möglichkeiten. Ich notiere die Nummer dieser Kugel und lege sie anschließend in die Urne zurück. Daher habe ich beim Ziehen der zweiten Kugel wiederum alle n Möglichkeiten. Und jetzt kommt's: Zu *jeder einzelnen* der n Möglichkeiten für die Wahl der ersten Kugel gibt es n Möglichkeiten für die zweite, insgesamt gibt es somit n mal n, also n^2 verschiedene Möglichkeiten für die Wahl der beiden ersten Kugeln. Und so geht das natürlich weiter, wenn ich eine dritte Kugel ziehe, habe ich erneut n Möglichkeiten, insgesamt können dann also n^3 verschiedene Situationen auftreten, und da es keinen Grund gibt, nach drei Kugeln zu stoppen, haben wir bereits für den ersten Fall eine allgemeine Formel:

Auswahl mit Zurücklegen mit Beachtung der Reihenfolge

Zur Auswahl von k Kugeln aus einer Menge von n Kugeln mit Zurücklegen und mit Beachtung der Reihenfolge gibt es

$$n^k$$

Möglichkeiten. Die Wahrscheinlichkeit dafür, eine bestimmte Auswahl zu treffen, ist also

$$\frac{1}{n^k} \; .$$

Hier ist also das gute alte Potenzieren mit natürlichen Zahlen aus dem ersten Kapitel wieder gefragt. Beispielsweise berechnet man, dass es für die Auswahl von vier Kugeln aus einer Urne mit 15 Kugeln insgesamt $15^4 = 50\,625$ verschiedene Möglichkeiten gibt; nicht schlecht, wie?

Natürlich sind diese Ergebnisse nicht auf die pure Kugelrechnerei beschränkt, das ist ja nur ein Gedankenmodell. Mit ein wenig Abstraktionsvermögen kann man damit auch ganz andere Situationen lösen. Hierzu ein Beispiel:

Beispiel 7.1:
An einem Tresor befinden sich fünf Zahlenräder, die jeweils auf eine der Zahlen von 1 bis 99 eingestellt werden können. Wie viele Möglichkeiten gibt es, hieraus eine Zahlenkombination zu bilden?

Das richtige Urnenmodell ist hier (wen wundert's?) die Auswahl mit Zurücklegen (da man ja beispielsweise das zweite Rädchen auf die gleiche Zahl wie das erste einstellen kann, diese also wieder zur Verfügung steht) und mit Beachtung der Reihenfolge. Somit gibt es

$$99^5 \;=\; 9\,509\,900\,499 \; ,$$

also fast 10 Milliarden Möglichkeiten. Viel Erfolg beim Tresorknacken! ■

Übungsaufgabe 7.1:
Aus einem Kartenspiel mit 32 verschiedenen Karten wird dreimal nacheinander eine Karte gezogen und anschließend wieder in den Stapel zurückgelegt. Wie wahrscheinlich ist es, dreimal Herz-Dame zu ziehen? ■

Als Nächstes wende ich mich der *Auswahl mit Zurücklegen ohne Beachtung der Reihenfolge* zu. Auch hier wird also eine einmal gezogene Kugel wieder in die Urne zurückgelegt, jetzt aber kommt es bei der Betrachtung des Endergebnisses nicht mehr auf die Reihenfolge an, in der die einzelnen Kugeln gezogen wurden; beispielsweise würden also die Züge 3-6-2 und 6-3-2 als dasselbe Ergebnis gewertet, da sie beide je einmal die Kugeln 2, 3 und 6 hervorgebracht haben.

Wenn es Ihnen schwer fällt, sich das vorzustellen, weil Sie lieber Würfelspiele machen als Kugeln zu ziehen, dann hilft vielleicht das Folgende. Das Problem, aus einer Menge von sechs Zahlen (Kugeln) zwei Stück mit Wiederholung ohne Beachtung der Reihenfolge auszuwählen, entspricht genau dem Problem, mit zwei nicht unterscheidbaren Würfeln ein bestimmtes Wurfergebnis zu erzielen: Das Werfen eines Würfels entspricht genau der Auswahl einer Zahl zwischen 1 und 6, und da die Würfel nicht unterscheidbar sein sollen, werden auch beispielsweise die Ergebnisse, mit dem „ersten" Würfel eine Zwei und mit dem „zweiten" eine Fünf sowie dem „ersten" Würfel eine Fünf und mit dem „zweiten" eine Zwei zu würfeln, als dasselbe Ereignis gewertet.

Bevor ich ernsthaft daran gehen kann zu berechnen, wie viele Möglichkeiten es in dieser Situation gibt, muss ich leider noch eine neue Notation einführen, denn mit simplem Potenzieren ist es jetzt leider nicht mehr getan. Erinnern Sie sich noch an die Definition der Fakultät, die im ersten Kapitel gegeben wurde? Wenn nicht, ist es nicht so schlimm, ich wiederhole das hier nochmal: Ist m eine natürliche Zahl, so ist $m!$ („m Fakultät") definiert als

$$m! \ = \ 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (m-1) \cdot m \ .$$

Als wäre das nicht schon schlimm genug, braucht man jetzt noch einen neuen Begriff, für dessen Definition ich auf die Fakultät zurückgreifen muss, nämlich den Binomialkoeffizienten. Dieser kombiniert zwei natürliche Zahlen, sagen wir m und l, und ist wie folgt definiert:

Binomialkoeffizient

Sind m und l zwei natürliche Zahlen mit $l \leq m$, so ist der **Binomialkoeffizient**

$$\binom{m}{l} \ ,$$

gesprochen „m über l", definiert als

$$\binom{m}{l} \ = \ \frac{m!}{l! \cdot (m-l)!} \ . \tag{7.1}$$

Sie sollten jetzt weder sich noch mich fragen, warum man das gerade *so* definiert hat; es stellte sich eben heraus, dass genau diese Kombination zweier Zahlen in sehr vielen Formeln gebraucht wird, und daher hat man diese abkürzende Notation eingeführt.

Beispielsweise werde ich den Binomialkoeffizienten benutzen, um die Anzahl der Auswahlen im gerade betrachteten Kontext zu berechnen. Zuvor allerdings will ich Sie noch auf eine Möglichkeit hinweisen, bei der Berechnung des Binomialkoeffizienten Arbeit und auch Fehlerquellen einzusparen: Da ja $l \leq m$ vorausgesetzt ist, sind sowohl l als auch $m - l$ Zahlen zwischen 0 und m, somit steckt sowohl $l!$ als auch $(m-l)!$ als Faktor in $m!$ drin, und daher kann man aus der Definition (7.1) folgende Identitäten herleiten:

$$\binom{m}{l} = \frac{(l+1)\cdot(l+2)\cdots(m-1)\cdot m}{(m-l)!} = \frac{(m-l+1)\cdot(m-l+2)\cdots(m-1)\cdot m}{l!} .$$

Zur Berechnung von $\binom{m}{l}$ kann man jede der Darstellungen benutzen, man sollte stets die nehmen, bei der weniger Rechenarbeit zu leisten ist; dies wiederum hängt von der Größe von m und l ab. Beispielsweise ist

$$\binom{7}{3} = \frac{(7-3+1)\cdot(7-3+2)\cdot 7}{3!} = \frac{5\cdot 6\cdot 7}{2\cdot 3} = 35$$

und

$$\binom{4}{2} = \frac{3\cdot 4}{2} = 6 .$$

Vielleicht haben Sie sich die ganze Zeit über schon gewundert, dass ich zur Definition des Binomialkoeffizienten die beteiligten natürlichen Zahlen nicht wie meist mit n und k bezeichnet habe. Das liegt einfach daran, dass im Kontext Urnenmodelle, in dem wir uns ja gerade befinden – ich hoffe, Sie hatten das nicht vergessen – n und k anderweitig belegt sind, nämlich als Anzahl von Kugeln. Der Zusammenhang wird jetzt hergestellt:

Auswahl mit Zurücklegen ohne Beachtung der Reihenfolge

Zur Auswahl von k Kugeln aus einer Menge von n Kugeln mit Zurücklegen, aber ohne Beachtung der Reihenfolge gibt es

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

Möglichkeiten. Die Wahrscheinlichkeit dafür, eine bestimmte Auswahl zu treffen, ist also

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

Beispiel 7.2:

Ich hatte in der Einleitung versprochen zu berechnen, wie viele verschiedene Würfe es mit zwei nicht unterscheidbaren Würfeln gibt. Wir hatten uns bereits darauf geeinigt, dass das richtige Urnenmodell hierfür die Auswahl mit Zurücklegen und ohne

Beachtung der Reihenfolge ist. Folglich gibt es

$$\binom{7}{2} = \frac{6 \cdot 7}{2} = 21$$

verschiedene Würfe. ■

Übungsaufgabe 7.2:
In einer Trommel liegen sieben Kugeln, die mit den Buchstaben A, B, E, H, M, N, T beschriftet sind. Man zieht nun nacheinander fünf Kugeln, notiert sich den Buchstaben und legt die gezogene Kugel wieder in die Trommel zurück. Wie groß ist die Wahrscheinlichkeit dafür, aus den notierten Buchstaben das Wort „MATHE“ bilden zu können? ■

Falls Sie übrigens verzweifelt auf weitere Übungsaufgaben warten, keine Sorge, die kommen noch. Ich stelle sie allerdings gemeinerweise erst nach der Vorstellung der beiden noch verbliebenen Urnenmodelle, und zwar vermischt, damit es schwerer wird!

Jetzt behandle ich zunächst die Auswahl von k aus n Kugeln *ohne* Zurücklegen; das bedeutet auch, dass ich jetzt stets $k \leq n$ voraussetzen muss, denn da ich eine bereits gezogene Kugel nicht mehr zurücklege, kann ich eben insgesamt nicht mehr als n Kugeln ziehen.

Schauen wir uns die *Auswahl ohne Zurücklegen mit Beachtung der Reihenfolge* an: Für die Auswahl der ersten Kugel habe ich offenbar genau n Möglichkeiten. Da diese nun aber nicht zurückgelegt wird, gibt es für die zweite Kugel noch $n-1$ Möglichkeiten, für die dritte noch $n-2$, usw. Wenn Sie das nun konsequent weiterdenken, dann stellen Sie fest, das es für die k-te Kugel noch $n-k+1$ Möglichkeiten gibt, und da man diese Einzelwerte nun multiplizieren muss, erhält man als Anzahl der Möglichkeiten, k Kugeln aus n auszuwählen:

$$n \cdot (n-1) \cdot (n-2) \cdots (n-k+2) \cdot (n-k+1) \ .$$

Das kann man mithilfe der Fakultät-Schreibweise noch etwas eleganter schreiben, und dieses tue ich gleich im nächsten Merkkasten.

Auswahl ohne Zurücklegen mit Beachtung der Reihenfolge

Zur Auswahl von k Kugeln aus einer Menge von n Kugeln ohne Zurücklegen, aber mit Beachtung der Reihenfolge gibt es

$$\frac{n!}{(n-k)!} = n \cdot (n-1) \cdot (n-2) \cdots (n-k+1) \tag{7.2}$$

Möglichkeiten. Die Wahrscheinlichkeit dafür, eine bestimmte Auswahl zu treffen, ist der Kehrwert hiervon, also

$$\frac{(n-k)!}{n!} \ .$$

Beispiel 7.3:
Bei einem Pferderennen mit 12 teilnehmenden Pferden wettet jemand zufällig auf die 5 Erstplatzierten und möchte die Wahrscheinlichkeit dafür wissen, dass er damit richtig liegt. Um diese zu ermitteln, kann man das gerade besprochene Urnenmodell anwenden, denn es handelt sich um Auswahl ohne Zurücklegen, da ein bereits durchs Ziel gelaufenes Pferd nicht nochmal einlaufen kann, und die Reihenfolge ist beim Pferdewetten natürlich wie kaum sonst zu beachten. Die gesuchte Wahrscheinlichkeit berechnet man also nach obiger Formel mit $n = 12$ und $k = 5$ zu

$$\frac{(12-5)!}{12!} = \frac{1}{8 \cdot 9 \cdot 10 \cdot 11 \cdot 12} = \frac{1}{95\,040} \approx 0{,}000\,010\,5\ .$$ ■

Übungsaufgabe 7.3:
In einer Trommel liegen sieben Kugeln, die mit den Buchstaben A, B, E, H, M, N, T beschriftet sind. Man zieht nun nacheinander fünf Kugeln und legt sie in dieser Reihenfolge ab. Wie groß ist die Wahrscheinlichkeit dafür, dass hierbei das Wort „MATHE“ entsteht? ■

Bevor ich zum letzten der vier Urnenmodelle übergehe, möchte ich noch auf einen Spezialfall der Auswahl ohne Zurücklegen mit Beachtung der Reihenfolge hinweisen, der oft fälschlicherweise, oder auch um ihn besonders zu betonen, als eigener Fall dargestellt wird. Sicherlich erinnern Sie sich, dass man per Definition gesetzt hat: $0! = 1$. Der Grund hierfür war und ist, dass dadurch viele Formeln auch für 0 richtig bleiben, und das trifft auch für (7.2) bzw. die daraus als Kehrwert abgeleitete Wahrscheinlichkeit zu. Setzt man nämlich $k = n$, was bedeutet, dass man *alle* Kugeln aus der Urne herausholt, so wird (7.2) zu

$$\frac{n!}{(n-n)!} = \frac{n!}{0!} = n!\ .$$

Es gibt also $n!$ Möglichkeiten, n Kugeln zu entnehmen und dabei die Reihenfolge zu beachten; man nennt dies auch das **vollständige Ziehen** und interpretiert es auch einfach als die Anzahl der Möglichkeiten, n Kugeln oder sonstige Dinge *anzuordnen.*

> **Anordnungen einer Menge**
>
> Es gibt $n!$ Möglichkeiten, um n unterscheidbare Elemente einer Menge anzuordnen.

Auch hierzu ein kleines Beispiel: Der Schachclub *Schachmatt 07* möchte ein Gruppenbild seiner aktuellen Mitglieder machen; dabei sollen sich die 5 Damen in der vorderen Reihe und die 9 Herren in der hinteren Reihe aufstellen. Wie viele Möglichkeiten gibt es, die Personen auf dem Bild anzuordnen?

Der kleine Trick ist hierbei, dass man die Anordnung der Damen und die der Herren zunächst als separate Probleme behandelt. Das ist aber leicht, denn da die 5 Damen vermutlich unterscheidbar sind, gibt es für ihre Anordnung $5! = 120$ Möglichkeiten und aus demselben Grund kann man für die Herren $9! = 362\,880$ Anordnungen finden. Da nun aber für *jede* Anordnung der Damen *sämtliche* Anordnungen der Herren durchgespielt werden können, muss ich zur Ermittlung der insgesamt möglichen Fälle die beiden gerade ermittelten Werte multiplizieren. Die gefragte Anzahl der Möglichkeiten ist also

$$5! \cdot 9! \;=\; 120 \cdot 362\,880 \;=\; 43\,545\,600\,.$$

Die Herrschaften werden also eine ganze Weile durchprobieren müssen, bis sie alle möglichen Aufstellungen getestet haben. ■

Nun komme ich aber endlich zum letzten der vier Urnenmodelle, das man scherzhaft mit „Auswahl ohne alles“ bezeichnen könnte, es handelt sich nämlich um die *Auswahl ohne Zurücklegen und ohne Beachtung der Reihenfolge*. Man wählt also im Modell aus einer Menge von n Kugeln k Stück aus, wobei man eine einmal ausgewählte Kugel nicht mehr zurücklegt, und auch am Ende nicht mehr beachtet, in welcher Reihenfolge man die Kugeln gezogen hat. Auch hierfür benötige ich den oben eingeführten Binomialkoeffizienten, genau genommen ist der gerade für diese Situation vor vielen Jahren eingeführt worden. Es gilt nämlich Folgendes:

Auswahl ohne Zurücklegen ohne Beachtung der Reihenfolge

Zur Auswahl von k Kugeln aus einer Menge von n Kugeln ohne Zurücklegen und ohne Beachtung der Reihenfolge gibt es

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

Möglichkeiten. Die Wahrscheinlichkeit dafür, eine bestimmte Auswahl zu treffen, ist also

$$\frac{1}{\binom{n}{k}} \;=\; \frac{k!(n-k)!}{n!}\,.$$

Vermutlich warten Sie ja schon die ganze Zeit über darauf, dass endlich die Lottozahlen kommen; nun gut, hier sind sie:

Beispiel 7.4:

Das Zahlenlotto, beispielsweise also „6 aus 49“, ist gerade das Standardbeispiel für die Auswahl ohne Zurücklegen ohne Beachtung der Reihenfolge: Eine einmal gezogene Kugel wird beim Lotto nicht wieder zurückgelegt, und auf die Reihenfolge, in der die Kugeln gezogen werden, kommt es nicht an (sonst wäre es auch noch schwieriger, einen Hauptgewinn zu erzielen!). Somit gibt es für die Auswahl von 6 aus 49 Kugeln genau

$$\binom{49}{6} \;=\; \frac{49!}{6!43!} \;=\; 13\,983\,816$$

Möglichkeiten, die Wahrscheinlichkeit, eine bestimmte davon zu treffen, ist also

$$\frac{1}{13\,983\,816} \approx 0{,}000\,000\,071\,5\ .$$

Spätestens jetzt wissen Sie, warum sehr wenige Mathematiker Lotto spielen. ■

Wenn man schon nicht Lotto spielen will, so kann man ja wenigstens Karten spielen; auch hierbei kann die Formel für die Anzahl der Auswahlen hilfreich sein:

Beispiel 7.5:
Ich möchte das Problem klären, wie viele verschiedene Möglichkeiten es gibt, die 32 Karten eines Skatspiels auf die drei Spieler (und den Skat) zu verteilen.

Dazu nehme ich einmal (im Gegensatz zu den internationalen Skatregeln, aber im Einklang mit dem Modell) an, dass zunächst 10 Karten an den ersten Spieler ausgeteilt werden. Ist der Stapel gut gemischt, schließen wir also aus, dass Tante Erna dabei ist, die immer nur einmal abhebt und nicht mischt, so gibt es hierfür

$$\binom{32}{10}$$

Möglichkeiten. Auch der zweite Spieler erhält 10 Karten, wofür jetzt aber nur noch 22 zur Auswahl stehen; folglich gibt es für das Blatt des zweiten Spielers

$$\binom{22}{10}$$

Möglichkeiten. Und da ich Sie um alles in Welt nicht langweilen will, sage ich jetzt ohne Umschweife, dass es für das Blatt des dritten Spielers nur noch

$$\binom{12}{10}$$

Möglichkeiten gibt. Haben aber alle drei Spieler ihre Karten, so ist der Skat (also die restlichen zwei Karten) eindeutig festgelegt, hierfür gibt es also keine Auswahlmöglichkeit mehr. Die anfangs gefragte Anzahl der Möglichkeiten für die Kartenverteilung erhalte ich nun also durch Multiplikation der gerade ermittelten Einzelmöglichkeiten; es ergibt sich

$$\binom{32}{10} \cdot \binom{22}{10} \cdot \binom{12}{10} = \frac{32! \cdot 22! \cdot 12!}{10! \cdot 22! \cdot 10! \cdot 12! \cdot 10! \cdot 2!} = \frac{32!}{10! \cdot 10! \cdot 10! \cdot 2!} \approx 2{,}753 \cdot 10^{15}.$$

■

Das war in meinem Alter ganz schön anstrengend; daher ziehe ich mich jetzt einmal kurz zurück und überlasse Ihnen das Feld:

Übungsaufgabe 7.4:

a) Ein Fußballverein hat 15 aktive Spieler. Wie viele Möglichkeiten hat der Trainer, hieraus eine Mannschaft von 11 Spielern zu formen? (Auch wenn es meinem Sohn, der aktiver Fußballer ist, weh tut, nehme ich hierbei zur Vereinfachung an, dass jeder Spieler jede Position, also auch die des Torwarts, einnehmen kann.)

b) Der Trainer aus Teil a) hat gekündigt, da er mit einer Mannschaft aus 15 Alleskönnern nicht arbeiten kann. Der neue Trainer analysiert das Spielerpotenzial genauer und stellt fest, dass es sich um 7 Verteidiger, 6 Stürmer und 2 Torleute handelt (das Mittelfeld hat er aufgegeben). Wie viele Möglichkeiten hat der Trainer, hieraus eine Mannschaft zu formen, die aus 5 Verteidigern, 5 Stürmern und einem Torwart besteht? ■

Ich hatte ja bereits angedroht, dass ich noch ein paar Übungsaufgaben formulieren wollte, bei denen Sie selbst entscheiden müssen, welches der oben formulierten Urnenmodelle angewandt werden kann; nun ist es so weit:

Übungsaufgabe 7.5:

a) Beim klassischen Fußballtoto muss man bei insgesamt 11 Spielen auf Unentschieden (0), Heimsieg (1) oder Auswärtssieg (2) tippen. Wie viele verschiedene Tipps sind hier möglich?
b) Sie haben die Wahl, entweder beim Lotto „5 aus 25“ oder beim Lotto „4 aus 20“ mitzuspielen. Wobei ist die Chance auf einen Hauptgewinn höher?
c) An einer Bushaltestelle besteigen vier Fahrgäste den Bus und finden sieben freie Sitzplätze vor. Wie viele Möglichkeiten haben sie, sich auf vier dieser Plätze zu verteilen?
d) Bei einem Sportturnier müssen die zwölf teilnehmenden Mannschaften auf drei Gruppen mit je vier Mannschaften verteilt werden. Wie viele Möglichkeiten hat der Veranstalter hierfür? ■

7.2 Relative Häufigkeit und klassische Definition der Wahrscheinlichkeit

In diesem Kapitel habe ich nun schon etwa 10 Seiten lang über Wahrscheinlichkeiten geschrieben und diese ja auch in gewissen Fällen berechnet. So richtig *definiert* habe ich den Begriff bisher aber noch nicht, und das ist auch kein Wunder, denn eine solche Definition ist gar nicht so leicht, zumindest dann, wenn sich die solchermaßen definierte Wahrscheinlichkeit mit den berechneten Zahlen decken soll. Das fängt schon ganz am Anfang an, wenn ich beispielsweise angenommen habe, dass eine bestimmte Seite einer Münze nach dem Wurf mit der „Wahrscheinlichkeit 1/2“ oben liegt. Das mag einer gewissen Lebenserfahrung entsprechen, aber es lässt sich durch kein mathematisches Theorem der Welt beweisen.

Um dem Begriff der Wahrscheinlichkeit näher zu kommen, ist es eine gute Idee, sich mit der so genannten Häufigkeit von Ereignissen zu befassen:

Zufallsversuch und zufälliges Ereignis

Ein **zufälliges Ereignis**, meist kurz als **Ereignis** bezeichnet, ist das Ergebnis eines Zufallsversuchs. Ein **Zufallsversuch** wiederum ist ein (zumindest theoretisch) beliebig oft wiederholbarer Versuch, dessen Ausgang nicht vorhersagbar ist.

Ein solcher Zufallsversuch ist beispielsweise der Münzwurf oder das Ziehen von Kugeln aus einer Urne, wobei dann „Wappen oben“ oder „2 und 5 gezogen“ zugehörige zufällige Ereignisse sind.

Mithilfe dieser Begrifflichkeiten kann ich nun die Häufigkeit(en) zufälliger Ereignisse definieren:

Absolute und relative Häufigkeit

Ein Zufallsversuch werde n-mal wiederholt und hierbei trete das zufällige Ereignis A genau a-mal auf.

Dann nennt man die Zahl

$$H_n(A) \;=\; a$$

die **absolute Häufigkeit** des Ereignisses A und die Zahl

$$h_n(A) \;=\; \frac{a}{n}$$

die **relative Häufigkeit** des Ereignisses A.

Von größerer Aussagekraft ist hierbei sicherlich die relative Häufigkeit $h_n(A)$, denn sie gibt an, in welchem Bruchteil aller n Fälle das Ereignis A eingetreten ist.

Da n eine positive Zahl und a eine ganze Zahl zwischen 0 und n ist, ist $h_n(A)$ eine reelle Zahl zwischen 0 und 1, genauer gesagt ist $h_n(A)$ gleich einer der Zahlen

$$0\,,\ \frac{1}{n}\,,\ \frac{2}{n}\,,\ \ldots\,,\ \frac{n-1}{n}\,,\ 1\ .$$

Tritt beispielsweise bei einem 2 000-mal durchgeführten Versuch 864-mal das Ereignis A ein, so ist

$$H_{2\,000}(A) \;=\; 864$$

und

$$h_{2\,000}(A) \;=\; \frac{864}{2\,000} \;=\; 0{,}432\ .$$

Ob Sie's glauben oder nicht, aber um Ihnen hier ein realistisches Zahlenbeispiel geben zu können, habe ich mich beim Schreiben dieser Zeilen hingesetzt und 100-mal eine Münze geworfen; dabei trat genau 47-mal das Ereignis W = Wappen oben ein. Für dieses Ereignis gilt also

$$H_{100}(W) \;=\; 47$$

und

$$h_{100}(W) = \frac{47}{100} = 0{,}47 \, .$$

Vielleicht ist Ihnen das Münzenwerfen zu langweilig? In diesem Fall lade ich Sie zum Würfeln ein:

Übungsaufgabe 7.6:
Beim Würfeln mit einem Standardwürfel traten die sechs Augenzahlen mit folgenden absoluten Häufigkeiten auf:

Augenzahl	1	2	3	4	5	6
Häufigkeit	137	140	127	120	140	133

Berechnen Sie die relativen Häufigkeiten der folgenden Ereignisse in diesem Zufallsversuch:

A_1 : Wurf einer 5

A_2 : Wurf einer geraden Zahl

A_3 : Wurf einer ungeraden Zahl, die nicht durch 3 teilbar ist

■

Ich hatte oben ein wenig damit angegeben, dass ich mir die Mühe gemacht habe und 100-mal eine Münze geworfen habe. Das ist zwar für einen alternden Mathematikprofessor eine ganz ansehnliche Leistung, aber es genügt natürlich noch nicht, um eine gesicherte Aussage über die Häufigkeit zu machen, die für das Auftreten von W zu erwarten ist.

Zum Glück gab und gibt es weitaus fleißigere Menschen als mich, die viele Tausend Würfe mit einer Münze auf sich genommen haben, in früheren Zeiten tatsächlich am Spieltisch, heute auch gerne als Computersimulation.

Das Ergebnis, und *das* ist tatsächlich eines der Mysterien der Wahrscheinlichkeitsrechnung, ist überall auf der Welt und zu allen Zeiten der Menschheit dasselbe: Nach einer gewissen Anzahl von Würfen pendelt sich die relative Häufigkeit des Wappenwurfs bei $0,5$ ein. Anders formuliert: In ziemlich genau der Hälfte aller Fälle wird die Münze oben das Wappen zeigen. Wohlgemerkt: Kein Mensch kann vorhersagen, wie der nächste, der übernächste, der drittnächste, ... Wurf ausgehen wird, aber man kann mit großer Sicherheit vorhersagen, dass man nach 100 000 Würfen etwa 50 000-mal das Wappen oben gesehen haben wird.

Bevor Sie mich fragen: Ich habe *keine* Ahnung, warum das so ist! Ich kann zwar mit einiger Mühe die verschachteltsten Wahrscheinlichkeiten zusammengesetzter, bedingter und sonstwie verunstalteter Ereignisse berechnen, aber ich kann Ihnen wirklich nicht sagen, warum sich diese verflixte Münze so verhält, wie sie es nun mal tut.

Das ist bei komplexeren Systemen nicht anders; man kann beispielsweise bei radioaktiven Materialien ziemlich genau vorher sagen, wie viel des anfangs vorhandenen Materials nach einer gewissen Zeiteinheit noch vorhanden und wie viel zerfallen ist, das sind relativ einfache Gleichungen, deren Ergebnisse aber dennoch mit der Wirklichkeit verblüffend gut übereinstimmen. Aber von *einem bestimmten* einzelnen Teilchen kann kein Mensch und keine Formel der Welt sagen, ob es in der nächsten Sekunde zerfällt oder nicht.

Wahrscheinlichkeit ist also etwas, was man nur „im Großen" berechnen kann; auch wenn es – das sei für die mathematischen Puristen unter Ihnen gesagt – das Pferd ein wenig von hinten aufzäumt, so kann man doch die Wahrscheinlichkeit eines Ereignisses intuitiv definieren als denjenigen Wert, auf den sich die relativen Häufigkeiten dieses Ereignisses für sehr große Versuchsanzahlen einpendeln. Berechnen kann man die Wahrscheinlichkeit eines Ereignisses auf diese Weise allerdings nicht, man erhält nur Schätzwerte; außerdem ist dies nur auf relativ einfache Versuche wie Münzwürfe oder Würfelspiele beschränkt, denn beispielsweise die Wahrscheinlichkeit eines Super-GAUs im Atomkraftwerk durch relative Häufigkeiten dieses Ereignisses, also mithilfe einer genügend langen Versuchsreihe zu ermitteln, halte ich nicht für empfehlenswert.

Es gibt eine etwas bessere Definition des Begriffs Wahrscheinlichkeit, die ich Ihnen nicht vorenthalten will. Man nennt sie auch die **klassische Definition von Wahrscheinlichkeit** und sie geht immerhin zurück auf den großen Mathematiker und Wahrscheinlichkeitstheoretiker Pierre S. de Laplace. Aber auch große Leute sind nicht perfekt, und diese Definition ist es auch nicht, da sie nur für Zufallsversuche mit *endlich vielen* Ausgängen anwendbar ist, und auch nur für solche, bei denen jeder einzelne Ausgang dieselbe Einzelwahrscheinlichkeit hat. Im Prinzip also genau die Situation, die ich im ersten Abschnitt über Kombinatorik angenommen hatte. Insofern ist diese Definition zumindest als erster Ansatz gar nicht so verkehrt, weshalb ich mich jetzt auch erst mal eine Weile damit befasse.

Klassische Definition der Wahrscheinlichkeit

Ein Zufallsversuch habe N verschiedene gleich wahrscheinliche Ausgänge, in genau k davon trete das Ereignis A ein. Dann nennt man die Zahl

$$p(A) = \frac{k}{N}$$

die **(klassisch definierte) Wahrscheinlichkeit** des Ereignisses A. Manchmal sagt man auch k sei die Anzahl der für A **günstigen Fälle**.

Der Buchstabe p steht für das englische probability und dürfte leichter verständlich sein als diese Definition. Um diese wiederum verständlicher zu machen, hier ein erstes Beispiel:

Beispiel 7.6:
In einer Urne sind 20 Kugeln enthalten, die von 1 bis 20 durchnummeriert sind. Wie groß ist die Wahrscheinlichkeit für das Ereignis U, eine Kugel zu ziehen, deren Nummer ungerade, aber nicht durch 3 teilbar ist?

In diesem einfachen Fall zählt man die günstigen Fälle einfach durch: Es sind die Nummern $1, 5, 7, 11, 13, 17, 19$, also sieben Stück, folglich ist

$$p(U) = \frac{7}{20} = 0{,}35 \ . \qquad \blacksquare$$

Übungsaufgabe 7.7:
Bei einer Tombola werden insgesamt 500 Lose verkauft; darunter ist ein Hauptgewinn, 25 hochwertige Gewinne, 140 Trostpreise, der Rest besteht aus Nieten. Wie hoch ist die Wahrscheinlichkeit für das Ereignis G, das darin besteht, überhaupt etwas zu gewinnen? ■

Einer der großen Vorteile der obigen Definition von Wahrscheinlichkeit ist die Tatsache, dass man damit „richtig rechnen" kann; ich zeige Ihnen hierfür zwei Beispiele.

Hat der Versuch, so wie in der Definition, N Ausgänge, wovon in genau k Fällen das Ereignis A eintritt, so bedeutet dies, dass in genau $N-k$ Fällen A *nicht eintritt*. Man nennt das Ereignis „A tritt nicht ein" meist das **Gegenereignis** von A und bezeichnet es mit $\overline{A}$; da es in genau $N-k$ der N Fälle eintritt, hat es die Wahrscheinlichkeit

$$p(\overline{A}) = \frac{N-k}{N} = 1 - \frac{k}{N} \ .$$

Da aber $\frac{k}{N}$ gerade die Wahrscheinlichkeit von $p(A)$ war, haben wir folgende Rechenregel hergeleitet:

Wahrscheinlichkeit des Gegenereignisses

Die Wahrscheinlichkeiten eines Ereignisses A und des Gegenereignisses $\overline{A}$ hängen wie folgt zusammen:

$$p(\overline{A}) = 1 - p(A) \ .$$

Dies will ich als erstes Beispiel für das erwähnte „Rechnen mit Wahrscheinlichkeiten" verstanden haben. Um es zu illustrieren, greife ich nochmals Beispiel 7.6 auf. Das Gegenereignis $\overline{U}$ des dort definierten Ereignisses U besteht gerade darin, eine Zahl zu ziehen, die entweder gerade oder durch 3 teilbar ist (oder beides). Die hierfür günstigen Fälle sind die Nummern $2, 3, 4, 6, 8, 9, 10, 12, 14, 15, 16, 18, 20$, also dreizehn Stück. Die Wahrscheinlichkeit von $\overline{U}$ ist somit

$$p(\overline{U}) = \frac{13}{20} = 0{,}65 \ ,$$

also gerade $1 - p(U) = 1 - 0{,}35$.

Für das zweite Beispiel für das „Rechnen mit Wahrscheinlichkeiten" brauchen wir zwei Ereignisse, sagen wir A und B, und eine neue Begrifflichkeit:

Unvereinbare Ereignisse

Zwei Ereignisse A und B heißen **unvereinbar**, wenn sie niemals gleichzeitig eintreten, wenn also das Eintreten von B unmöglich ist, falls A eintritt, und umgekehrt.

Beispielsweise sind beim Würfeln mit einem Standardwürfel die Ereignisse

$$A = \text{Wurf einer 1}$$

und

$$B = \text{Wurf einer geraden Zahl}$$

unvereinbar.

Aus zwei (oder mehr) Ereignissen kann man ein neues basteln, indem man das Eintreten von mindestens einem der beiden bereits als neues Ereignis feiert; formal nennt man dieses neue Ereignis die Summe der beiden Ausgangsereignisse und symbolisiert es durch das Zeichen „$\cup$“; wenn Ihnen hierbei unangenehme Assoziationen zur Mengenlehre kommen, so hat das durchaus seine Berechtigung (die Assoziationen, nicht unbedingt die Tatsache, dass diese unangenehm sind): Tatsächlich wird die ganze moderne Wahrscheinlichkeitsrechnung axiomatisch mithilfe der Mengenlehre aufgebaut, indem man Ereignisse als Mengen auffasst und ihre Wahrscheinlichkeit als das „Maß“ dieser Menge, gewissermaßen also ihre Größe, definiert. Ich werde im nächsten Abschnitt, auf den wir gerade zusteuern, einen Schritt in Richtung dieser axiomatischen Sichtweise tun, aber um Ihnen (und mir) die Einführung der vollen Mengenlehre zu ersparen, werde ich mich dabei auf das wahrscheinlichkeitsrechnerisch Nötigste beschränken. Das hat dann leider den kleinen Nachteil, dass ein paar Begriffe (wie hier die Summe von Ereignissen) eingeführt werden, deren voller Hintergrund bzw. deren Namensursache im Verborgenen bleiben, aber ich denke, damit können Sie an dieser Stelle leben.

Jetzt aber:

Summe von Ereignissen

Sind A und B zwei Ereignisse, so kann man ein neues Ereignis $\mathbf{A \cup B}$ definieren, das man als **Summe von A und B** bezeichnet. Dieses Ereignis tritt ein, wenn *mindestens eines* der beiden Ereignisse A oder B eintritt.

Die Summe der beiden oben definierten Ereignisse wäre somit

$$A \cup B \; = \text{Wurf einer 1 oder einer geraden Zahl},$$

also

$$A \cup B \; = \text{Wurf einer 1, 2, 4 oder 6}.$$

Die Wahrscheinlichkeit des Ereignisses $A \cup B$ kann man hier direkt erkennen: Da es in vier von sechs möglichen Fällen eintritt, ist

$$p(A \cup B) \;=\; \frac{4}{6} \;=\; \frac{2}{3}.$$

Dieses Ergebnis bekommt man aber auch auf einem anderen Wege heraus: Das Ereignis A, also der Wurf der Zahl 1, hat offenbar die Wahrscheinlichkeit $1/6$, während das Ereignis B, also Wurf einer geraden Zahl, die Wahrscheinlichkeit $1/2$ besitzt. Da beide Ereignisse unvereinbar sind, also niemals gleichzeitig eintreten können, kann ich ja mal versuchen, diese Wahrscheinlichkeiten zu addieren, und unter Benutzung der Regeln für das Bruchrechnen erhalte ich:

$$p(A) + p(B) \;=\; \frac{1}{6} + \frac{1}{2} \;=\; \frac{1}{6} + \frac{3}{6} \;=\; \frac{2}{3},$$

also das identische Ergebnis. Dies ist nicht nur in diesem Beispiel, sondern generell richtig und das sollte man – als zweite Rechenregel für Wahrscheinlichkeiten – festhalten:

Summe unvereinbarer Ereignisse

Sind A und B unvereinbare Ereignisse, so kann man die Wahrscheinlichkeit ihrer Summe wie folgt berechnen:

$$p(A \cup B) \;=\; p(A) + p(B)\ . \tag{7.3}$$

Schlagwortartig kann man also sagen: „Die Wahrscheinlichkeit der Summe ist gleich der Summe der Wahrscheinlichkeiten.“ Dies gilt aber, daran sei nochmals erinnert, nur für unvereinbare Ereignisse!

Übungsaufgabe 7.8:
Die sechs Seiten eines Würfels seien mit den Zahlen 1, 1, 3, 3, 4, 5 bedruckt. Wir betrachten die drei Ereignisse

$$\begin{aligned} A &= \text{ Wurf einer geraden Zahl,} \\ B &= \text{ Wurf einer durch 3 teilbaren Zahl,} \\ C &= \text{ Wurf der Zahl 5.} \end{aligned}$$

Zeigen Sie, dass diese drei Ereignisse paarweise unvereinbar sind, dass also jedes mögliche Zweierpärchen, das man aus diesen Ereignissen bilden kann, unvereinbar ist, und berechnen Sie auf zwei verschiedene Arten die Wahrscheinlichkeit der Vereinigung von je zweien dieser Ereignisse. ■

7.3 Axiomatische Definition der Wahrscheinlichkeit

Ich glaube, inzwischen sind Sie so weit in mathematische Denkweisen vorgedrungen, dass ich Sie mit einer axiomatischen Definition konfrontieren kann – keine Sorge, ich bin ja bei Ihnen. Ein Axiom ist laut Mathematik-Lexikon „eine Aussage, die wegen ihres Inhalts grundlegend ist und als evident gilt und daher keines Beweises bedarf“. Mit anderen Worten: Ein Axiom wird als Grundlage einer Theorie ohne weitere Begründung oder gar Motivation einfach hingeknallt und danach werden daraus Folgerungen gezogen und die ganze Theorie aufgebaut.

Im Fall der Wahrscheinlichkeitsrechnung hat dies zu Beginn des 20. Jahrhunderts der russische Mathematiker A. N. Kolmogoroff getan; er hat sich dabei von den oben angegebene Rechenregeln leiten lassen, die für „vernünftige“ Wahrscheinlichkeiten offenbar gelten müssen, und dann aufgeschrieben, was für Eigenschaften eine Abbildung von der Menge der Ereignisse in die der reellen Zahlen besitzen muss, damit man sie als „Wahrscheinlichkeit“ bezeichnen kann. Herausgekommen ist dabei Folgendes:

Axiomatische Definition der Wahrscheinlichkeit

Auf der Menge aller Ereignisse, die als Ergebnis eines Zufallsversuchs entstehen können (dies können jetzt auch unendlich viele sein), definiere man eine Funktion p, die jedem Ereignis eine reelle Zahl zuordnet. Diese Funktion bzw. ihre Werte bezeichnet man als **(axiomatisch definierte) Wahrscheinlichkeit**, wenn sie folgende Eigenschaften besitzt:

(1) Für jedes Ereignis A gilt $0 \le p(A) \le 1$, die Funktion p nimmt also nur Werte zwischen 0 und 1 an.

(2) Das so genannte **sichere Ereignis**, das alle möglichen Ereignisse in sich vereinigt und das man traditionell mit Ω („Omega“) bezeichnet, hat den Wert 1: $p(\Omega) = 1$.

(3) Für jedes Paar unvereinbarer Ereignisse A und B gilt

$$p(A \cup B) = p(A) + p(B) \; .$$

(4) Eigenschaft (3) gilt ebenfalls für die Summe beliebig vieler, also auch unendlich vieler Ereignisse, wenn sie paarweise unvereinbar sind.

Ich denke, ich habe mir jetzt erst mal ein Stück Schokolade verdient und Sie ein paar erläuternde Bemerkungen, die Ihnen diese Definition verständlicher machen sollen:

(a) Es gibt in der Mathematik viele abstruse Abbildungen, die gemäß dieser Definition eine „Wahrscheinlichkeit“ darstellen, die aber wenig mit dem zu tun haben, was man landläufig darunter verstehen wird; falls dies nur Ihr Vorurteil über Mathematik bestätigt, kann ich’s jetzt leider auch nicht ändern. In diesem Buch werde ich mich aber nur mit „normalen“ Wahrscheinlichkeiten befassen, und auch nur mit solchen, die nur endlich viele Ereignisse zulassen. Falls Sie übrigens daran zweifeln sollten, dass es überhaupt solche mit unendlich vielen Ereignisse gibt, dann spielen Sie doch mal mit einem Partner das (von mir gerade erfundene) Spiel „Denke dir eine beliebige reelle Zahl aus!“ Da es unendlich viele Zahlen gibt, hat dieses Zufallsexperiment auch ebenso viele Ereignisse.

(b) Zurück zur Analyse unserer Axiome: Die Eigenschaft (1) besagt, dass eine Wahrscheinlichkeit stets eine Zahl zwischen 0 und 1 ist, (2) sagt also, dass das sichere Ereignis Ω die maximale Wahrscheinlichkeit hat. Noch ein Wort zu diesem „sicheren Ereignis“: Dieses besteht, wie schon geschrieben, aus der Zusammenfassung *aller möglichen* Ereignisse und tritt somit mit Sicherheit ein; Beispiele hierfür sind das Werfen einer der Zahlen 1, 2, 3, 4, 5, 6 beim Würfeln oder das Ziehen einer der Zahlen $1, 2, \ldots, 49$ beim Standardlotto.

(c) Das Gegenteil des sicheren ist das **unmögliche Ereignis**, symbolisiert durch das Zeichen $\emptyset$. Beispiele hierfür sind das Werfen einer 7 beim Würfeln, das Ziehen einer negativen Zahl beim Lotto oder das Schreiben eines tippfehlerfreien Mathematikbuches.

(d) Vielleicht wundert es Sie, dass man zusätzlich zu (3) noch (4) fordern muss, obwohl es doch so aussieht, als könne man dies aus (3) herleiten. Nun, ich kann hier nur sagen, ohne dies in diesem Rahmen beweisen zu können oder wollen, dass diese Herleitung *nicht* möglich ist, falls es sich um unendlich viele Ereignisse handelt, was ich ja zugelassen hatte.

Oben hatte ich Werbung damit gemacht, dass man für axiomatisch definierte Wahrscheinlichkeiten Rechenregeln hat, und es wird nun Zeit, dieses Versprechen einzulösen; um zu verdeutlichen, dass diese zu den Axiomen als Regeln *hinzukommen,* zähle ich einfach weiter.

Rechenregeln für Wahrscheinlichkeiten

Neben den oben genannten Axiomen gelten für Wahrscheinlichkeiten die folgenden Regeln:

(5) Das unmögliche Ereignis hat die Wahrscheinlichkeit null: $p(\emptyset) = 0$.
(6) Ist $\overline{A}$ das Gegenereignis zu A, so gilt $p(\overline{A}) = 1 - p(A)$.
(7) Ist ein Ereignis A Teil eines anderen Ereignisses B, impliziert also das Eintreten von A auf alle Fälle das von B, so gilt $p(A) \leq p(B)$.

Da das unmögliche Ereignis gerade das Gegenereignis zum sicheren Ereignis Ω ist, folgt (5) übrigens gerade aus (6) und dem Axiom $p(\Omega) = 1$, aber es ist üblich, diese Eigenschaft dennoch separat aufzuschreiben.

Um die in (7) benutzte Sprechweise zu illustrieren, bemühe ich nochmals das Würfelspiel und definiere die Ereignisse

$$A = \text{Wurf einer } 2$$

und

$$B = \text{Wurf einer geraden Zahl.}$$

Sicherlich ist A Teil von B, denn 2 ist eine gerade Zahl; auch die zweite Sprechweise wird hierdurch deutlich, denn das Ereignis A, eine 2 gewürfelt zu haben, *impliziert*, dass eine gerade Zahl gewürfelt wurde, also B. Und auch die behauptete Ungleichung stimmt hier, denn $p(A) = 1/6$ ist sicherlich kleiner als $p(B) = 1/2$.

Nun wird es aber höchste Zeit für ausführlicheres Beispiel:

Beispiel 7.7:
Diesmal würfeln wir mit einem Würfel aus der Geheimwerkstatt von Al Capone. Für das Würfeln der Augenzahlen 2 bis 6 hat dieser folgende Wahrscheinlichkeiten:

$$p(6) = p(5) = \frac{1}{5} \quad \text{und} \quad p(4) = p(3) = p(2) = \frac{1}{6} \ .$$

(a) Mit welcher Wahrscheinlichkeit wird mit diesem Würfel eine 1 gewürfelt?

Um dies zu beantworten, muss ich lediglich beachten, dass die Summe aller Wahrscheinlichkeiten, die ja das sichere Ereignis ist, 1 sein muss. Da die Summe der angegebenen fünf Wahrscheinlichkeiten gerade

$$\frac{2}{5} + \frac{3}{6} = \frac{9}{10}$$

ist, bleibt nur

$$p(1) = \frac{1}{10}$$

übrig.

(b) Wie groß ist die Wahrscheinlichkeit, mit diesem Würfel *keine* 6 zu würfeln?

Das ist ganz einfach: „Keine 6“ zu würfeln ist gerade das Gegenereignis davon, eine 6 zu würfeln, was wiederum nach Angabe die Wahrscheinlichkeit 1/5 hat. Somit ist

$$p(\text{keine } 6) = 1 - \frac{1}{5} = \frac{4}{5} \ .$$

(c) Mit welcher Wahrscheinlichkeit würfelt man mit diesem Würfel eine ungerade Zahl?

Da die Würfe der drei ungeraden Zahlen jeweils unvereinbar sind (man kann ja nur entweder eine 1 oder eine 3 oder eine 5 würfeln), muss man zur Beantwortung dieser Frage nach Axiom (3) nur die drei Einzelwahrscheinlichkeiten aufaddieren; es folgt

$$p(\text{ungerade Zahl}) = p(1) + p(3) + p(5) = \frac{1}{10} + \frac{1}{6} + \frac{1}{5} = \frac{7}{15} \ . \qquad \blacksquare$$

Den eigentlichen „Knaller“ unter den Rechenregeln für Wahrscheinlichkeiten habe ich mir für den Schluss aufbewahrt, nämlich die Formel für die Wahrscheinlichkeit einer Summe von Ereignissen, die nicht unbedingt unvereinbar sind. Um diese nun formulieren zu können, brauche ich noch einen letzten neuen Begriff in diesem Abschnitt, nämlich den des Produkts von Ereignissen:

Produkt von Ereignissen

Sind A und B zwei Ereignisse, so kann man ein neues Ereignis $\mathbf{A} \cap \mathbf{B}$ definieren, das man als **Produkt von A und B** bezeichnet. Dieses Ereignis tritt ein, wenn *sowohl A als auch B* eintritt.

Das Produkt der beiden Ereignisse

$$A = \text{Wurf einer 1 oder 2}$$

und

$$B = \text{Wurf einer geraden Zahl}$$

wäre also das Ereignis

$$A \cap B = \text{Wurf einer 2} \ ,$$

denn nur in diesem Fall ist sowohl A als auch B eingetreten.

Nun bin ich in der Lage, Ihnen die folgende Formel anzugeben, die mit Sicherheit die am häufigsten gebrauchte Rechenregel beim Umgang mit Wahrscheinlichkeiten ist:

Summe beliebiger Ereignisse

Für beliebige (also nicht unbedingt unvereinbare) Ereignisse A und B gilt

$$\mathbf{p}(\mathbf{A} \cup \mathbf{B}) \ = \ \mathbf{p}(\mathbf{A}) + \mathbf{p}(\mathbf{B}) - \mathbf{p}(\mathbf{A} \cap \mathbf{B}) \ . \tag{7.4}$$

Die Fülle von Regeln, die Sie inzwischen kennen gelernt haben, schreit nun wirklich nach ausführlichen Beispielen, und da wir mit dem Stoff dieses Abschnitts „durch" sind, will ich auch den ganzen Rest nur noch damit bestreiten.

Beispiel 7.8: Bei einem Zufallsversuch sind insgesamt vier Ereignisse A, B, C, D möglich, die alle paarweise unvereinbar sein sollen. Bekannt sind folgende Daten:

$$p(A \cup B) = \frac{13}{21} \ , \quad p(A \cup C) = \frac{8}{15} \ , \quad p(B \cup C) = \frac{17}{35} \ .$$

Man berechne nun die Einzelwahrscheinlichkeiten der vier Ereignisse.

Da die Ereignisse unvereinbar sind, kann ich die spezielle Formel (7.3) verwenden; ich schreibe zunächst einmal auf, was ich weiß: Es ist

$$\begin{aligned} p(A) + p(B) &= p(A \cup B) &= \frac{13}{21} \ , \\ p(A) + p(C) &= p(A \cup C) &= \frac{8}{15} \ , \\ p(B) + p(C) &= p(B \cup C) &= \frac{17}{35} \ , \end{aligned}$$

und das ist ein wunderschönes lineares Gleichungssystem mit drei Zeilen zur Berechnung der drei unbekannten Einzelwahrscheinlichkeiten. Mit den Methoden, die Sie im Kapitel über Lineare Algebra kennen gelernt haben, finden Sie leicht die Lösungen

$$p(A) = \frac{1}{3}\ , \quad p(B) = \frac{2}{7}\ , \quad p(C) = \frac{1}{5}$$

heraus. Da es insgesamt nur noch *ein* weiteres Ereignis, nämlich D, geben soll, ist dieses gerade das Gegenereignis dieser drei. Es folgt

$$p(D) \ = \ 1 - (p(A) + p(B) + p(C)) \ = \ 1 - (\frac{1}{3} + \frac{2}{7} + \frac{1}{5}) \ = \ 1 - \frac{86}{105} \ = \ \frac{29}{105}\ .$$

■

Und weil's so schön ist, gleich noch ein weiteres Beispiel:

Beispiel 7.9: In einem Kunststoff verarbeitenden Betrieb wird festgestellt, dass 5 % der produzierten Teile Verformungen und 8 % Farbunechtheiten aufweisen; 3 % aller Teile haben sogar beide Fehler. Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewähltes Teil
(a) mindestens einen der beiden Fehler,
(b) höchstens einen der beiden Fehler,
(c) gar keinen Fehler

aufweist?

Bei solchen „Textaufgaben“ sollte man erst einmal die Prozentangaben in Wahrscheinlichkeiten übersetzen. Ich bezeichne das Ereignis, dass ein zufällig ausgewähltes Teil eine Verformung aufweist, mit V, entsprechend mit F die Farbunechtheit. Dann gilt, da p % gerade $p/100$ sind,

$$p(V) = \frac{5}{100} = 0{,}05\ , \quad p(F) = \frac{8}{100} = 0{,}08 \quad \text{sowie} \ \ p(V \cap F) = \frac{3}{100} = 0{,}03\ .$$

Letzteres, da „beide Fehler“ ja gerade das gleichzeitige Eintreten von V und F bedeutet. Nun braucht man nur noch die genannten Rechenregeln geeignet einzusetzen:

(a) Dies ist das Ereignis $V \cup F$ und nach (7.4) ist

$$p(V \cup F) \ = \ p(V) + p(F) - p(V \cap F) \ = \ 0{,}05 + 0{,}08 - 0{,}03 \ = \ 0{,}1\ .$$

(b) Dies ist das Gegenereignis zu „beide Fehler gleichzeitig haben“, also zu $V \cap F$. Somit ist

$$p(\text{höchstens ein Fehler}) \ = \ 1 - p(V \cap F) \ = \ 0{,}97\ .$$

(c) Auch hier kommt man mit dem Gegenereignis-Argument am schnellsten ans Ziel; dieses ist nun $V \cup F$, somit ist

$$p(\text{kein Fehler}) \ = \ 1 - p(V \cup F) \ = \ 0{,}9\ .$$

■

Übungsaufgabe 7.9:
In der Unterstufe eines Gymnasiums werden die Fremdsprachen Englisch und Latein angeboten, jeder Schüler muss mindestens eine Sprache erlernen. Von den insgesamt 50 Schülern erlernen 35 Englisch und 25 Latein. Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Schüler

(a) beide Sprachen erlernt,
(b) nur Englisch erlernt,
(c) nur eine Sprache erlernt? ■

Übungsaufgabe 7.10:
Zeigen Sie, dass die Formel (7.3) für die Summe unvereinbarer Ereignisse als Spezialfall in der Formel (7.4) enthalten ist. ■

7.4 Bedingte Wahrscheinlichkeiten

Dieser letzte Abschnitt wird recht kurz, das verspreche ich Ihnen, aber er ist notwendig, denn man kann das Kapitel Wahrscheinlichkeitsrechnung nicht abschließen, ohne über bedingte Wahrscheinlichkeiten gesprochen zu haben. Diesen Begriff will ich Ihnen nun näher bringen und ich beginne wie meist mit einem Beispiel: Die Wahrscheinlichkeit, beim Standardwürfeln eine 6 zu würfeln, ist, wie Sie nun schon längst wissen, gleich 1/6. Nehmen wir nun aber einmal an, jemand würde Ihnen garantieren, dass auf jeden Fall eine gerade Zahl fällt (das wäre also die „Bedingung“), so sollte man doch annehmen, dass dies die Wahrscheinlichkeit für eine 6 auf 1/3 erhöht, da dann ja nur noch drei Zahlen zur Auswahl stehen.

Dies ist auch tatsächlich der Fall, und um dieses allgemein berechnen zu können, benutzt man die Formel für die bedingte Wahrscheinlichkeit:

Bedingte Wahrscheinlichkeit

Ich nehme an, A und B seien zwei Ereignisse desselben Zufallsversuchs und das Ereignis B habe nicht die Wahrscheinlichkeit null: $p(B) > 0$. Dann heißt $p(A|B)$, definiert durch

$$p(A|B) = \frac{p(A \cap B)}{p(B)} \tag{7.5}$$

die **Wahrscheinlichkeit des Ereignisses A unter der Bedingung B** oder kurz die **bedingte Wahrscheinlichkeit von A bzgl. B**.

Es wird also die Wahrscheinlichkeit des Produkts von A und B durch diejenige von B dividiert, und da diese nach Voraussetzung nicht null ist, kann dabei nichts passieren.

Ich greife das Eingangsbeispiel wieder auf: A sei das Ereignis „Wurf einer 6" mit $p(A) = 1/6$ und B sei das Ereignis „Wurf einer geraden Zahl" mit $p(B) = 1/2$. Um Formel (7.5) anwenden zu können, brauche ich noch die Wahrscheinlichkeit des Produkts $A \cap B$ von A und B. Dieses besteht gerade aus dem Wurf einer 6 und somit ist $p(A \cap B) = 1/6$. Insgesamt erhalten wir als Wahrscheinlichkeit für den Wurf einer 6 unter der Bedingung, dass eine gerade Zahl gewürfelt wird:

$$p(A|B) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

in Übereinstimmung mit der „Intuition".

Generell ist das Würfeln sehr gut dazu geeignet, die Formel für die bedingte Wahrscheinlichkeit zu illustrieren, da man die Ergebnisse recht gut interpretieren kann; das macht fast jedes mir bekannte Lehrbuch so und warum sollte ich mich dagegen sträuben?

Beispiel 7.10:
Ich würfle mit einem Standardwürfel und nehme als Bedingung B das „Würfeln einer ungeraden Zahl". Es ist also $p(B) = 1/2$. Nun betrachte ich drei verschiedene Würfelereignisse A_1, A_2, A_3 und berechne deren Wahrscheinlichkeit sowie die Wahrscheinlichkeit dieser Ereignisse unter der Bedingung B.

(a) A_1 sei das Ereignis „Wurf von $1, 2$ oder 3". Dann ist $A_1 \cap B$ = „Wurf einer 1 oder 3" und damit

$$p(A_1) = \frac{1}{2} \text{ und } p(A_1|B) = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3} .$$

Weiß man also, dass eine ungerade Zahl geworfen wird (Ereignis B), so steigt die Wahrscheinlichkeit des Ereignisses A_1 an.

(b) A_2 sei das Ereignis „Wurf von 2, 3 oder 4". Dann ist $A_2 \cap B$= „Wurf einer 3" und damit

$$p(A_2) = \frac{1}{2} \text{ und } p(A_2|B) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3} .$$

Weiß man also, dass eine ungerade Zahl geworfen wird, so sinkt die Wahrscheinlichkeit des Ereignisses A_2 ab.

(c) A_3 sei das Ereignis „Wurf von 5 oder 6". Dann ist $A_3 \cap B$= „Wurf einer 5" und damit

$$p(A_3) = \frac{1}{3} \text{ und } p(A_3|B) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3} .$$

Weiß man also, dass eine ungerade Zahl geworfen wird, so beeinflusst dies die Wahrscheinlichkeit des Ereignisses A_3 nicht. ■

Um die Richtigkeit oder Plausibilität einer Formel zu testen, sollte man immer spezielle Fälle ansehen; im Falle der Formel (7.5) will ich daher zum Schluss zwei solcher Spezialfälle testen:

(a) Sind A und B unvereinbare Ereignisse, so ist $A \cap B$ unmöglich und somit $p(A \cap B) = 0$. Damit ist aber auch $p(A|B) = 0$, und das ist auch gut so, denn dies ist ja die Wahrscheinlichkeit dafür, dass A eintritt unter der Bedingung, dass B bereits eingetreten ist; sind aber wie angenommen A und B unvereinbar, dann *kann* A nicht mehr eintreten, falls B bereits eingetreten ist.

(b) Ist $A = B$, so ist auch $A \cap B = B$, und somit $p(A|B) = 1$. Auch das ist gut so, denn wenn ich die Bedingung stelle, dass B eintritt und $A = B$ ist, so muss natürlich auch A eintreten.

Um den Umgang mit der bedingten Wahrscheinlichkeit noch ein klein wenig zu üben, möchte ich Sie herzlich bitten, die letzte Übungsaufgabe noch zu bearbeiten.

Übungsaufgabe 7.11:
Ein Betrieb erhält seine Zulieferteile von zwei verschiedenen Firmen, und zwar 75 % von Firma A und 25 % von Firma B. Eine Stichprobe hat ergeben, dass die Wahrscheinlichkeit dafür, dass ein zufällig entnommenes Teil intakt ist und von Firma B stammt, $p = 0,23$ beträgt. Wie viel Prozent der von Firma B gelieferten Teile sind intakt? ■

8 Komplexe Zahlen

Ein wichtiges Thema der ersten Kapitel war das Wurzelziehen, es wurden viele Beispiele und Übungsaufgaben gemacht, und so irgendwie wurde Ihnen dabei wohl auch das Gefühl vermittelt, Sie könnten nun schon Wurzeln aus allen reellen Zahlen berechnen.

Das war gelogen.

Na ja, zumindest war es, wenn es so herübergekommen sein sollte, nicht die volle Wahrheit. Die kommt jetzt: Bis hierhin können Sie Wurzeln aus allen *positiven* reellen Zahlen berechnen und natürlich auch aus der Null, denn $\sqrt{0} = 0$. Wurzeln aus *negativen* Zahlen können Sie aber noch nicht berechnen, und das liegt nicht an Ihnen, sondern an der Tatsache, dass es im Reellen einfach keine Wurzeln aus negativen Zahlen gibt: Es existiert zum Beispiel keine reelle Zahl x mit der Eigenschaft

$$x \cdot x \;=\; -1\ .$$

Um so etwas zu realisieren, müssen wir uns in den Bereich der komplexen Zahlen begeben, und ich bitte Sie, mir nun dahin zu folgen; keine Angst, ich bin bei Ihnen – falls das ein Trost für Sie ist.

Falls nicht, tröstet Sie vielleicht das Folgende: Ich war vor einigen Wochen mit meinen Kindern in einem Freizeitpark in Süddeutschland und musste bei dieser Gelegenheit auch mit der „höchsten Achterbahn Europas“ (Aussage des Betreibers, dem ich aufs Wort glauben will) fahren („Papa, du hast doch keine Angst, oder?“).

Nach dieser Erfahrung haben komplexe Zahlen nichts Erschreckendes mehr an sich, glauben Sie mir!

8.1 Die imaginäre Einheit i und die Menge der komplexen Zahlen

Die Einführung der komplexen Zahlen steht und fällt mit der Einführung einer Zahl, die mit sich selbst multipliziert -1 ergibt. Im Reellen existiert so etwas nicht, das hatte ich in der Einleitung nochmals herausgestellt. Auf der anderen Seite wäre es sehr schön, so etwas zu haben, denn schon die Lösung der einfach aussehenden Gleichung

$$x^2 + 1 \;=\; 0$$

erfordert die Existenz einer solchen Zahl.

G. Walz et al., *Brückenkurs Mathematik*

An dieser Stelle machen es die Mathematiker auch nicht anders als – beispielsweise – die Geisteswissenschaftler: Wenn etwas gebraucht wird, was noch nicht existiert, wird es eben definiert:

Imaginäre Einheit

Als **imaginäre Einheit** bezeichnet man diejenige (im Reellen nicht vorkommende, daher „imaginäre“) Zahl i, die die Eigenschaft

$$i^2 = -1$$

hat. Mit anderen Worten: es ist

$$i = \sqrt{-1}\,.$$

Wenn Ihnen diese Definition nicht ganz geheuer ist, so kann ich das nachvollziehen; ich kann Sie jedoch beruhigen und Ihnen versichern, dass das Rechnen mit dieser neu definierten Zahl i zu korrekten und – beispielsweise bei der Lösung von Differenzialgleichungen – mit der wirklichen Welt übereinstimmenden Ergebnissen führen wird.

Aber so weit sind wir noch nicht. Vielleicht denken Sie jetzt: Na fein, nun können wir die Wurzel aus -1 berechnen, aber was ist mit Wurzel aus -2, aus -3 usw.? Sollen wir für jede negative Zahl einen neuen Buchstaben einführen und die Lösung neu benennen?

Das ist sicherlich nicht nötig, als erste Anwendung der neuen Zahl i zeige ich Ihnen jetzt, dass man die Wurzel aus *jeder* negativen Zahl berechnen kann, wenn man nur diejenige aus -1 beherrscht: Stehen Sie beispielsweise vor der Aufgabe, $\sqrt{-9}$ zu berechnen, so können Sie folgende Gleichungskette aufstellen:

$$\sqrt{-9} = \sqrt{(-1)\cdot 9} = \sqrt{-1}\cdot\sqrt{9} = i\cdot\sqrt{9} = 3i\,.$$

Damit haben Sie also die Wurzel aus -9 ermittelt; und da dieses Vorgehen offenbar nicht von der speziellen Wahl -9 abhängt, wissen Sie jetzt, wie man die Wurzel aus einer beliebigen negativen Zahl $-a$ berechnet: Es ist

$$\sqrt{-a} = \sqrt{(-1)\cdot a} = \sqrt{-1}\cdot\sqrt{a} = i\cdot\sqrt{a}.$$

Überhaupt ist die Zahl i der Schlüssel zur Einführung eines neuen, über die bekannten reellen Zahlen hinausgehenden Zahlenbereichs, der Menge der komplexen Zahlen:

Komplexe Zahlen

Ist i die oben definierte komplexe Einheit und sind a und b beliebige reelle Zahlen, so nennt man eine Zahl der Form

$$a + ib$$

eine **komplexe Zahl**.

Die Menge aller komplexen Zahlen bezeichnet man mit $\mathbb{C}$, also

$$\mathbb{C} \ = \ \{a + bi; \ a, b \in \mathbb{R}, i = \sqrt{-1}\} \ .$$

Die Zahl a nennt man den **Realteil**, die Zahl b den **Imaginärteil** der komplexen Zahl $a + ib$.

Beispiele komplexer Zahlen sind $1 + 2i$, $-\sqrt{3} - i$ und $\pi + \pi i$. Aber auch i selbst ist eine komplexe Zahl, denn es lässt sich in der Form $0 + 1i$ schreiben; und schließlich ist jede reelle Zahl x auch eine komplexe, denn man kann sie in der Form $x + 0i$ schreiben. Die komplexen Zahlen stellen also eine Erweiterung der reellen dar, ebenso wie die reellen Zahlen ihrerseits eine Erweiterung der rationalen und die rationalen eine der ganzen Zahlen waren.

Im nächsten Abschnitt zeige ich Ihnen, wie man mit komplexen Zahlen rechnet. Vorher jedoch ... na ja, Sie wissen schon:

Übungsaufgabe 8.1:
Welche der folgenden Ausdrücke sind komplexe Zahlen?
a) $-2 - 3i$,
b) i^2 ,
c) die Lösungen x der Gleichung $x^2 + 2 \ = \ 0$. ■

8.2 Grundrechenarten für komplexe Zahlen

In diesem Abschnitt will ich Ihnen zeigen, wie man komplexe Zahlen addiert, subtrahiert, multipliziert und dividiert, und beginne dabei natürlich mit den beiden einfachsten Rechenarten, der Addition und der Subtraktion.

Sind z_1 und z_2 komplexe Zahlen, dann haben sie nach Definition die Darstellung $z_1 = a_1 + ib_1$ und $z_2 = a_2 + ib_2$. Die Summe $z_1 + z_2$ ist demnach einfach gleich

$$z_1 + z_2 \ = \ (a_1 + ib_1) + (a_2 + ib_2) \, ,$$

wobei ich mathematisch unnötige Klammern gesetzt habe, um die Zusammensetzung dieses Ausdrucks zu verdeutlichen. In dieser Form ist das Ergebnis zwar noch nicht so recht als komplexe Zahl erkennbar, aber natürlich kann man es umstellen, um dies zu erreichen: Es ist

$$z_1 + z_2 \ = \ (a_1 + ib_1) + (a_2 + ib_2) \ = \ (a_1 + a_2) + i \cdot (b_1 + b_2)\,,$$

also eine komplexe Zahl reinsten Wassers.

Damit hätten wir die Addition von komplexen Zahlen bereits geklärt (was nicht schwer war) und die Subtrakion ist auch nicht schlimmer: Ersetzt man das Pluszeichen zwischen z_1 und z_2 durch ein Minuszeichen, so ergibt sich

$$z_1 - z_2 \ = \ (a_1 + ib_1) - (a_2 + ib_2) \ = \ (a_1 - a_2) + i \cdot (b_1 - b_2)\,,$$

auch kein Problem.

So ist also beispielsweise

$$(3 - 2i) + (1 + 4i) \ = \ 4 + 2i$$

und

$$(-2 + 4i) - (-3 + 3i) \ = \ 1 + i\ .$$

Mit Übungsaufgaben zu diesem Thema will ich Ihnen gar nicht erst kommen, aus dem Stadium sind wir raus. Gehen wir stattdessen zur nächsthöheren Grundrechenart über, dem Multiplizieren: Was ist das Produkt zweier komplexer Zahlen? Auch hier muss man keine Scheu haben und auch keine Tricks anwenden, man multipliziert einfach aus und schleppt dabei die ominöse Zahl i mit, „vergisst“ also sozusagen, was für ein unvorstellbares Objekt sich dahinter verbirgt, und rechnet munter drauf los: Einfaches Ausmultiplizieren ergibt

$$z_1 \cdot z_2 \ = \ (a_1 + ib_1) \cdot (a_2 + ib_2) \ = a_1a_2 + ib_1a_2 + ia_1b_2 + i^2b_1b_2\ . \qquad (8.1)$$

Das sieht noch nicht sehr nach einer komplexen Zahl aus, aber das bekommen wir hin: Jetzt „erinnert“ man sich daran, dass ja i nicht irgendein Parameter, sondern dass $i^2 = -1$ ist. Setzt man dies in (8.1) ein, so erhält man

$$z_1 \cdot z_2 \ = \ a_1a_2 - b_1b_2 + i(b_1a_2 + a_1b_2)\,, \qquad (8.2)$$

wobei ich gleich noch den Faktor i bei den beiden gemischten Termen ausgeklammert habe.

So multipliziert man also komplexe Zahlen. Es bleibt Ihrem persönlichen Geschmack überlassen, ob Sie sich die Formel (8.2) merken wollen oder ob Sie im konkreten Fall den gerade vorgeführten Weg (also ausmultiplizieren und dann $i^2 = -1$ setzen) nehmen wollen – das Ergebnis ist dasselbe.

Als kleines Beispiel berechne ich das Produkt

$$(3 - 2i) \cdot (1 + 4i) \ = \ (3 + 8) + i(-2 + 12) \ = \ 11 + 10i\ .$$

Das Beste habe ich mir bis zum Schluss aufgehoben: Die Division komplexer Zahlen, also die vierte Grundrechenart. Hier tasten wir uns ganz langsam an das Problem heran und sehen zunächst einmal, wie man den Kehrwert einer komplexen Zahl berechnet, was also

$$\frac{1}{a+ib}$$

ist. Um diese komplexe Zahl zu berechnen, gibt es einen einfachen Trick, den ich Ihnen jetzt einmal ohne allzu viel Herumgerede zeigen will: Man erweitert den Bruch zunächst mit der Zahl $a - ib$ und erhält somit

$$\frac{1}{a+ib} = \frac{a-ib}{(a+ib)(a-ib)} .$$

Wenn Sie jetzt denken, dass dadurch alles nur noch schlimmer geworden sei, so kann ich Sie verstehen, darf Ihnen aber versichern, dass sich alles gleich zum Guten wenden wird: Multipliziert man nämlich nun den Nenner aus, so erhält man

$$\frac{1}{a+ib} = \frac{a-ib}{(a+ib)(a-ib)} = \frac{a-ib}{a^2+b^2} \tag{8.3}$$

und dieser Nenner ist eine *positive reelle Zahl*. Durch eine solche kann man aber immer dividieren, und somit haben wir das Problem gelöst. Zerlegt man nämlich nun noch den letzten Bruch in (8.3) in zwei Teile, so ergibt sich eine komplexe Zahl reinsten Wassers, die nach Herleitung das Ergebnis der Division ist. Als Merkregel:

Kehrwert einer komplexen Zahl

Für eine beliebige komplexe Zahl $z = a + ib \neq 0$ ist

$$\frac{1}{z} = \frac{1}{a+ib} = \frac{a}{a^2+b^2} - i \cdot \frac{b}{a^2+b^2} .$$

So ist zum Beispiel

$$\frac{1}{2+3i} = \frac{2}{13} - i \cdot \frac{3}{13}$$

und

$$\frac{1}{1-i} = \frac{1}{2} + i \cdot \frac{1}{2} .$$

Wenn Sie mir das nicht glauben – wozu ich Ihnen ehrlich gesagt immer raten würde – so können Sie die Ergebnisse kontrollieren, indem Sie den Wert im Nenner auf der linken Seite mit dem Ergebnis auf der rechten Seite multiplizieren; als Ergebnis muss 1 herauskommen.

Im ersten Beispiel stimmt das auch, denn

$$(2+3i)\left(\frac{2}{13} - i \cdot \frac{3}{13}\right) = \frac{4}{13} + \frac{6i}{13} - \frac{6i}{13} - i^2 \cdot \frac{9}{13} = \frac{13}{13} = 1 ,$$

da, wie inzwischen hinlänglich bekannt sein dürfte, $-i^2 = 1$ ist. Das zweite Beispiel testen Sie bitte nach derselben Methode selbst.

Bisher können wir ja nur Kehrwerte von komplexen Zahlen bilden, wie sieht es aber mit der allgemeinen Division aus; wie also berechnet man Werte der Form

$$\frac{c+id}{a+ib}\ ?$$

Nun, man macht das ganz genau so, wie gerade bei der Kehrwertbildung gezeigt (weshalb ich dies auch getan habe), man erweitert also mit der Zahl $a - ib$. Das führt zunächst auf

$$\frac{c+id}{a+ib} = \frac{(c+id)(a-ib)}{(a+ib)(a-ib)} = \frac{(c+id)(a-ib)}{a^2+b^2},$$

und führt man die Multiplikation im Zähler noch aus, so erhält man das gewünschte Ergebnis, das ich gleich in einem Regelkasten formulieren will:

Division komplexer Zahlen

Sind $z_1 = c + id$ und $z_2 = a + ib$ beliebige komplexe Zahlen mit $z_2 \neq 0$, so ist

$$\frac{z_1}{z_2} = \frac{ac+bd+i(ad-bc)}{a^2+b^2}. \tag{8.4}$$

Damit hätten wir die vier Grundrechenarten erledigt; was jetzt kommt, ist klar:

Übungsaufgabe 8.2:
Gegeben seien die komplexen Zahlen $z_1 = 2 + 5i$, $z_2 = -1 - 2i$ und $z_3 = 1 - 3i$. Berechnen Sie die folgenden komplexen Zahlen:

$$z_1 \cdot z_2\ , \quad \frac{z_1}{z_2 \cdot z_3}\ , \quad \frac{z_1+z_2}{z_2-z_3}\ .$$

■

8.3 Die Gauß'sche Zahlenebene und die trigonometrische Form komplexer Zahlen

Vielleicht haben Sie sich ja schon die ganze Zeit über gefragt, wo man diese neue Zahlenmenge, eben die komplexen Zahlen, überhaupt noch auf der Zahlengeraden unterbringen soll. Zur Erinnerung: Die Zahlengerade ist die Menge aller reellen Zahlen und umgekehrt hat jede reelle Zahl ihren Platz auf der Zahlengeraden. Mit anderen Worten: Die Zahlengerade ist proppenvoll mit reellen Zahlen und für andere Dinge, beispielsweise die komplexen Zahlen, ist kein Platz mehr.

Das ist auch völlig richtig und dennoch kein Problem: Es ist eine ebenso einfache wie geniale Idee (wobei ich persönlich glaube, dass *wirklich* geniale Ideen immer einfach sind, allerdings nicht unbedingt umgekehrt), die auf den großen C. F. Gauß zurückgeht, hier sozusagen eine zweite Dimension zu eröffnen und die Menge der komplexen Zahlen in einer Ebene, der so genannten **Gauß'schen Zahlenebene**, darzustellen.

Man interpretiert hierfür Real- und Imaginärteil der jeweiligen Zahl als ihre beiden Koordinaten in der Ebene, zeichnet also die Zahl $a + ib$ an die Stelle (a, b) im Koordinatensystem ein. In Abbildung 8.1 sehen Sie drei komplexe Zahlen und ihre Position in der Gauß'schen Ebene.

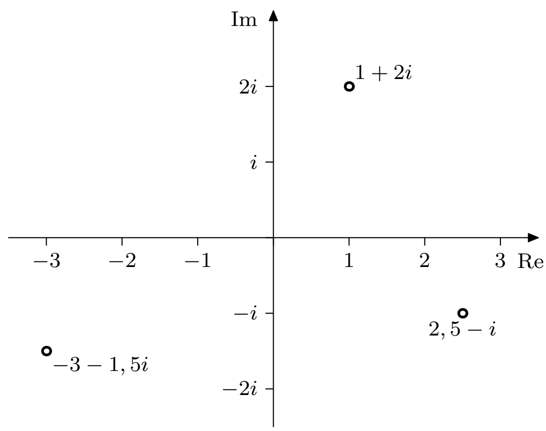

Abb. 8.1 Einige komplexe Zahlen in der Gauß'schen Zahlenebene

Wo ist denn nun eigentlich die reelle Zahlengerade hingeraten? Nun, die Menge der reellen Zahlen ist ja gerade die Menge aller komplexen Zahlen mit $b = 0$; dementsprechend ist die gute alte reelle Zahlengerade als waagrechte Koordinatenachse in der Gauß'schen Zahlenebene wiederzufinden.

Die Gauß'sche Zahlenebene dient aber natürlich nicht nur zum bloßen Hinmalen der komplexen Zahlen, sondern kann auch zum „grafischen Rechnen" verwendet werden, insbesondere zum Wurzelziehen aus komplexen Zahlen, was ansonsten ein ziemlich schwieriges Unterfangen wäre. Hierzu braucht man eine andere Darstellung komplexer Zahlen, die so genannte trigonometrische Form. Diese beruht auf der Tatsache, dass man einen Punkt in der Ebene (und als einen solchen betrachtet man ja jetzt jede komplexe Zahl) nicht nur durch seine kartesischen Koordinaten (also die Werte a und b) eindeutig beschreiben kann, sondern auch dadurch, dass man angibt, wie weit er vom Nullpunkt weg ist und welchen Winkel er – beispielsweise – mit der positiven reellen Achse einschließt.

Das war ein ziemliches Satzungetüm, eines von der Sorte „grammatikalisch korrekt, aber in keiner Weise verständlich". Ich will das daher lieber gleich noch einmal in mathematisch-präziser Form angeben:

Den Abstand einer komplexen Zahl vom Nullpunkt der Gauß'schen Zahlenebene bezeichnet man auch als ihren Betrag, denn man kann diesen Abstand auch interpretieren als den Betrag (also die Länge) des Vektors, der vom Nullpunkt zu dieser Zahl zeigt.

Wie dem auch sei, diesen Betrag berechnet man auf alle Fälle gemäß dem Satz des Kollegen Pythagoras wie folgt:

Betrag einer komplexen Zahl

Der **Betrag** der komplexen Zahl $z = a + ib$ ist die reelle Zahl

$$|z| = \sqrt{a^2 + b^2} .$$

Beispielsweise ist

$$|2 + 3i| = \sqrt{2^2 + 3^2} = \sqrt{13}$$

und

$$|-3 - 4i| = \sqrt{(-3)^2 + (-4)^2} = \sqrt{9 + 16} = 5 .$$

Wie berechnet man nun den Winkel, den eine komplexe Zahl, genauer gesagt der gerade erwähnte Vektor, mit der positiven Achse einschließt? Hierzu schauen wir uns am besten einmal die komplexe Zahl $z = a + ib$ in Abbildung 8.2 an: Sie sehen, dass der Winkel φ als spitzer Winkel in einem rechtwinkligen Dreieck gedeutet werden kann, dessen Gegenkathete b und dessen Ankathete a ist. Folglich ist der Tangens dieses Winkels gleich

$$\tan\varphi = \frac{\text{Gegenkathete}}{\text{Ankathete}} = \frac{b}{a} .$$

Wenn Sie mit dem Tangens nicht mehr oder noch nicht so sehr vertraut sind, dann vertrauen Sie mir doch bitte einfach und akzeptieren Sie die gerade gemachte Aussage. Sie ist richtig. Allerdings bin ich ja nicht am Tangens von φ interessiert, sondern an φ selbst. Diesen Winkel bekomme ich aber nun ganz einfach, indem ich auf beide Seiten die Umkehrfunktion des Tanges anwende, das ist der Arcustangens. Somit ist

$$\varphi = \arctan\left(\frac{b}{a}\right) .$$

Falls Sie auf Ihrem Taschenrechner übrigens keine Taste mit der Inschrift „arctan" entdecken, so schauen Sie einmal nach „inv tan" oder „$\tan^{-1}$", dies alles bedeutet das Gleiche, nämlich die Umkehrfunktion des Tangens, eben den Arcustangens.

Damit hätten wir also die Berechnung des Winkels geklärt, allerdings nur, falls die betreffende komplexe Zahl im ersten Quadranten ist. In allen anderen Fällen muss

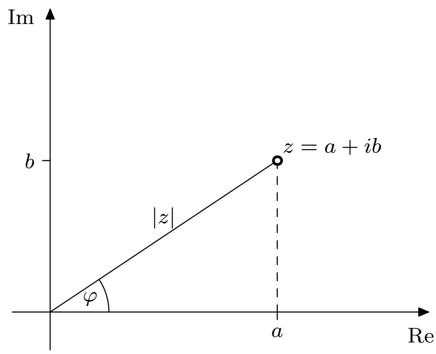

Abb. 8.2 Komplexe Zahl und zugehöriger Winkel

man noch gewisse Winkelzahlen addieren, deren Herleitung hier einfach zu weit führen würde; ich liste daher die Berechnung dieser Winkel kommentarlos in einer kleinen Tabelle auf:

Winkel einer komplexen Zahl

Der Winkel φ, den eine komplexe Zahl $z = a + ib$ mit der positiven reellen Achse einschließt, ist wie folgt zu berechnen:

$$\varphi = \begin{cases} \arctan\left(\frac{b}{a}\right), & \text{falls } a > 0 \text{ und } b \geq 0\,, \\ \arctan\left(\frac{b}{a}\right) + 360°, & \text{falls } a > 0 \text{ und } b < 0\,, \\ \arctan\left(\frac{b}{a}\right) + 180°, & \text{falls } a < 0\,, \\ 90°, & \text{falls } a = 0 \text{ und } b > 0\,, \\ 270°, & \text{falls } a = 0 \text{ und } b < 0\,, \\ 0°, & \text{falls } a = 0 \text{ und } b = 0\,. \end{cases}$$

Beispielsweise ist der Winkel von $2 + 3i$ gleich

$$\varphi = \arctan\frac{3}{2} = 56{,}31°$$

und derjenige von $-3 - 2i$ gleich

$$\varphi = \arctan\frac{-2}{-3} + 180° = \arctan\frac{2}{3} + 180° = 213{,}69°\,.$$

Somit haben wir die beiden Ingredienzien (zugegeben, ich habe mal wieder heimlich im Fremdwörterbuch gelesen) der erwünschten trigonometrischen Form einer komplexen Zahl beisammen und können das nun formal hinschreiben:

Trigonometrische Form einer komplexen Zahl

Als **trigonometrische Form** der komplexen Zahl z bezeichnet man die Darstellung

$$z \;=\; |z| \cdot (\cos\varphi + i\sin\varphi) \;.$$

Dabei ist $|z|$ der Betrag von z und φ der nach obiger Definition berechnete Winkel.

Übrigens nennt man die Darstellungsweise $a + ib$ einer komplexen Zahl auch deren **Normalform**, um sie von der trigonometrischen Form zu unterscheiden.

Auch hierzu natürlich wieder ein paar Beispiele: Ich benutze hierfür die beiden komplexen Zahlen, deren Winkel ich oben bereits ausgerechnet hatte, von denen ich also nur noch den Betrag bestimmen muss (man wird älter!).

Der Betrag von $z = 2 + 3i$ ist

$$|z| \;=\; \sqrt{2^2 + 3^2} \;=\; \sqrt{13} \;\approx\; 3{,}6056 \;,$$

den Winkel hatte ich oben zu $56{,}31°$ berechnet. Somit ist die trigonometrische Form dieser Zahl gleich

$$z \;=\; \sqrt{13} \cdot (\cos 56{,}31° + i\sin 56{,}31°) \;.$$

Zur Kontrolle können Sie Ihren Taschenrechner anwerfen und nachrechnen, dass

$$3{,}6056 \cdot (\cos 56{,}31° + i\sin 56{,}31°) \;=\; 3{,}6056 \cdot (0{,}5547 + i \cdot 0{,}8321) \;=\; 2 + i \cdot 3$$

ist.

Auch der Betrag der zweiten Beispielzahl, $z = -3 - 2i$, ist $\sqrt{13}$; hierfür ergibt sich somit die trigonometrische Form

$$z \;=\; 3{,}6056 \cdot (\cos 213{,}69° + i\sin 213{,}69°) \;.$$

Die Kontrollrechnung vertraue ich diesmal Ihnen an.

Offen gestanden hatte ich mit dem Gedanken gespielt, Sie an dieser Stelle damit zu verblüffen, dass ich *keine* Übungsaufgaben zu diesem Thema stelle. Ich habe mich aber sehr schnell dagegen entschieden, denn gerade die trigonometrische Form ist sehr wichtig und sollte daher unbedingt eingeübt werden.

Übungsaufgabe 8.3:
Bestimmen Sie die trigonometrische Form der folgenden komplexen Zahlen:
a) $z = 1 - i$,
b) $z = -5 - 3i$. ■

8.4 Potenzen und Wurzeln komplexer Zahlen

Im letzten Kapitel hatte ich schon davon gesprochen, dass die trigonometrische Form komplexer Zahlen manche Rechenoperationen einfacher durchführbar bzw. überhaupt erst möglich macht; dies will ich Ihnen jetzt zeigen.

Addition und Subtraktion sind in der Normalform komplexer Zahlen so einfach durchführbar, dass sie keiner weiteren Vereinfachung bedürfen. Etwas anders sieht es schon bei der Multiplikation und der Division aus, beide können in der Normalform etwas aufwendig sein und sind daher für Vereinfachungen immer zu haben. Eine solche bietet die trigonometrische Form:

Multiplikation und Division komplexer Zahlen in trigonometrischer Form

Es seien

$$z_1 = |z_1| \cdot (\cos\varphi_1 + i\sin\varphi_1)$$

und

$$z_2 = |z_2| \cdot (\cos\varphi_2 + i\sin\varphi_2)$$

zwei komplexe Zahlen in trigonometrischer Form.

Dann lautet ihr Produkt

$$z_1 \cdot z_2 = |z_1| \cdot |z_2| \cdot (\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)) \tag{8.5}$$

sowie – falls $z_2 \neq 0$ – ihr Quotient

$$\frac{z_1}{z_2} = \frac{|z_1|}{|z_2|} \cdot (\cos(\varphi_1 - \varphi_2) + i\sin(\varphi_1 - \varphi_2)) \ . \tag{8.6}$$

Die beiden Formeln sind übrigens gar nicht so schwer herzuleiten, man führt einfach die Multiplikation bzw. Division in Normalform durch und benutzt die Additionstheoreme für Sinus und Cosinus, aber das will ich Ihnen und mir hier dennoch ersparen.

Man multipliziert also komplexe Zahlen, indem man ihre Beträge multipliziert (was einfach ist, da es sich um positive reelle Zahlen handelt), und ihre Winkel addiert; Entsprechendes gilt für die Division.

Auch dieses will ich wieder durch ein kleines Beispiel untermauern. Da mir partout keines zu komplexer Pizza-Rechnung oder Ähnlichem einfällt, belasse ich es bei einem simplen Zahlenbeispiel. Ich verwende die beiden weiter oben bereits als Beispiel für die trigonometrische Form benutzten Zahlen

$$z_1 = 2 + 3i = 3{,}6056 \cdot (\cos 56{,}31° + i\sin 56{,}31°)$$

und

$$z_2 = -3 - 2i = 3{,}6056 \cdot (\cos 213{,}69^\circ + i \sin 213{,}69^\circ) .$$

Multiplikation der beiden Beträge und Addition der Winkel ergibt hier

$$z_1 \cdot z_2 = 13 \cdot (\cos 270^\circ + i \sin 270^\circ) ,$$

also

$$z_1 \cdot z_2 = 13 \cdot (0 - i) = -13i ,$$

was man durch direkte Multiplikation der beiden Zahlen gemäß der Formel (8.2) auch bestätigen kann.

Um einen Quotienten der beiden Zahlen zu berechnen, muss ich zunächst die Beträge dividieren; dies ergibt 1, und insgesamt erhalte ich

$$\frac{z_1}{z_2} = 1 \cdot (\cos(-157{,}38^\circ) + i \sin(-157{,}38^\circ)) = -0{,}9230 + i \cdot 0{,}3846 .$$

Na ja, zugegeben, allzu prickelnd war das noch nicht, eine simple Multiplikation oder Division komplexer Zahlen bekommt man auch in der Normalform noch ganz gut hin.

Ein wenig attraktiver wird die trigonometrische Form schon, wenn es sich um wiederholte Multiplikation, sprich also Potenzierung einer komplexen Zahl handelt. Will man beispielsweise $(3 + 7i)^5$ berechnen (keine Ahnung, warum man das tun sollte, aber um Anwendung ihrer Ergebnisse haben sich die Mathematiker noch selten gekümmert, dafür sind andere da), so kann man dies natürlich mithilfe der Formel (8.2) für das Multiplizieren komplexer Zahlen tun, aber man hat schon einige Arbeit damit. Sehr viel eleganter geht das durch wiederholte Anwendung der Formel (8.5) für die Multiplikation in trigonometrischer Form. Dies ergibt:

Potenzierung komplexer Zahlen

Ist

$$z = |z| \cdot (\cos \varphi + i \sin \varphi)$$

eine komplexe Zahl und n eine natürliche Zahl, so ist

$$z^n = |z|^n \cdot (\cos n\varphi + i \sin n\varphi) .$$

Man potenziert also eine komplexe Zahl, indem man ihren Betrag potenziert (hatte ich schon erwähnt, dass das eine einfache Operation ist, da der Betrag eine positive reelle Zahl ist?) und ihren Winkel mit n multipliziert.

Übrigens erhält man als Spezialfall dieser Regel für komplexe Zahlen, deren Betrag gleich 1 ist, die Formel

$$(\cos\varphi + i\sin\varphi)^n = (\cos n\varphi + i\sin n\varphi) ,$$

die man auch die **de Moivre'sche Formel** nennt, benannt nach einem französischen Mathematiker namens – Sie werden es nicht glauben – de Moivre, der von 1667 bis 1754 lebte.

Als erstes Beispiel für diese neue Potenzierungsregel berechne ich die oben schon erwähnte Zahl $(3 + 7i)^5$. Die trigonometrische Form von $3 + 7i$ lautet (rechnen Sie es lieber nach, die nächste Auflage dieses Buches kann noch Verbesserungen vertragen!)

$$3 + 7i = \sqrt{58} \cdot (\cos 66{,}8014° + i\sin 66{,}8014°)$$

und somit ist

$$(3 + 7i)^5 = \sqrt{58}^5 \cdot (\cos 334{,}0070° + i\sin 334{,}0070°) = 23027{,}9907 - 11228{,}0199i .$$

Da man aber ja weiß, dass Real- und Imaginärteil der Zahl $(3+7i)^5$ ganze Zahlen sein müssen, da sie bei Rechnung mit der Normalform durch wiederholte Multiplikation ganzer Zahlen entstehen, runde ich dies lieber gleich zu $23028 - 11228i$ und dieses Ergebnis wiederum kann man durch explizites Ausmultiplizieren bestätigen.

Als zweites Beispiel berechne ich einmal die vierte Potenz von i, also i^4. Die trigonometrische Form von i lautet – streng nach Vorschrift –

$$i = 1 \cdot (\cos 90° + i\sin 90°) ,$$

und somit ist

$$i^4 = 1^4 \cdot (\cos 360° + i\sin 360°) = 1 + i \cdot 0 ,$$

also $i^4 = 1$. Zugegeben, das kann man auch schneller direkt sehen, denn

$$i^4 = (i^2)^2 = (-1)^2 = 1 ,$$

aber damit sollte ja auch nur die Korrektheit der Potenzierungsformel illustriert werden.

Weitere Beispiele folgen nun im Do-it-yourself-Verfahren:

Übungsaufgabe 8.4:
Berechnen Sie die folgenden komplexen Zahlen und geben Sie diese in Normalform an:
a) $(-1 + 2i)^5$,
b) $\left(\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right)^8$. ■

Bisher hatten wir ja schon einige Vorteile der trigonometrischen Form gesehen, aber die wahre Stärke dieses Zugangs zeige ich Ihnen erst jetzt (toll, nicht wahr, erst mal ein paar Seiten hinhalten, und dann erst mit dem großen Vorteil herausrücken! Na ja, das nennt man andernorts „Marketing“.) Was ich sagen will ist Folgendes: Bisher konnte ich Ihnen einige Techniken – sprich Durchführung der Grundrechenarten – zeigen, die in der trigonometrischen Form leichter von der Hand gehen als in der Standardform. Jetzt aber kommt etwas, das durch die trigonometrische Form überhaupt erst möglich, das also in der Normalform praktisch gar nicht durchführbar ist: das Wurzelziehen aus komplexen Zahlen.

Bevor ich Ihnen das genauer näher bringe, hätte ich aber gerne einen Trommelwirbel ...? Nun gut, notfalls mache ich den auch selbst. Denn das, was jetzt folgt, ist so einen Aufwand allemal wert: Während man im Reellen sozusagen nie so genau weiß, wie viele verschiedene n-te Wurzeln es aus einer gegebenen Zahl x gibt – das hängt davon ab, ob n gerade oder ungerade ist und natürlich vor allem davon, ob x positiv oder negativ ist – so gibt es hier im Komplexen eine ganz einfache Regel ohne jede Ausnahme: Ist n irgendeine natürliche Zahl und z irgendeine von null verschiedene komplexe Zahl, dann gibt es genau n verschiedene n-te Wurzeln aus z. Punkt, fertig, aus, keine Ausnahmen nötig und erlaubt!

Aus jeder komplexen Zahl kann man also genau drei verschiedene dritte Wurzeln ziehen, genau sieben verschiedene siebte Wurzeln, genau neunzehn verschiedene neunzehnte Wurzeln und, wenn man unbedingt will, sogar zweihundertdreiundvierzig verschiedene zweihundertdreiundvierzigste Wurzeln.

Und es kommt sogar noch besser: Es gibt sogar eine einfache Formel, um diese Wurzeln zu berechnen. Diese gebe ich nun gleich an:

Wurzeln aus komplexen Zahlen

Es sei z eine beliebige komplexe Zahl ungleich null in der trigonometrischen Form

$$z \;=\; |z| \cdot (\cos\varphi + i \sin\varphi)$$

und n eine beliebige natürliche Zahl.

Es gibt genau n n-te Wurzeln aus z, die ich mit $\zeta_0,\ \zeta_1,\ \ldots,\ \zeta_{n-1}$ bezeichnen will.

Diese Wurzeln berechnet man wie folgt: Es ist

$$\zeta_k \;=\; \sqrt[n]{|z|} \cdot \left(\cos\left(\frac{\varphi + k \cdot 360°}{n}\right) + i \cdot \sin\left(\frac{\varphi + k \cdot 360°}{n}\right)\right) \tag{8.7}$$

für $k = 0,\ 1,\ \ldots,\ n-1$.

ζ ist übrigens der griechische Buchstabe „zeta“, also das griechische „z“; schließlich müssen sich die mühevoll durchgestandenen fünf Jahre Griechisch-Unterricht ja irgendwann einmal auszahlen.

Die gerade angegebene Formel (8.7) ist ein wahres Wunderwerk der Mathematik: Sie „produziert“ ohne weitere Rückfrage und/oder Fallunterscheidung alle möglichen

Wurzeln einer komplexen Zahl. Der kleine Nachteil einer solchen Wunderformel ist, dass man sich erst noch an den Umgang mit ihr gewöhnen muss. Dies macht man am besten mit ein paar Beispielen:

Als Erstes werde ich die dritten Wurzeln aus der Zahl

$$z = 8 \cdot (\cos 21^\circ + i \cdot \sin 21^\circ)$$

berechnen. Da ich an den dritten Wurzeln interessiert bin, ist hier also $n = 3$, und somit durchläuft der Index k die Werte $0, 1, 2$. Freundlicherweise ist z bereits in trigonometrischer Form angegeben, sodass ich den Betrag 8 direkt ablesen kann, und aus dem gleichen Grund ist der Winkel $\varphi = 21^\circ$ unschwer erkennbar.

Den Betrag aller drei dritten Wurzeln erhalte ich, indem ich aus $|z|$, also aus 8, die dritte Wurzel ziehe, was ohne weitere Mühe 2 ergibt. Um die erste der drei dritten Wurzeln, ζ_0, zu bestimmen, muss ich in der obigen Formel nun noch $\varphi = 21^\circ$, $n = 3$ und $k = 0$ setzen. Dies ergibt

$$\zeta_0 = 2 \cdot (\cos 7^\circ + i \cdot \sin 7^\circ)\,,$$

und wenn man dies noch ausrechnet, so erhält man die Normalform

$$\zeta_0 = 2 \cdot (0{,}9925 + i \cdot \sin 0{,}1219) = 1{,}985 + i \cdot 0{,}2438\ .$$

Ebenso berechnet man die beiden weiteren dritten Wurzeln zu

$$\zeta_1 = 2 \cdot (\cos 127^\circ + i \cdot \sin 127^\circ) = -1{,}2036 + 1{,}5972i$$

und

$$\zeta_2 = 2 \cdot (\cos 247^\circ + i \cdot \sin 247^\circ) = -0{,}7815 - 1{,}8410i\ .$$

Dieses erste Beispiel war – eben *weil* es das erste war – ein wenig atypisch, da die Zahl, aus der Wurzeln gezogen werden sollten, schon in trigonometrischer Form gegeben war. Das ist natürlich üblicherweise nicht so, und daher werde ich jetzt ein „ernsthaftes" Beispiel durchrechnen, nämlich die dritten Wurzeln aus

$$-1 + i$$

berechnen. Der Betrag dieser Zahl ist gleich $\sqrt{(-1)^2 + 1^2} = \sqrt{2}$, der Winkel ist

$$\arctan\left(\frac{1}{-1}\right) + 180^\circ = -45^\circ + 180^\circ = 135^\circ\,.$$

Damit ergeben sich die folgenden drei dritten Wurzeln:

$$\begin{aligned}
\zeta_0 &= \sqrt[6]{2} \cdot (\cos 45^\circ + i \sin 45^\circ) = 0{,}7937 + 0{,}7937i\,,\\
\zeta_1 &= \sqrt[6]{2} \cdot (\cos 165^\circ + i \sin 165^\circ) = -1{,}0842 + 0{,}2905i\,,\\
\zeta_2 &= \sqrt[6]{2} \cdot (\cos 285^\circ + i \sin 285^\circ) = 0{,}2905 - 1{,}0842i\,.
\end{aligned}$$

Vielleicht wundern Sie sich über die sechste Wurzel, die im Betrag vorkommt? Nun, nach Vorschrift musste ich die dritte Wurzel aus dem Betrag von $-1+i$ ziehen, dieser wiederum ist die (zweite) Wurzel aus 2, insgesamt ergibt sich also die sechste Wurzel.

Auf die Gefahr hin, dass ich Sie langweile (was übrigens ein gutes Zeichen wäre, denn dann haben Sie's schon verstanden), gebe ich nun noch ein drittes Beispiel: Ich berechne die zweiten Wurzeln (also Quadratwurzeln) aus -1.

Noch irgendwelche Fragen, Watson? Nun ja, vielleicht verblüfft es Sie, dass hier weit und breit kein „i" zu entdecken ist, die Zahl also nicht unbedingt als komplexe Zahl zu identifizieren ist. Dennoch ist sie eine, denn ganz zu Anfang des Kapitels hatten wir uns ja darauf geeinigt, dass die reellen Zahlen als Teilmenge in den komplexen enthalten sind; sie zeichnen sich eben genau dadurch aus, dass ihr Imaginärteil null ist.

Das macht aber überhaupt nichts, ich hatte oben behauptet, dass die Wurzelberechnungsformel ohne jede Fallunterscheidung und Ausnahme funktioniert, und das zeige ich Ihnen jetzt am Beispiel $z = -1$.

Der Betrag dieser Zahl ist 1, das sieht man entweder mit bloßem Auge oder rechnet es streng nach der Formel $\sqrt{(-1)^2 + 0}$ aus. Ähnlich ist es mit dem Winkel: Da die Zahl -1 auf der negativen reellen Achse liegt, schließt sie mit der positiven Achse natürlich den Winkel 180° ein; sieht man das nicht direkt ein, so kann man natürlich auch die obige Tabelle verwenden. In jedem Fall erhält man die trigonometrische Form

$$-1 \;=\; 1 \cdot (\cos 180^\circ + i \sin 180^\circ)\ .$$

Nun geht's los: Um den Betrag brauche ich mich offenbar nicht zu kümmern, der ist und bleibt 1, es geht lediglich um die Winkel: Für $k = 0$ erhalte ich

$$\zeta_0 \;=\; \cos\frac{180^\circ}{2} + i\sin\frac{180^\circ}{2} \;=\; \cos 90^\circ + i \sin 90^\circ \;=\; 0 + i \;=\; i\,.$$

Die erste Quadratwurzel aus -1 ist also i. Na toll, das sollte niemanden überraschen, so war die imaginäre Einheit i ja gerade definiert worden. Immerhin, wir haben bestätigt, dass die Wurzelberechnungsformel auch in diesem Fall funktioniert.

Wie lautet nun aber die zweite Quadratwurzel aus -1? Auch diese kann man natürlich erraten, aber man kann sie auch errechnen, und genau das will ich nun auch tun: Für $k = 1$ (und das ist auch schon der letzte Wert, denn für $n = 2$ ist eben $n - 1 = 1$) ergibt die Formel

$$\zeta_1 \;=\; \cos\frac{540^\circ}{2} + i\sin\frac{540^\circ}{2} \;=\; \cos 270^\circ + i \sin 270^\circ \;=\; 0 + i(-1) \;=\; -i\,,$$

wobei man diese Werte von Sinus und Cosinus jeder Formelsammlung oder Tabelle entnehmen kann.

Die zweite Quadratwurzel aus -1 ist also $-i$, und das kann man auch direkt verifizieren, denn

$$(-i)^2 = (-1)^2 \cdot i^2 = -1 \; .$$

Übungsaufgabe 8.5:
Berechnen Sie
a) alle zweiten Wurzeln aus $-2 + 3i$,
b) alle dritten Wurzeln aus 8,
und geben Sie das Ergebnis in Normalform an. ■

8.5 Vollständige Lösung quadratischer und biquadratischer Gleichungen

Vermutlich schon in Ihrer Schulzeit, spätestens jedoch im dritten Kapitel dieses Buches haben Sie gesehen, wie man die Lösungen einer quadratischen Gleichung berechnet. Zur Erinnerung: Die beiden Lösungen der quadratischen Gleichung

$$x^2 + px + q = 0$$

sind die Zahlen

$$x_1 = -\frac{p}{2} + \sqrt{\frac{p^2}{4} - q}$$

und

$$x_2 = -\frac{p}{2} - \sqrt{\frac{p^2}{4} - q}\,,$$

wobei bisher, da wir es nur mit reellen Zahlen zu tun hatten, stets der Zusatz kam, *falls der Radikand* $\frac{p^2}{4} - q$ *nicht negativ ist, ansonsten hat die Gleichung keine Lösung.*

Genau dieser Zusatz kann jetzt entfallen, da inzwischen klar ist, wie man Wurzeln aus negativen Zahlen berechnet; ganz zu Anfang des Kapitels habe ich herausgestellt, dass die Wurzel aus einer negativen Zahl $-a$ wie folgt berechnet werden kann:

$$\sqrt{-a} = \sqrt{(-1) \cdot a} = \sqrt{-1} \cdot \sqrt{a} = i \cdot \sqrt{a}.$$

Wendet man dies auf den oben genannten Fall der p-q-Formel an, so erhält man die folgenden Lösungsformel quadratischer Gleichungen:

Vollständige Lösung quadratischer Gleichungen

Die quadratische Gleichung

$$x^2 + px + q = 0$$

hat die Lösungen

$$x_1 = -\frac{p}{2} + \sqrt{\frac{p^2}{4} - q}$$

und

$$x_2 = -\frac{p}{2} - \sqrt{\frac{p^2}{4} - q} \, ,$$

falls der Radikand $\frac{p^2}{4} - q$ nicht negativ ist. Andernfalls sind die Lösungen die komplexen Zahlen

$$x_1 = -\frac{p}{2} + i \cdot \sqrt{-\left(\frac{p^2}{4} - q\right)}$$

und

$$x_2 = -\frac{p}{2} - i \cdot \sqrt{-\left(\frac{p^2}{4} - q\right)} \, .$$

Als Beispiel bestimme ich die Lösungen der Gleichung

$$x^2 - 2x + 2 = 0 \, . \tag{8.8}$$

Hier ist $p = -2$ und $q = 2$, also

$$\frac{p^2}{4} - q = 1 - 2 = -1 \, .$$

Der Radikand ist also negativ und somit sind die beiden Lösungen der Gleichung (8.8) nach der zweiten Möglichkeit zu berechnen. Es ergibt sich

$$x_1 = 1 + i$$

und

$$x_2 = 1 - i \, .$$

Zur Kontrolle kann man diese Lösungen nun wieder in die Gleichung (8.8) einsetzen; so erhält man z. B.

$$x_1^2 - 2x_1 + 2 = (1+i)^2 - 2(1+i) + 2 = 1 + 2i + i^2 - 2 - 2i + 2 = 0 \, .$$

Den Test der zweiten Lösung x_2 übernehmen Sie bitte selbst.

Ganz ähnliche Bemerkungen wie für quadratische Gleichungen gelten auch für biquadratische, also solche der Form

$$x^4 + bx^2 + c \;=\; 0\,,$$

die Sie im dritten Kapitel kennen gelernt haben. Dort hatten Sie auch gesehen, dass eine solche Gleichung im Reellen bis zu vier Lösungen haben kann, aber eben nur dann, wenn gewisse Zahlen nicht negativ sind. Genau dieser „wenn"-Zusatz kann nun im Komplexen entfallen: Hier hat *jede* biquadratische Gleichung vier Lösungen, und die können nach folgendem Verfahren berechnet werden:

Vollständige Lösung biquadratischer Gleichungen

Um die Lösungen der biquadratischen Gleichung

$$x^4 + bx^2 + c \;=\; 0$$

zu bestimmen, führt man die Substitution $u = x^2$ durch und berechnet die Lösungen u_1 und u_2 der quadratischen Gleichung

$$u^2 + bu + c \;=\; 0$$

wie oben gezeigt. Dann berechnet man die Zahlen

$$x_{11} = \sqrt{u_1} \quad \text{und} \quad x_{12} = -\sqrt{u_1}$$

sowie ebenso

$$x_{21} = \sqrt{u_2} \quad \text{und} \quad x_{22} = -\sqrt{u_2}\ .$$

Die Zahlen x_{11}, x_{12}, x_{21} und x_{22} sind Lösungen der biquadratischen Gleichung.

Um diese Berechnungsweise zu illustrieren, rechne ich das letzte Beispiel zu biquadratischen Gleichungen aus dem dritten Kapitel, das dort keine (reelle) Lösung hatte, mit diesem neu erworbenen Wissen über komplexe Zahlen nochmals durch: Es handelt sich um die Gleichung

$$x^4 + 2x^2 + 4 \;=\; 0\ . \tag{8.9}$$

Die Substitution $u = x^2$ führt hier auf die quadratische Gleichung

$$u^2 + 2u + 4 \;=\; 0$$

mit den beiden Lösungen

$$u_{1/2} \;=\; -1 \pm \sqrt{1-4} \;=\; -1 \pm i\sqrt{3}\,.$$

Um nun die vier Lösungen der Ausgangsgleichung (8.9) zu erhalten, muss ich aus u_1 und u_2 jeweils die beiden Quadratwurzeln ziehen, und das heißt zunächst, die trigonometrische Form bestimmen.

Ich beginne mit $u_1 = -1 + i\sqrt{3}$. Der Betrag dieser Zahl ist $|u_1| = \sqrt{1+3} = 2$, der Winkel lautet

$$\varphi = \arctan\left(\frac{\sqrt{3}}{-1}\right) + 180° = -60° + 180° = 120°.$$

Damit erhalte ich als Wurzeln hieraus und somit als Lösungen der biquadratischen Gleichung (8.9):

$$x_{11} \;=\; \sqrt{2} \cdot (\cos 60° + i \sin 60°) \;=\; 0{,}7071 + i \cdot 1{,}2247$$

und

$$x_{12} \;=\; \sqrt{2} \cdot (\cos 240° + i \sin 240°) \;=\; -0{,}7071 - i \cdot 1{,}2247\,.$$

Die beiden Wurzeln aus $u_2 = -1 - i\sqrt{3}$ kann man auf dieselbe Art und Weise berechnen (womit ich elegant zu Ihrer letzten Übungsphase überleite) und erhält schließlich

$$x_{21} \;=\; \sqrt{2} \cdot (\cos 120° + i \sin 120°) \;=\; -0{,}7071 + i \cdot 1{,}2247$$

sowie

$$x_{22} \;=\; \sqrt{2} \cdot (\cos 300° + i \sin 300°) \;=\; 0{,}7071 - i \cdot 1{,}2247\,.$$

Und noch ein letztes Mal bitte ich Sie dringend, diese Ergebnisse nachzurechnen, einerseits, damit Sie diese Techniken üben, aber auch, weil ich mich leider andauernd verrechne – weshalb ich ja auch Professor geworden bin, als einfacher Mathematiker hätte ich im Leben keine Chance.

Übungsaufgabe 8.6:
Bestimmen Sie alle Lösungen der folgenden Gleichungen:
a) $x^2 + x + \frac{5}{2} \;=\; 0,$
b) $x^4 - x^2 - 2 \;=\; 0.$ ■

Ganz am Ende eines solchen Buches ist es immer eine gute Idee, nochmals den ganzen Verlauf Revue passieren zu lassen. Am Anfang hatten Sie sich – neben Pizza und Rotwein – mit den Grundrechenarten herumzuplagen, Sie haben gelernt, was Terme und Funktionen sind, Sie können jetzt auch kompliziertere Gleichungen lösen und mit Ungleichungen umgehen, geometrische Grundstrukturen sind Ihnen nicht mehr fremd, das Ableiten wie auch das Integrieren von Funktionen sind Ihnen vertraut geworden, sogar ganze Gleichungssysteme können Sie lösen und damit lineare geometrische Objekte im Raum beschreiben und auch die Grundlagen der Wahrscheinlichkeitsrechnung lernten Sie kennen. Ebenso ist Ihnen der Umgang mit komplexen Zahlen, für die meisten Mathematik-Einsteiger ein echter Stolperstein, nunmehr vertraut.

Finden Sie nicht auch, dass das eine ganz ansehnliche Leistung ist? Also, ich schon, und mir können Sie in dieser Hinsicht vertrauen. Wenn Sie dieses Buch durchgearbeitet und die Inhalte großenteils verstanden haben, braucht Ihnen vor keiner Art Mathematik im Studium oder im Leben bange zu sein; und wer sonst kann das schon von sich behaupten?

Lösungen der Übungsaufgaben

Zu Ihrer Selbstkontrolle sind hier die Lösungen der Übungsaufgaben angegeben; ausführliche Lösungswege finden Sie auf den Internetseiten des Verlags.

Übungen Kapitel 1

Lösung von Aufgabe 1.1:
Es ist

$$a + b + c = a + (b + c) = (b + c) + a = (c + b) + a = c + b + a$$

Lösung von Aufgabe 1.2:
a) $(3 - 6)(-2 - 3) = (-3)(-5) = 15$
b) $-(8 - 3)(-2 + 5) = (-5)3 = -15$
c) $(9 - 3 - 8)(3 - 1 - 7)(-2 + 1) = (-2)(-5)(-1) = -10$

Lösung von Aufgabe 1.3:
$\frac{231}{22} = \frac{21 \cdot 11}{2 \cdot 11} = \frac{21}{2}$
$\frac{52}{28} = \frac{4 \cdot 13}{4 \cdot 7} = \frac{13}{7}$
$\frac{16}{64} = \frac{1}{4}$

Lösung von Aufgabe 1.4:
$\frac{9}{4} \cdot \frac{6}{3} = \frac{9}{2}$
$\frac{13}{7} \cdot \frac{21}{26} = \frac{3}{2}$
$\frac{52}{76} \cdot \frac{19}{13} = 1$

Lösung von Aufgabe 1.5:

$\frac{3}{4} : \frac{7}{8} = \frac{6}{7}$

$\frac{52}{76} : \frac{13}{19} = 1$

$\frac{143}{11} : \frac{130}{22} = \frac{11}{5}$

Lösung von Aufgabe 1.6:

a) $(((2-3)\cdot 4)-2)\cdot(-2) \; = \; 12$
b) $2-((1-4)\cdot(3-2)+4) \; = \; 1$

Lösung von Aufgabe 1.7:

a) $\left(\frac{1}{2}\right)^3 \cdot \left(\frac{2}{3}\right)^5 \cdot \left(\frac{3}{4}\right)^2 \; = \; \frac{1}{108}$

b) $(a^2b^3c^{-1})^2 \cdot (c^2)^2 \; = \; a^4b^6c^2$

Lösung von Aufgabe 1.8:

a) $\sqrt[4]{81} \; = \; 3$
b) $\sqrt[3]{2^2 \cdot 2} \; = \; 2$
c) $\sqrt[5]{(a^5)^2} \; = \; a^2$

Lösung von Aufgabe 1.9:

a) $4^{\frac{1}{3}} \cdot 2^{\frac{1}{3}} \cdot 16^{\frac{1}{4}} \; = \; 4$
b) $120^{\frac{1}{2}} \cdot 900^{\frac{1}{4}} \; = \; 60$
c) $\sqrt{0,16} \; = \; 0,4$

Lösung von Aufgabe 1.10:

a) $(a^2+b-c)(a+bc)(a^2b) \; = \; a^5b + a^3b^2 - a^3bc + a^4b^2c + a^2b^3c - a^2b^2c^2$
b) $(1+x+x^2+x^3+x^4)(1-x) \; = \; 1-x^5$

Lösung von Aufgabe 1.11:

a) $34xyz^2 - 17x^3yz + 51y^4 \; = \; 17y(2xz^2 - x^3z + 3y^3)$
b) $2a^2bc - 4ab^2c + 8abc^2 \; = \; 2abc(a-2b+4c)$

Lösung von Aufgabe 1.12:

a) $\dfrac{4a^3(bc)^2}{(2abc-4ab)a^2} \;=\; \dfrac{2bc^2}{c-2}$

b) $\dfrac{(3x^2yz)^3}{9(xy^2z)^3} \;=\; 3\left(\dfrac{x}{y}\right)^3$

Lösung von Aufgabe 1.13:

a) $\displaystyle\sum_{n=1}^{33} 3n$

b) $\displaystyle\sum_{n=0}^{10}(2n+1)$

Lösung von Aufgabe 1.14:

$$\sum_{i=-1}^{1}\;\sum_{j=-2i}^{i} i^2\cdot j \;=\; -2$$

Lösung von Aufgabe 1.15:

a) $8! = 40\,320$

b) $9\cdot 10\cdot 11 = \dfrac{11!}{8!}$

Übungen Kapitel 2

Lösung von Aufgabe 2.1:

$f([-1,2]) = [-1,-\dfrac{1}{4}]$

Lösung von Aufgabe 2.2:

a) $f:\mathbb{N}\to\mathbb{N}\ ,\ f(x)=x^2$ ist eine Funktion, denn das Quadrat einer natürlichen Zahl ist wieder eine natürliche Zahl.

b) $g:[-1,1]\to\mathbb{R}\ ,\ g(x)=\frac{1}{x}$ ist keine Funktion, denn für $x=0$ existiert kein Wert.

c) $h:\mathbb{Q}\to\mathbb{Q}\ ,\ h(x)=2x^4-\frac{5}{3}x^3+\frac{1}{2}x$ ist eine Funktion, denn die Werte sind stets rationale Zahlen.

Lösung von Aufgabe 2.3:
Die Verkettung ist in beiden Fällen möglich; im ersten Fall ergibt sich

$$(g \circ f)(x) = (\sqrt{x} + 1)^2 \ ,$$

im zweiten

$$(g \circ f)(x) = \sqrt{\frac{1 + 2x^2}{1 + x^2}} \ .$$

Lösung von Aufgabe 2.4:
Die Funktion f ist monoton steigend, die Funktionen g und h sind monoton fallend.

Lösung von Aufgabe 2.5:
Da alle drei Funktionen streng monoton sind, sind die Umkehrfunktionen definiert; sie lauten

$$\begin{aligned} f^{-1} : \mathbb{R} \to \mathbb{R} \ , \quad f^{-1}(x) &= \ 10x + 170 \, , \\ g^{-1} : \mathbb{R} \to \mathbb{R} \ , \quad g^{-1}(x) &= \ 170 - 17x \, ; \\ h^{-1} : [\frac{1}{17}, 1] \to \mathbb{R} \ , \quad h^{-1}(x) &= \ \frac{1}{x} \, . \end{aligned}$$

Lösung von Aufgabe 2.6:
f und h sind keine Polynome, $g(x) = x + 1$ ist ein Polynom.

Lösung von Aufgabe 2.7:
Der Funktionswert $p(0) = 117$ ist positiv, während beispielsweise die Werte $p(-1\,000)$ und $p(1\,000)$ negativ sind; also muss dazwischen je eine Nullstelle sein.

Lösung von Aufgabe 2.8:
f ist wegen des Terms x^x keine rationale Funktion, g ist eine rationale Funktion mit maximalem Definitionsbereich $\mathbb{R} \setminus \{1\}$.

Lösung von Aufgabe 2.9:
Die Werte sind folgender Tabelle zu entnehmen:

a	x = −2	x = 0	x = 1	x = 2	x = 10
0,5	4	1	0,5	0,25	≈ 0,000 97
1	1	1	1	1	1
2	0,25	1	2	4	1 024
4	0,0625	1	4	16	1 048 576

Lösung von Aufgabe 2.10:
Sie sollte Bank A wählen, da sie dort etwa 9 540,15 Euro erhält.

Lösung von Aufgabe 2.11:

$$\left(\frac{1}{2}\right)^{\frac{10\,000}{5\,776}} \approx 0{,}301 \text{ Gramm}$$

Lösung von Aufgabe 2.12:

$\log_{10}(0{,}001) = -3$, $\log_7(\sqrt[4]{7^3}) = \log_7(7^{\frac{3}{4}}) = \dfrac{3}{4}$.

Lösung von Aufgabe 2.13:

a) nach 1 000 Jahren
b) nach 1 160,96 Jahren

Übungen Kapitel 3

Lösung von Aufgabe 3.1:
Es ist die Gleichung

$$12 + 6x = 60$$

zu lösen, es folgt $x = 8$. Sie können also 8 Gäste einladen, wenn Sie selbst hungern, oder 7, wenn Sie auch eine Pizza essen.

Lösung von Aufgabe 3.2:
a) $x = 1$
b) $x = -2$
c) $x = \frac{1}{9}$

Lösung von Aufgabe 3.3:

a) $x_1 = x_2 = -\sqrt{2}$
b) $x_1 = \sqrt{3}$, $x_2 = -3$

Lösung von Aufgabe 3.4:

a) Die Gleichung ist linear, die Lösung ist $x_1 = 5/2$.
b) Die Gleichung ist allgemein gültig, wird also von jeder reellen Zahl x gelöst, allerdings darf man die Nullstellen der Nenner nicht einsetzen. Die Lösungsmenge ist somit $\mathbb{L} = \mathbb{R} \setminus \{-2, 1\}$.

Lösung von Aufgabe 3.5:

a) $r_1(x) = \dfrac{x(x+12)(x-5)}{3(x-4)(x-5)} = \dfrac{x(x+12)}{3(x-4)}$

b) $r_2(x) \;=\; \dfrac{(x-2)(x+2)}{2(x-2)} \;=\; \dfrac{x+2}{2}$

Lösung von Aufgabe 3.6:
Vereinfachte Gleichung:

$$x^2 - 2x - 1 \;=\; 0.$$

Lösungen: $x_{1/2} = 1 \pm \sqrt{2}$.

Lösung von Aufgabe 3.7:

a) Unlösbar (bereits die substituierte Gleichung hat keine Lösung).
b) Zwei Lösungen: $x_{2,1} = -3$, $x_{2,2} = 3$.
c) Vier Lösungen: $x_{1,1} = -2$, $x_{1,2} = 2$, $x_{2,1} = -3$, $x_{2,2} = 3$.

Lösung von Aufgabe 3.8:

a) Zwei Lösungen: $x_1 = -2$, $x_2 = 2$.
b) Zwei Lösungen: $x_1 = 1$, $x_2 = 1/2$.
c) Lösung ist $x_1 = 7$, als *Scheinlösung* tritt noch $x_2 = 2$ auf.

Lösung von Aufgabe 3.9:

a) $x = 0$
b) $x = 2$
c) $x \approx 5{,}419$

Lösung von Aufgabe 3.10:
Etwa 352,25 Jahre.

Lösung von Aufgabe 3.11:

a) $\mathbb{L} = \{x \in \mathbb{R} \,|\, x > 6\}$
b) $\mathbb{L} = \{x \in \mathbb{R} \,|\, x > -9/2\}$
c) $\mathbb{L} = \{x \in \mathbb{R} \,|\, x < -1 \;\text{ oder }\; x > 2\}$

Übungen Kapitel 4

Lösung von Aufgabe 4.1:
Der Flächeninhalt F der Dreiecke in a), b) und c) ist jeweils: $F = \frac{1}{2} \cdot 4 \cdot 2 = 4$. In d) berechnet man mit Herons Formel ebenfalls $F = 4$.

Lösung von Aufgabe 4.2:
Die verbleibende Kathete hat die Länge $b = 3$.

Lösung von Aufgabe 4.3:
Die Katheten besitzen die Längen $a = (-1+\sqrt{3})$ und $b = 3-\sqrt{3}$, während die Länge der Hypotenuse $c = 2(\sqrt{3}-1)$ ist.

Lösung von Aufgabe 4.4:
Für die Länge der Höhe h auf die Seite mit Länge b gilt nach dem Satz des Pythagoras $h = \sqrt{a^2 - b^2/4}$. Einsetzen in die Formel für den Flächeninhalt führt auf $F = bh/2 = b\sqrt{4a^2 - b^2}/4$.

Lösung von Aufgabe 4.5:
Die verbleibenden Größen sind $q = 1, h = 1$ und $a = \sqrt{2}$.

Lösung von Aufgabe 4.6:
Die beiden verbleibenden Innenwinkel sind $\alpha = 60°$ und $\beta = 30°$. Die Hypotenusenlänge ist $c = 6$.

Lösung von Aufgabe 4.7:
Die möglichen Höhen sind $h_c = 4\ \sin(30°) = 2$ und $h_a = 5\ \sin(30°) = 5/2$. In beiden Fällen berechnet sich der Flächeninhalt des Dreiecks als $F = 5$.

Lösung von Aufgabe 4.8:
Es gelten: a) $\sin(x) = 1$ genau dann, wenn $x = \frac{\pi}{2} + 2k\pi$, wobei k eine beliebige ganze Zahl ist, b) $\cos(x) = -1$ genau dann, wenn $x = \pi + 2k\pi$, wobei k eine beliebige ganze Zahl ist.

Lösung von Aufgabe 4.9.:
Alle reellen Zahlen $x \neq \frac{\pi}{2} + k\pi$, wobei k eine ganze Zahl ist, denn dort gilt $\cos(x) \neq 0$.

Lösung von Aufgabe 4.10:
Nach dem zweiten Additionstheorem gelten:

$$\begin{aligned}
\cos(x) &= \cos(\tfrac{x+y}{2} + \tfrac{x-y}{2}) = \cos(\tfrac{x+y}{2})\cos(\tfrac{x-y}{2}) - \sin(\tfrac{x+y}{2})\sin(\tfrac{x-y}{2}),\\
\cos(y) &= \cos(\tfrac{x+y}{2} + \tfrac{y-x}{2}) = \cos(\tfrac{x+y}{2})\cos(\tfrac{y-x}{2}) - \sin(\tfrac{x+y}{2})\sin(\tfrac{y-x}{2}).
\end{aligned}$$

Wegen der Symmetrieeigenschaften der Cosinus- und Sinusfunktion gilt zudem

$$\cos(\tfrac{y-x}{2}) = \cos(\tfrac{x-y}{2}) \text{ und } -\sin(\tfrac{y-x}{2}) = \sin(\tfrac{x-y}{2}),$$

woraus die Behauptung folgt.

Lösung von Aufgabe 4.11:
Der Cosinussatz ergibt $\cos(\alpha) = \cos(\beta) = 0{,}75$. Hieraus folgt $\alpha = \beta \approx 41{,}41°$ und $\gamma \approx 97{,}18°$.

Lösung von Aufgabe 4.12:
Nein! Man betrachte beispielsweise die vorgegebene Längen $a = 1, b = 2$ und $c = 3$. Nimmt man an, es gibt ein Dreieck mit diesen Seitenlängen, so müsste gemäß der Dreiecksungleichung $a + b > c$ gelten. Dies ist aber für die vorgegebenen Längen *nicht* der Fall.

Lösung von Aufgabe 4.13:
Man berechnet $\sin(\beta) = 4\ \sin(10^\circ) \approx 0{,}6946$. Also: $\beta \approx 43{,}99^\circ$. Somit: $\gamma \approx 126{,}01^\circ$. Schließlich: $c \approx 2\ \sin(126{,}01^\circ)/\sin(10^\circ) \approx 9{,}3167$. Ist $\alpha = 40^\circ$, so wäre $\sin(\beta) = 4\ \sin(40^\circ) \approx 2{,}5712$. Dies kann aber wegen $|\sin(x)| \leq 1$ nicht sein – ein solches ebenes Dreieck existiert nicht.

Lösung von Aufgabe 4.14:
Sind zwei Seitenlängen b und c eines Dreiecks und der diese beiden Seiten einschließende Winkel α gegeben, so gilt zunächst $F = bc\ \sin(\alpha)/2$. Die Verwendung der Beziehung $\sin(\alpha) = a/2r$ ergibt dann $F = abc/4r$, wobei r der Radius des Umkreises ist.

Lösung von Aufgabe 4.15:
Es gilt $\alpha = 180^\circ - \beta = 60^\circ$ und somit ist die Länge der Diagonale d im Rhomboid $d = \sqrt{a^2 + b^2 - 2ab\cos(\alpha)} = \sqrt{19}$. Der Flächeninhalt des Parallelogramms ist $F = \sin(\alpha)\, ab = 15\sqrt{3}/2$.

Lösung von Aufgabe 4.16:
Die Lösung ergibt sich aus Abbildung L.1.

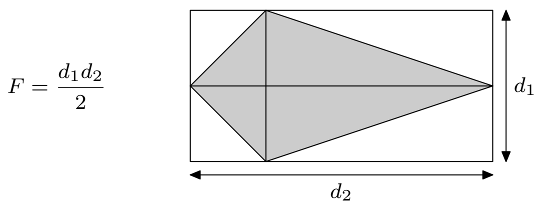

Abb. L.1 Der Flächeninhalt F eines Drachens (graue Fläche) mit Diagonalen der Länge d_1 und d_2 ist die Hälfte des Flächeninhalts eines Rechtecks mit Seitenlängen d_1 und d_2.

Lösung von Aufgabe 4.17:
Für die Länge der Höhe h auf die Seite mit Länge a gilt $h = \sin(\alpha)c = \sin(45^\circ)\sqrt{2} = 1$ und somit ergibt sich $F = (a + b)h/2 = 5$.

Lösung von Aufgabe 4.18:
Aus der Flächeninhaltsformel für das Trapez folgt, dass für die Länge der Höhe gilt $h = \sqrt{6}/3$. Damit berechnen sich die Längen a und b der parallelen Seiten des Trapezes als $a = h = \sqrt{6}/3$ und $b = 2\ h = (2\sqrt{6})/3$. Da das Trapez gleichschenklig ist, bestimmt man die verbleibende Seitenlänge c mittels des Satz des Pythagoras $c = \sqrt{30}/6$. Aus $\sin(\alpha) = h/c$ folgt dann $\alpha \approx 63{,}43^\circ$ und wegen der Gleichschenkligkeit des Trapezes gilt für den verbleibenden Winkel $\beta \approx 116{,}57^\circ$.

Lösung von Aufgabe 4.19:
a) Aus $F_{3\,600} = 450\, b^2 \sin((0{,}1)^\circ)/\sin^2((0{,}05)^\circ)$ berechnet man $F_{3\,600} \approx 1\,031\,323{,}769\, b^2$
b) Die Summe der Längen der Seiten des regelmäßigen N-Ecks ist $U_N = Nb = 2N\ \sin(180^\circ/N)$, da $a = 1$. Für $N = 3\,600$ ergibt sich $U_{3\,600} = 7\,200\ \sin((0{,}05)^\circ) \approx 6{,}283\,184\,51$.

Lösung von Aufgabe 4.20:
Die Fläche jeder Pizzaportion ist 200π cm^2 ≈ 628 cm^2 und der entsprechende Rand hat die Länge 10π cm $\approx 31{,}42$ cm. Das „Angebot" des Pizzabäckers nehmen Sie natürlich an, denn nach diesem ist die Gesamtfläche der Kreisring-Pizza $(2500-625)\pi$ cm^2 – jedem stehen dann $234{,}38\pi$ cm^2 ≈ 736 cm^2 Pizza zu. Der Rand jeder Portion hat dann allerdings die Länge $\approx 58{,}90$ cm.

Lösung von Aufgabe 4.21:
Der Ansatz $\pi 10^2 2/3 = \pi(2b)b$ führt auf den Nebenachsenradius $b = 10\sqrt{3}/3 \approx 5{,}77$ cm und den Hauptachsenradius $r = 20\sqrt{3}/3 \approx 11{,}55$ cm der Hamburger-Ellipse.

Übungen Kapitel 5

Lösung von Aufgabe 5.1:

$$\begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} - 4 \cdot \begin{pmatrix} -2 \\ 1 \\ 4 \end{pmatrix} + 2 \cdot \begin{pmatrix} 12 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 33 \\ -6 \\ -9 \end{pmatrix},$$

$$-\begin{pmatrix} 3 \\ -4 \end{pmatrix} + 3 \cdot \left[\begin{pmatrix} 2 \\ 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} -1 \\ 4 \end{pmatrix}\right] = \begin{pmatrix} 9 \\ -17 \end{pmatrix}.$$

Lösung von Aufgabe 5.2:

a) $\begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -12 \end{pmatrix} = -9, \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 5 \\ 6 \end{pmatrix} = -3.$

b) $\begin{pmatrix} -1 \\ 2 \end{pmatrix} \cdot \left(\begin{pmatrix} 3 \\ 4 \end{pmatrix} + \begin{pmatrix} -2 \\ 5 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 9 \end{pmatrix} = 17$

und

$$\begin{aligned} \begin{pmatrix} -1 \\ 2 \end{pmatrix} \cdot \left(\begin{pmatrix} 3 \\ 4 \end{pmatrix} + \begin{pmatrix} -2 \\ 5 \end{pmatrix}\right) &= \begin{pmatrix} -1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 4 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} -2 \\ 5 \end{pmatrix} \\ &= 5 + 12 = 17. \end{aligned}$$

Lösung von Aufgabe 5.3:

a) $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 7 \end{pmatrix}$ sind linear unabhängig.

b) $\begin{pmatrix} -2 \\ 5 \end{pmatrix}, \begin{pmatrix} 12 \\ 17 \end{pmatrix}, \begin{pmatrix} 234 \\ 1 \end{pmatrix}$ sind linear abhängig.

c) $\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \\ -4 \end{pmatrix}$ sind linear abhängig.

Lösung von Aufgabe 5.4:

a) $2 \cdot \begin{pmatrix} 1 & 0 & -2 \\ -3 & 5 & 1 \\ 4 & 2 & -1 \end{pmatrix} + 3 \cdot \begin{pmatrix} -6 & 3 & 4 \\ 1 & 2 & -2 \\ -3 & -4 & 0 \end{pmatrix} = \begin{pmatrix} -16 & 9 & 8 \\ -3 & 16 & -4 \\ -1 & -8 & -2 \end{pmatrix}.$

b) $-3 \cdot \begin{pmatrix} -1 & 5 & 2 \\ -2 & 1 & 0 \\ 3 & 1 & -2 \end{pmatrix} + \begin{pmatrix} 2 & -5 & -2 \\ 2 & -1 & 0 \\ -3 & -1 & 3 \end{pmatrix} = \begin{pmatrix} 5 & -20 & -8 \\ 8 & -4 & 0 \\ -12 & -4 & 9 \end{pmatrix}.$

Lösung von Aufgabe 5.5:

$$2 \cdot \begin{pmatrix} 1 & 4 \\ -2 & 6 \end{pmatrix} - 3 \cdot \left[\begin{pmatrix} 0 & 2 \\ 3 & -1 \end{pmatrix} + 4 \cdot \begin{pmatrix} 1 & 0 \\ -1 & -2 \end{pmatrix} \right] = \begin{pmatrix} -10 & 2 \\ -1 & 39 \end{pmatrix}.$$

Lösung von Aufgabe 5.6:

a) $\begin{pmatrix} 1 & 0 & -2 \\ -3 & 5 & 1 \\ 4 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} -6 & 3 & 4 \\ 1 & 2 & -2 \\ -3 & -4 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 11 & 4 \\ 20 & -3 & -22 \\ -19 & 20 & 12 \end{pmatrix}.$

b) $\begin{pmatrix} -1 & 5 & 2 \\ -2 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & -5 & -2 & 1 \\ 2 & -1 & 0 & 2 \\ -3 & -1 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 8 & 9 \\ -2 & 9 & 4 & 0 \end{pmatrix}.$

Lösung von Aufgabe 5.7:
Die Ergebnismatrix lautet

$$\begin{pmatrix} -17 & 38 & -5 \\ -11 & 25 & 26 \\ -38 & 35 & 7 \end{pmatrix}.$$

Dabei spielt es keine Rolle, ob man zuerst die Matrizenoperationen innerhalb der eckigen Klammer bearbeitet und dann die Matrizen multipliziert oder gleich ausmultipliziert und dann zusammenfasst.

Lösung von Aufgabe 5.8:
Hätte $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ eine Inverse $A^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, so müsste gelten:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = A \cdot A^{-1} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+2c & b+2d \\ 2a+4c & 2b+4d \end{pmatrix}.$$

Also wäre $a + 2c = 1$ und gleichzeitig $2a + 4c = 0$. Wegen $2a + 4c = 2(a + 2c)$ ist das unmöglich.

Lösung von Aufgabe 5.9:
Im ersten Gleichungssystem ergibt sich $x = 1$, $y = -1$, $z = 2$, im zweiten ist $x = -1$, $y = 7$, $z = -\frac{1}{2}$, $u = 3$.

Lösung von Aufgabe 5.10:
Das Gleichungssystem ist unlösbar.

Lösung von Aufgabe 5.11:
Das erste Gleichungssystem hat die Lösungen

$$x = z + 1, y = z - 1,\ z \in \mathbb{R} \text{ beliebig.}$$

Das zweite Gleichungssystem hat die Lösungen

$$x = u, y = 2u, z = -2u,\ u \in \mathbb{R} \text{ beliebig.}$$

Lösung von Aufgabe 5.12:
Die Vektoren sind linear unabhängig.

Lösung von Aufgabe 5.13:

$$A^{-1} = \begin{pmatrix} 1 & 1 & -4 \\ 2 & 2 & -7 \\ 0 & 1 & -4 \end{pmatrix}.$$

Lösung von Aufgabe 5.14:

a) $\overrightarrow{AB} = \begin{pmatrix} -3 \\ -3 \\ 2 \end{pmatrix}, \overrightarrow{CD} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$

b) $\overrightarrow{\mathbf{0}A} = \begin{pmatrix} 3 \\ 6 \\ -1 \end{pmatrix}, \overrightarrow{\mathbf{0}B} = \begin{pmatrix} 4 \\ 7 \\ 9 \end{pmatrix}.$

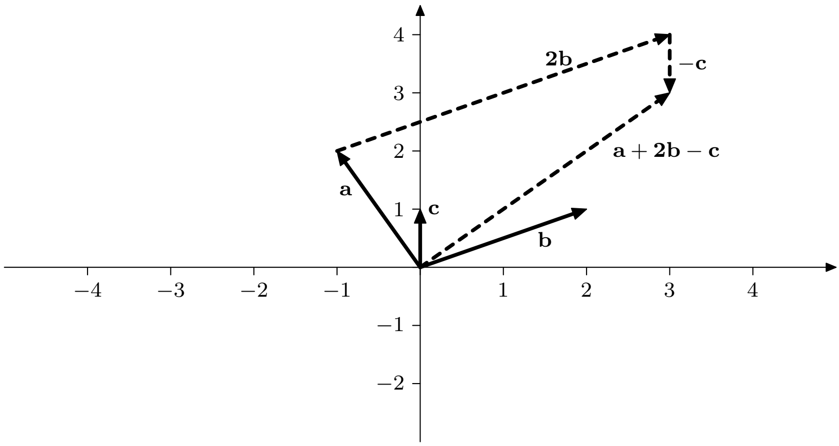

Abb. L.2 Vektoren $\mathbf{a}, \mathbf{b}, \mathbf{c}$ sowie $\mathbf{a} + 2\mathbf{b} - \mathbf{c}$.

Lösung von Aufgabe 5.15:

$\mathbf{a} + 2\mathbf{b} - \mathbf{c} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}, \ 3\mathbf{a} + \mathbf{b} + \mathbf{c} = \begin{pmatrix} -1 \\ 8 \end{pmatrix}.$

Die erste Rechnung ist in der graphischen Darstellung eingetragen.

Lösung von Aufgabe 5.16:
Die Gleichung lautet

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \lambda \cdot \begin{pmatrix} -3 \\ -1 \\ 2 \end{pmatrix} + \mu \cdot \begin{pmatrix} -1 \\ 0 \\ -2 \end{pmatrix} \text{ mit } \lambda, \mu \in \mathbb{R}.$$

Lösung von Aufgabe 5.17:

$$\begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} \times \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -6 \\ -16 \\ -1 \end{pmatrix}$$

und

$$\begin{pmatrix} -6 \\ -16 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix} = \begin{pmatrix} -6 \\ -16 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix} = 0.$$

Lösung von Aufgabe 5.18:
Die Gleichung lautet $3x + 6y + 9z = 12$ oder, wenn man auf beiden Seiten durch 3 teilt, $x + 2y + 3z = 4$.

Übungen Kapitel 6

Lösung von Aufgabe 6.1:
a) $f'(x) = 2\,x$, also: $f'(3) = 6$ und $f'(-2) = -4$
b) $x = -1$

Lösung von Aufgabe 6.2:
$(x+h)^3 - x^3 = x^3 + 3\,x^2h + 3\,xh^2 + h^3 - x^3 = h(3\,x^2 + 3\,xh + h^2)$ und somit folgt für den Grenzwert des Quotienten der Differenzen

$$\lim_{h\to 0} 3\,x^2 + 3\,xh + h^2 = 3\,x^2\,.$$

Lösung von Aufgabe 6.3:
$(x+h)^2 + (x+h) - x^2 - x = 2\,hx + h^2 + h = h(2\,x + h + 1)$ und somit folgt für den Grenzwert der Differenzenquotienten, also den Differentialquotienten:

$$\lim_{h\to 0} 2\,x + h + 1 = 2\,x + 1\ .$$

Lösung von Aufgabe 6.4:

a) $f(x) = |x-2|$ ist an der Stelle $x = 2$ nicht ableitbar, also nicht ableitbar
b) $f(x) = 1/2((x-2)^3 + |x-2|^3)$ ist ableitbar und insbesondere in $x = 2$ ableitbar: $f'(2) = 0$.

Lösung von Aufgabe 6.5:
$(3\,\cos(x) - 2\,\sin(x))' = 3\cos'(x) - 2\,\sin'(x) = -3\sin(x) - 2\,\cos(x)$.

Lösung von Aufgabe 6.6:

a) $f'(x) = -3\,x^2 - 4\,x + \frac{1}{8}$
b) $f_t'(x) = 4\,t^2x^3 - 4\,t^4x$
c) $f_x'(t) = -8\,x^2t^3 + 2\,(x^4+1)t$

Lösung von Aufgabe 6.7:

a) $f'(x) = -5\,x^4 + 16\,x^3 - 6\,x^2$
b) $f'(x) = (2x+1)\,\cos(x) - (x^2+x)\sin(x)$
c) $f'(x) = x(2+x)\,e^x$
d) $f'(x) = (\cos(x) - \sin(x))\,e^x$
e) $f'(x) = -2\sin(x)\cos(x)$

Lösung von Aufgabe 6.8:

a) $f'(x) = 16\,(2x-4)(x^2-4x+2)^{15}$
b) $f'(x) = (2x-1)\,e^{x^2-x}$
c) $f'(x) = (n/m)\,x^{(n/m)-1}$, wobei $x > 0$
d) $f'(x) = 2x\,(x^2-3)\,(x^2+1)^{-3}$

Lösung von Aufgabe 6.9:

a) $f'(x) = (-2x^2 + 16x + 1)/(2\,x^2 + 1)^2$, x beliebig
b) $f'(x) = (-2x^2 - 16x - 1)/(2\,x^2 - 1)^2$, $x \neq \sqrt{2}/2$ und $x \neq -\sqrt{2}/2$
c) $f(x) = (\cos(x)x - \sin(x))/x^2$, $x \neq 0$
d) $f(x) = (2x(x^2 - 3))/(x^2 + 1)^3$, x beliebig, dieselbe (Ableitungs-)Funktion wie in Aufgabe 6.8 d)
e) $f(x) = (-x^2 - 3x - 1)/e^x$, x beliebig

Lösung von Aufgabe 6.10:

a) $p'(x) = 16x^3 - 12x$, $p''(x) = 48x^2 - 12$, $p^{(3)}(x) = 96x$, $p^{(4)}(x) = 96$, $p^{(i)}(x) = 0$, $i \geq 5$
b) $f_t'(x) = 3t^2\,x^2 - 2t\,x + t$, $f_t''(x) = 6t^2\,x - 2t$, $f_t'''(x) = 6t^2$, $f_t^{(i)}(x) = 0$, $i \geq 4$

Lösung von Aufgabe 6.11:
$(e^{x^2-x})''' = (8x^3 - 12x^2 + 18x - 7)\,e^{x^2-x}$ und $(e^{x^2-x})'''' = (16x^4 - 32x^3 + 72x^2 - 56x + 25)\,e^{x^2-x}$

Lösung von Aufgabe 6.12:
$f'(x) = (-1/2)\,|x|$ ist in $x = 0$ nicht ableitbar.

Lösung von Aufgabe 6.13:
$(e^{x^2-x})' = 0$ hat nur die Lösung $x = 1/2$, der Wert der zweiten Ableitung $(e^{x^2-x})'' = (4x^2 - 4x + 3)e^{x^2-x}$ in $x = 1/2$ ist > 0, $T = (1/2|e^{-1/4})$ ist der Tiefpunkt des Graphen von e^{x^2-x}

Lösung von Aufgabe 6.14:
a) $f'(x) = (1/3)\,x^3$, $f''(x) = x^2$, $f'''(x) = 2x$, Test von Wendestellen $f''(x) = 0$ liefert Kandidaten $x = 0$, jedoch ist $f'''(0) = 0$, $x = 0$ ist kein Extremum von f', also keine Wendestelle von f.

b) $(e^{x^2-x})'' = (4x^2 - 4x + 3)e^{x^2-x} = 0$ führt auf quadratische Gleichung $x^2 - x + 3/4 = 0$, die keine reellen Lösungen hat – also sind keine Kandidaten für Wendestellen vorhanden.

Lösung von Aufgabe 6.15:
Der Definitionsbereich von f_t ist $\mathbb{R} \setminus \{0\}$. Die Nullstellen von f_t sind $x = -t$ und $x = t$. f_t besitzt keine lokalen Extrema und Wendestellen. Der Graph von f_t ist achsensymmetrisch zur y-Achse. $a_{f_t}(x) = -t$ ist horizontale Asymptote von f_t.

Lösung von Aufgabe 6.16:
Es handelt sich um eine quadratische Fläche mit Seitenlänge $\sqrt{2} \approx 1.41$ Kilometer.

Lösung von Aufgabe 6.17:

a) $\int_1^3 x^2\,dx = 26/3$
b) $\int_1^t x^2\,dx = t^3/3 - 1/3 = 7/3$ und $t > 1$ implizieren $t = 2$

Lösung von Aufgabe 6.18:

a) Orientierter Flächeninhalt: $\int_{-1}^{1} x^3 - x\,dx = 0$, absoluter Flächeninhalt: $\int_{-1}^{1} |x^3 - x|\,dx = \int_{-1}^{0} x^3 - x\,dx - \int_{0}^{1} x^3 - x\,dx = 1/2$

b) $\int_{-1}^{1} 4x^4 - 6x^2 + e^8\,dx = -12/5 + 2\,e^8$, $\int_{-1}^{1} t^3x^3 - tx^2 + tx + t^3\,dx = 2t^3 - (2/3)t$

Lösung von Aufgabe 6.19:
$\int_0^t -x^2 + 2x\,dx = -t^3/3 + t^2 = F_0(t)$, erste Ableitung F_0' von F_0 stimmt mit Integrand überein: $F_0'(t) = -t^2 + 2t$.

Lösung von Aufgabe 6.20:

a) $H'(x) = -x^2 + 2x = h(x)$

b) Da $(e^{x^2-x})' = (2x-1)\,e^{x^2-x}$ gilt, folgt, dass $F(x) = e^{x^2-x} + c$ Stammfunktionen von f sind (c ist hier eine beliebige Konstante).

Lösung von Aufgabe 6.21:

a) $\int_0^2 -x^2 + 2x\,dx = [-x^3/3 + x^2]_0^2 = 4/3$

b) $\int_1^2 (2x-1)\,e^{x^2-x}\,dx = e^{x^2-x}|_1^2 = e - 1$

c) $\int_1^2 (1-x^2)/(1+x^2)^2\,dx = [x/(1+x^2)]_1^2 = -1/10$

d) Aus Übungsaufgabe 6.7 b) folgt: $F(x) = \cos(x)e^x$ ist Stammfunktion von $f(x) = (\cos(x) - \sin(x))\,e^x$. Somit: $\int_0^\pi (\cos(x) - \sin(x))\,e^x\,dx = \cos(x)e^x|_0^\pi = -(e^\pi + 1)$.

e) $F(x) = -1/x$ ist Stammfunktion von $f(x) = 1/x^2$. Somit: $\int_1^2 1/x^2\,dx = 1/x|_1^2 = 1/2$.

Lösung von Aufgabe 6.22:
$\int^t \sin(x)\,x\,dx + c = \sin(t) - t\cos(t) + c$ und $\int^t x^2\,e^x\,dx = (t^2 - 2t + 2)\,e^t + c$, wobei c eine beliebige Konstante ist.

Lösung von Aufgabe 6.23:
$\int^t 3x^2(x^3+4)^{88}\,dx + c = \int^{t^3+4} \bar{x}^{88}\,d\bar{x} + c = (t^3+4)^{89}/89 + c$ und $\int^t (2x-1)\,e^{x^2-x}\,dx + c = \int^{t^2-t} e^{\bar{x}}\,d\bar{x} + c = e^{t^2-t} + c$, wobei c eine beliebige Konstante ist.

Lösung von Aufgabe 6.24:
Definiert man $g(x) = x^2 + 1$, so gilt $g'(x) = 2x$. Anwenden der Substitutionsregel ergibt dann: $1/2\,\int_1^2 2x/(x^2+1)\,dx = 1/2\,\int_{g(1)}^{g(2)} 1/\bar{x}\,d\bar{x} = 1/2\,(\ln(5) - \ln(2))$.

Übungen Kapitel 7

Lösung von Aufgabe 7.1:
Es gibt $32^3 = 32\,768$ Möglichkeiten, die gesuchte Wahrscheinlichkeit ist somit

$$\frac{1}{32\,768} \,.$$

Lösung von Aufgabe 7.2:
Wegen $n = 7$ und $l = 5$ gibt es

$$\binom{11}{5} \;=\; 462$$

Möglichkeiten, von denen nur eine einzige das gewünschte Wort ergibt. Die gesuchte Wahrscheinlichkeit ist somit

$$\frac{1}{462} \;\approx\; 0{,}002\,16 \,.$$

Lösung von Aufgabe 7.3:
Die gesuchte Wahrscheinlichkeit ist

$$\frac{(7-5)!}{7!} \;=\; \frac{1}{3\cdot 4\cdot 5\cdot 6\cdot 7} \;=\; \frac{1}{2\,520} \;\approx\; 0{,}000\,396\,8 \,.$$

Lösung von Aufgabe 7.4:
a) Der Trainer hat

$$\binom{15}{11} \;=\; \frac{15!}{11!\cdot 4!} \;=\; 1\,365$$

Möglichkeiten.
b) Der neue Trainer hat

$$\binom{7}{5}\cdot\binom{6}{5}\cdot\binom{2}{1} \;=\; \frac{7!\cdot 6!\cdot 2!}{5!\cdot 2!\cdot 5!\cdot 1!\cdot 1!} \;=\; 252$$

Möglichkeiten.

Lösung von Aufgabe 7.5:
a) Es gibt

$$3^{11} \;=\; 177\,147$$

verschiedene Tipps.
b) Bei '5 aus 25' ist die Wahrscheinlichkeit für einen Haupttreffer gleich

$$\binom{25}{5} \;=\; \frac{1}{53\,130}\,,$$

bei '4 aus 20' ist sie dagegen

$$\binom{20}{4} \;=\; \frac{1}{4\,845},$$

also wesentlich höher.

c) Insgesamt haben die Fahrgäste

$$\frac{7!}{(7-4)!} \;=\; 840$$

Möglichkeiten.

d) Der Veranstalter hat

$$\binom{12}{4} \cdot \binom{8}{4} \;=\; \frac{12! \cdot 8!}{4! \cdot 8! \cdot 4! \cdot 4!} \;=\; 34\,650$$

Möglichkeiten.

Lösung von Aufgabe 7.6:
Durch Addition der einzelnen Häufigkeiten ergibt sich $n = 797$. Damit wird

$$h_{797}(A_1) = \frac{140}{797} \approx 0{,}176,$$
$$h_{797}(A_2) = \frac{393}{797} \approx 0{,}493,$$
$$h_{797}(A_3) = \frac{277}{797} \approx 0{,}348.$$

Lösung von Aufgabe 7.7:
Es sind insgesamt 166 Gewinne zu ziehen, die gesuchte Wahrscheinlichkeit also

$$p(G) \;=\; \frac{166}{500} \;=\; 0{,}332\ .$$

Alternativ kann man auch wie folgt rechnen: Die Wahrscheinlichkeit dafür, eine Niete zu ziehen, ist

$$\frac{334}{500} \;=\; 0{,}668\ ,$$

und da dies gerade die Wahrscheinlichkeit dafür ist, *keinen* Gewinn zu ziehen, folgt

$$p(G) \;=\; 1 - 0{,}668 \;=\; 0{,}332\ .$$

Lösung von Aufgabe 7.8:
Die Unvereinbarkeit der drei Ereignisse sieht man unmittelbar:

Die Einzelwahrscheinlichkeiten sind $p(A) = 1/6$, $p(B) = 1/3$, $p(C) = 1/6$. Somit ist

$$p(A) + p(B) = 1/2,$$
$$p(B) + p(C) = 1/2,$$
$$p(A) + p(C) = 1/3.$$

Dies erhält man auch durch Berechnung von $p(A \cup B)$, $p(B \cup C)$, und $p(A \cup C)$.

Lösung von Aufgabe 7.9:
Es sei E = „Schüler lernt Englisch" und L = „Schüler lernt Latein". Dann ist

$$p(E) = \frac{35}{50} = 0{,}7 \quad \text{und} \quad p(L) = \frac{25}{50} = 0{,}5\ .$$

Da jeder Schüler mindestens eine Sprache erlernt, ist $p(E \cup L) = 1$. Es folgt:
a) $1 = p(E) + p(L) - p(E \cap L)$, also $p(E \cap L) = 0{,}2$.
b) $p(\text{„nur Englisch"}) = p(E) - p(E \cap L) = 0{,}5$.
c) $p(\text{„nur eine Sprache"}) = p(\text{„nur Englisch"}) + p(\text{„nur Latein"}) = 0{,}5 + 0{,}3 = 0{,}8$.

Lösung von Aufgabe 7.10:
Sind A und B unvereinbar, so ist $A \cap B$ umöglich, also $p(A \cap B) = 0$. Damit stimmt (7.4) mit (7.3) überein.

Lösung von Aufgabe 7.11:
Es sei B das Ereignis „Teil stammt von Firma B" und OK das Ereignis „Teil ist intakt". Dann ist nach Angabe $p(OK \cap B) = 0{,}23$, und es folgt

$$p(OK|B) \;=\; \frac{p(OK \cap B)}{p(B)} \;=\; \frac{0{,}23}{0{,}25} \;=\; 0{,}92\,,$$

es sind also 92% der von B gelieferten Teile intakt.

Übungen Kapitel 8

Lösung von Aufgabe 8.1:
Alle drei Ausdrücke sind komplexe Zahlen.

Lösung von Aufgabe 8.2:
$z_1 \cdot z_2 = 8 - 9i$,
$\dfrac{z_1}{z_2 \cdot z_3} = -\dfrac{9}{50} - \dfrac{37}{50}\,i$,
$\dfrac{z_1 + z_2}{z_2 - z_3} = \dfrac{1+3i}{-2+i} \;=\; \dfrac{1}{5} - \dfrac{7}{5}\,i\ .$

Lösung von Aufgabe 8.3:
a) $1 - i = 1{,}4142 \cdot (0{,}7071 - 0{,}7071i)$,
b) $-5 - 3i = 5{,}8309 \cdot (-0{,}8575 - 0{,}5145i)$.

Lösung von Aufgabe 8.4:
a) $(-1+2i)^5 \;=\; -41 - 38i$,
b) $\left(\dfrac{1}{\sqrt{2}} + \dfrac{i}{\sqrt{2}}\right)^8 \;=\; 1.$

Lösung von Aufgabe 8.5:

a) Die beiden zweiten Wurzeln aus $-2+3i$ sind

$$0{,}8960 + 1{,}6741i \quad \text{und} \quad -0{,}8960 - 1{,}6741i\,.$$

b) Die drei dritten Wurzeln aus 8 lauten 2, $-1+i\sqrt{3}$ und $-1-i\sqrt{3}$.

Lösung von Aufgabe 8.6:

a) $x_{1/2} \;=\; -\frac{1}{2} \pm i \cdot \frac{3}{2}$.
b) Die vier Lösungen lauten i, $-i$, $\sqrt{2}$ und $-\sqrt{2}$.

Literatur

Fritzsche, K.: *Mathematik für Einsteiger*, Spektrum Akademischer Verlag, Heidelberg 1995

Knorrenschild, M.: *Vorkurs Mathematik*, Fachbuchverlag Leipzig 2004

Reinhard, F., Soeder, H.: dtv-Atlas Mathematik, Band 1: *Grundlagen Algebra und Geometrie*, Deutscher Taschenbuch Verlag, München 1998

Reinhard, F., Soeder, H.: dtv-Atlas Mathematik, Band 2: *Analysis und Angewandte Mathematik*, Deutscher Taschenbuch Verlag, München 1998

Rießinger, Th.: *Übungsaufgaben zur Mathematik für Ingenieure*, Springer Verlag, Berlin, Heidelberg 2004

Rießinger, Th.: *Mathematik für Ingenieure*, Springer Verlag, Berlin, Heidelberg 2005

Rommelfanger, H.: *Mathematik für Wirtschaftswissenschaftler*, Bände 1 und 2, Elsevier, München 2002/2004

Schwarze, J.: *Mathematik für Wirtschaftswissenschaftler*, Band 3, Lineare Algebra, Lineare Optimierung und Graphentheorie, Verlag Neue Wirtschafts-Briefe 2005

Walz, G. (Hrsg.): *Lexikon der Mathematik in 6 Bänden*, Spektrum Akademischer Verlag, Heidelberg 2002

Walz, G.: *Mathematik für Fachhochschule, Duale Hochschule und Berufsakademie*, Spektrum Akademischer Verlag, Heidelberg 2010.

Stichwortverzeichnis