

Mathematische Grundlagen 2: Lineare Algebra und Differential- und Integralrechnung

Dr. Viktoriya Ozornova

20. Januar 2018

Warnung

Dieses Skript enthält unter Umständen Fehler und Ungenauigkeiten jeder Art. Falls Sie solche entdecken, wäre es nett, wenn Sie mir Ihre Anmerkungen per E-Mail an

`viktoriya.ozornova@rub.de`

schicken könnten. Die Fehler werden nur in der neusten Version des Skriptes korrigiert. Dieses Skript erhebt keinerlei Anspruch auf Originalität. Insbesondere habe ich bei der Erstellung dieses Skriptes die vorherigen Materialien von Emanuele Delucchi und Martin Dlugosch herangezogen, sowie die Bücher „Mathematik für Informatiker“ von G. Teschl und S. Teschl, „Mathematik für Informatiker“ von B. Kreußler und G. Pfister und „Mathematik für Informatiker“ von P. Hartmann.

Ferner möchte ich den Leuten danken, die bei der Erstellung dieses Skriptes auf die eine oder andere Weise hilfreich waren, insbesondere Roman Bruckner, Emanuele Delucchi, Martin Dlugosch, Eva-Maria Feichtner, Dmitry Feichtner-Kozlov, Tim Haga, Damien Imbs, Jan-Philipp Litza, Lennart Meier, Jan Senge, Ingolf Schäfer, Kirsten Schmitz sowie den vielen Studierenden der Informatik, die mir zu dem Skript Rückmeldung gegeben haben.

Inhaltsverzeichnis

I	Lineare Algebra	4
1	Lösen linearer Gleichungssysteme	4
2	Vektorraum \mathbb{R}^n und seine Unterräume	25
3	Lineare Unabhängigkeit und Basen	44
4	Lineare Abbildungen und Matrizen	74
5	Matrizenmultiplikation	94
6	Invertieren von Matrizen	101
7	Eigenwerte und Eigenvektoren	116
8	Skalarprodukt, Abstände und Winkel	129
II	Analysis: Differential- und Integralrechnung	150

9	Reelle und komplexe Zahlen	151
10	Folgen, Reihen und Grenzwerte	166
11	Die Ableitung einer Funktion	183
12	Exponentialfunktion und Logarithmus	189
13	Taylor-Polynome	201
14	Trigonometrische Funktionen	207
15	Integralrechnung	227
III	Numerische Verfahren	236
16	Vektoriteration	239
17	Lineare Regression	250

Teil I

Lineare Algebra

Lineare Algebra beschäftigt sich mit Studium von Objekten, die von Natur her etwas wie eine Gerade sind: Etwa mit Geraden in der Ebene, mit Ebenen in dem dreidimensionalen Raum und mit höherdimensionalen Analoga, die wir im Laufe dieses Kurses kennenlernen werden. Ferner kennen Sie eventuell lineare Gleichungssysteme oder lineare Funktionen aus dem Schulunterricht oder aus Mathematischen Grundlagen 1. Genauer können die Aufgaben der linearen Algebra erläutert werden, wenn wir zunächst einige Grundbegriffe kennengelernt haben.

1 Lösen linearer Gleichungssysteme

Zunächst wollen wir den Begriff des linearen Gleichungssystems erläutern. Diesen werden wir nicht in voller Strenge definieren; allerdings werden wir einen sehr ähnlichen Begriff präzise einführen.

Allgemein passiert es sowohl in mathematischen Sachverhalten als auch in der Anwendungspraxis, dass man eine Anzahl von Unbekannten bzw. Variablen hat, für die man die möglichen Werte ermitteln möchte. Dabei werden diese Variablen typischerweise irgendwelche Bedingungen erfüllen müssen. Diese Bedingungen können im Allgemeinen ganz unterschiedlicher Natur sein und müssen nicht durch Gleichungen vorgegeben sein. Will man die Anzahl der Mitglieder eines Komitees oder eines Parlaments bestimmen, so muss diese Anzahl beispielsweise eine ganze, nicht-negative (oder häufig positive) Zahl sein: Es können nicht 4,2 oder -7 Personen in einem Komitee vertreten sein. Andere Einschränkungen werden beispielsweise in Form von Ungleichungen vorgegeben; eine Geschwindigkeitsbeschränkung auf 60 Stundenkilometer würde etwa durch die Ungleichung $v \leq 60$ (bei entsprechender Wahl von Bezeichnungen und Einheiten) gegeben sein.

Wir wollen uns allerdings momentan auf den vergleichsweise einfachen Fall beschränken, dass alle Bedingungen, die wir an unsere Variablen vorgegeben haben, Gleichungen sind. Wir wollen unsere Situation durch eine weitere Annahme zusätzlich vereinfachen. Bei einem allgemeinen Gleichungssystem stehen nämlich die Chancen schlecht, es exakt lösen zu können. Hat man etwa das Gleichungssystem

$$\begin{cases} \sin(x) + e^y = 2, \\ 2^{xy} = 22, \end{cases}$$

so wird es im Allgemeinen nicht möglich - oder sehr schwierig - sein, ein solches Gleichungssystem zu lösen. Insbesondere gibt es kein allgemeines Verfahren, das zur Lösung solcher Gleichungssysteme eingesetzt werden kann.

Manchmal kann man trotzdem eine exakte Lösung finden, und im Allgemeinen kann man durch numerische Methoden versuchen, eine Näherungslösung zu finden.

Wir wollen uns jedoch auf den besonders einfachen Fall von linearen Gleichungssystemen einschränken, für den es tatsächlich ein effektives Lösungsverfahren gibt. Wir wollen dabei keine „komplizierten“ Funktionen in den Gleichungen erlauben: Beispielsweise keine Variablen in Exponenten, keine trigonometrische Funktionen, keine höheren Potenzen von Variablen (also z.B. keine quadratischen oder kubischen Terme), keine Produkte von Variablen... Wir wollen etwas genauer erläutern, was nun erlaubt ist. Hierbei ist die zulässige Form der Gleichungen unter anderem durch die Geradengleichungen in der Ebene motiviert, die sie vielleicht in der Form $y = 2x + 7$ oder in der Form $2x + 3y = 5$ kennen. Ähnlich wollen wir nur zulassen, dass Variablen mit festen reellen Zahlen multipliziert werden, die Ergebnisse addiert und subtrahiert werden dürfen und zusätzlich noch Konstanten (also ebenfalls feste reelle Zahlen) addiert werden dürfen.

Beispiel. Das Gleichungssystem

$$\begin{cases} x = 3 - 2y, \\ 2x + 3y - 5 = 0, \end{cases}$$

ist ein lineares Gleichungssystem in den Variablen x, y .

Wir wollen nun eine präzise Definition von linearen Gleichungssystemen in Standardform geben. Da der Unterschied zu beliebigen Gleichungssystemen gering ist, werden wir „in Standardform“ häufig weglassen. Hierbei wollen wir erreichen, dass alle Variablen auf den linken Seiten von Gleichungen sind, zusammengefasst sind und in gleicher Reihenfolge in jeder Gleichung erscheinen, während die konstanten Terme alle auf der rechten Seite erscheinen. Wir werden auch für die festen reellen Zahlen, mit denen die Variablen multipliziert werden, in einem festgelegten Muster nummerieren, das an ein doppeltes Array in einigen Programmiersprachen erinnert.

Definition 1.1. Ein **lineares Gleichungssystem in Standardform** in n Variablen mit k Gleichungen ist ein Gleichungssystem der Form

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2, \\ \vdots \\ a_{k1}x_1 + a_{k2}x_2 + a_{k3}x_3 + \dots + a_{kn}x_n = b_k, \end{cases}$$

wobei x_1, x_2, \dots, x_n die Variablen sind, a_{ij} für $1 \leq i \leq k$ und $1 \leq j \leq n$ feste reelle Zahlen (genannt **Koeffizienten**) sind und b_i für $1 \leq i \leq k$ ebenfalls

festen reellen Zahlen sind. Der Koeffizient a_{ij} steht in der i -ten Gleichung bei der Variablen x_j .

Falls alle Konstanten auf der rechten Seite 0 sind, also $b_i = 0$ für $1 \leq i \leq k$, nennt man das lineare Gleichungssystem **homogen**. Andernfalls, also wenn $b_i \neq 0$ für mindestens ein i zwischen 1 und k (einschließlich) ist, nennt man das lineare Gleichungssystem **inhomogen**.

Beispiel. Das eben betrachtete lineare Gleichungssystem

$$\begin{cases} x = 3 - 2y, \\ 2x + 3y - 5 = 0 \end{cases}$$

ist nicht in Standardform. Das lässt sich jedoch leicht durch Äquivalenzumformungen in den jeweiligen Gleichungen beheben:

$$\begin{aligned} & \begin{cases} x = 3 - 2y & | + 2y, \\ 2x + 3y - 5 = 0 & | + 5 \end{cases} \\ \Leftrightarrow & \begin{cases} x + 2y = 3, \\ 2x + 3y = 5. \end{cases} \end{aligned}$$

Das ist nun ein lineares Gleichungssystem in Standardform, das dieselbe Lösungsmenge besitzt wie das ursprüngliche (auf den Begriff der Lösungsmenge gehen wir später nochmal detaillierter ein).

Um die Nummerierung der Koeffizienten zu verdeutlichen, wollen wir die „Nummern“ der Koeffizienten in diesem Gleichungssystem festhalten:

$$\begin{array}{lll} a_{11} = 1 & a_{12} = 2 & b_1 = 3, \\ a_{21} = 2 & a_{22} = 3 & b_2 = 5. \end{array}$$

Dieses lineare Gleichungssystem ist inhomogen.

Beispiel. Das lineare Gleichungssystem

$$\begin{cases} 2x = 3y, \\ 4w = 5z \end{cases}$$

in den Variablen x, y, z, w ist nicht in Standardform. Hier erreichen wir wiederum durch Äquivalenzumformungen der einzelnen Gleichungen (und Einfügen der Koeffizienten 0 für die fehlenden Variablen) die folgende Standardform:

$$\begin{cases} 2x - 3y + 0z + 0w = 0, \\ 0x + 0y - 5z + 4w = 0. \end{cases}$$

(Hierbei müssen wir festlegen, in welcher Reihenfolge die Variablen aufgezählt werden sollen. Häufig wird das alphabetisch sein, hier aber beispielsweise nicht.)

Die Nummerierung der Koeffizienten ist in diesem Fall

$$\begin{array}{ccccc} a_{11} = 2 & a_{12} = -3 & a_{13} = 0 & a_{14} = 0 & b_1 = 0, \\ a_{21} = 0 & a_{22} = 0 & a_{23} = -5 & a_{24} = 4 & b_2 = 0. \end{array}$$

Dieses lineare Gleichungssystem ist homogen.

Als nächstes wollen wir uns Gedanken darüber machen, was es bedeutet, ein solches lineares Gleichungssystem zu lösen. Jede Lösung eines linearen Gleichungssystems wie in der Definition 1.1 wird ein (geordnetes) Tupel reeller Zahlen (x_1, \dots, x_n) sein, das die Eigenschaft hat, jeder der Gleichungen des linearen Gleichungssystems zu genügen. Genauer heißt das: Setzen wir dieses Zahlentupel anstelle von entsprechenden Variablen in die linke Seite jeder Gleichung des linearen Gleichungssystems ein, so erhalten wir genau die rechte Seite der entsprechenden Gleichung. Das lineare Gleichungssystem zu lösen, heißt, dass man alle Lösungen davon findet und nachweist, dass keine weitere Lösungen existieren. Bei vielen Methoden wird der zweite Teil „automatisch“ erledigt; das ist jedoch nicht notwendigerweise der Fall: Betrachtet man das uns schon bekannte lineare Gleichungssystem

$$\begin{cases} x + 2y = 3, \\ 2x + 3y = 5, \end{cases}$$

so stellt man durch Ausprobieren beispielsweise fest, dass das Paar $(1, 1)$ eine Lösung des Gleichungssystems ist. Das beantwortet jedoch nicht die Frage, ob weitere Lösungen existieren.

Es ist ferner a priori nicht ganz klar, was es heißt, alle Lösungen zu finden. Sollte das Gleichungssystem keine oder genau eine Lösung besitzen, so ist es einfach, alle Lösungen anzugeben. Häufig besitzt allerdings ein lineares Gleichungssystem unendlich viele Lösungen, und in diesem Fall gilt es, die Menge aller Lösungen in möglichst einfacher Form anzugeben. Wir werden später nochmal genauer darauf eingehen.

Beispiel. Wir wollen uns nochmal mit dem linearen Gleichungssystem

$$\begin{cases} x + 2y = 3, \\ 2x + 3y = 5, \end{cases}$$

beschäftigen. Um das zu lösen, verwenden wir nun eine Methode, die häufig als erstes im Schulunterricht behandelt wird: Das sukzessive Auflösen von Gleichungen nach einer Variablen und Einsetzen dieser in die übrigen Gleichungen. Dies ist eine gültige Methode, die gerade beim Lösen kleiner linearer Gleichungssysteme „per Hand“ Vorteile haben kann, die jedoch einige Nachteile hat, wie wir später diskutieren werden.

In diesem Fall bietet es sich an, die erste Gleichung nach x aufzulösen und das Ergebnis danach in die zweite Gleichung einzusetzen. Wir erhalten:

$$\begin{aligned}
 & \begin{cases} x + 2y = 3 & | - 2y, \\ 2x + 3y = 5 & . \end{cases} \\
 \Leftrightarrow & \begin{cases} x = 3 - 2y, \\ 2x + 3y = 5 \end{cases} \\
 \xLeftrightarrow{\text{Eins.}} & \begin{cases} x = 3 - 2y, \\ 2 \cdot (3 - 2y) + 3y = 5 \end{cases} \\
 \xLeftrightarrow{(1)} & \begin{cases} x = 3 - 2y, \\ 6 - 4y + 3y = 5 \end{cases} \\
 \xLeftrightarrow{(2)} & \begin{cases} x = 3 - 2y, \\ 6 - y = 5 & | - 6 | \cdot (-1) \end{cases} \\
 \Leftrightarrow & \begin{cases} x = 3 - 2y, \\ y = 1 \end{cases} \\
 \xLeftrightarrow{\text{Eins.}} & \begin{cases} x = 1, \\ y = 1. \end{cases}
 \end{aligned}$$

Dabei haben wir im Schritt (1) die Klammer aufgelöst und im Schritt (2) die Terme zusammengefasst. Da wir das ursprüngliche lineare Gleichungssystem stets äquivalent umgeformt haben, haben wir dabei die Lösungsmenge nicht verändert. Somit haben wir gezeigt, dass die Lösungsmenge (also die Menge aller Paare (x, y) , die dem ursprünglichen linearen Gleichungssystem genügt) die Menge $\{(1, 1)\}$ ist.

Die im letzten Beispiel vorgestellte Lösungsmethode hat mehrere Nachteile. Beim Lösen der linearen Gleichungssysteme per Hand kann man sich gut vorstellen, dass diese Methode schon bei 4 oder 5 Gleichungen sehr unübersichtlich wird, und erst recht bei größeren Gleichungssystemen. Ferner lässt diese Methode sich vergleichsweise schlecht automatisieren. Zuletzt sei noch angemerkt, dass die allgemeine Lösungsmethode, die wir als nächstes verwenden - das Gauß'sche Eliminationsverfahren - auch in einigen theoretischeren Fragestellungen von Vorteil sein wird; zu diesem Punkt werden wir möglicherweise später nochmal zurückkommen.

Gauß'sches Eliminationsverfahren

Um das allgemeine Verfahren zum Lösen linearer Gleichungssysteme zu erläutern, führen wir zunächst die elementaren Zeilentransformationen ein.

Definition 1.2. Die **elementaren Zeilentransformationen** eines linearen Gleichungssystem (in Standardform) sind die folgenden drei Operationen:

- (Z1) Vertauschen zweier Zeilen,
- (Z2) Multiplikation einer Zeile mit einer festen reellen Zahl $a \neq 0$,
- (Z3) Addition eines Vielfachen der Zeile j zu der Zeile $i \neq j$.

Der folgende Satz erläutert einen Teil der Wichtigkeit der elementaren Zeilentransformationen.

Satz 1.3. *Ein lineares Gleichungssystem, das aus einem anderen durch Anwendung der elementaren Zeilentransformationen hervorgegangen ist, hat dieselbe Lösungsmenge wie das ursprüngliche.*

Wir werden diesen Satz nicht formal beweisen; allerdings ist es eine präzisere Formulierung eines Sachverhaltes, der vielfach schon aus dem Schulunterricht bekannt ist.

Die einfachste Form des Gauß'schen Eliminationsverfahrens sieht nun wie folgt aus: Man wende diese elementaren Zeilentransformationen so lange auf ein lineares Gleichungssystem an, bis es so einfach ist, dass man die Lösungsmenge direkt ablesen kann. Das ist noch kein richtiges Verfahren, wie man es dem Computer geben könnte, stellt sich allerdings manchmal zum Lösen kleiner linearer Gleichungssysteme „per Hand“ als sinnvoller heraus. Bevor wir eine systematische Methode dafür angeben, wollen wir zunächst ein Beispiel davon betrachten, wie das grundsätzlich funktionieren kann.

Beispiel. Wir betrachten das lineare Gleichungssystem

$$\begin{cases} 2x_1 - 3x_2 + 10x_3 = -2, \\ x_1 - 2x_2 + 3x_3 = -2, \\ -x_1 + 3x_2 + x_3 = 4. \end{cases}$$

Nun wenden wir die elementaren Zeilentransformationen an, um dieses Gleichungssystem möglichst weit zu vereinfachen.

$$\begin{array}{l} \\ \\ \xrightarrow{(Z1)I \leftrightarrow II} \\ \\ \xrightarrow{(Z3)II - 2 \cdot I, III + I} \end{array} \begin{cases} 2x_1 - 3x_2 + 10x_3 = -2, \\ x_1 - 2x_2 + 3x_3 = -2, \\ -x_1 + 3x_2 + x_3 = 4 \\ \\ x_1 - 2x_2 + 3x_3 = -2, \\ 2x_1 - 3x_2 + 10x_3 = -2, \\ -x_1 + 3x_2 + x_3 = 4 \\ \\ x_1 - 2x_2 + 3x_3 = -2, \\ x_2 + 4x_3 = 2, \\ x_2 + 4x_3 = 2 \end{cases}$$

$$\underbrace{(Z3)_{III-II}} \rightarrow \begin{cases} x_1 - 2x_2 + 3x_3 = -2, \\ x_2 + 4x_3 = 2, \\ 0 = 0 \end{cases}$$

Nun ist die letzte Gleichung eine Aussage, die unabhängig von den Werten (x_1, x_2, x_3) , die wir einsetzen, erfüllt sein wird. Somit haben wir nur die anderen zwei Gleichungen für die restlichen drei Variablen zu erfüllen. Nun gibt es zwei Möglichkeiten, weiter zu verfahren: Wir fangen mit der einfacheren an, in der wir nun die Gleichungen durch Rückeinsetzen auflösen. Dabei ist es nun deutlich einfacher als vorher, denn jetzt hat jede Gleichung die Form $x_j + \text{weitere Terme} = \text{reelle Zahl}$, und wir können die Gleichungen von unten nach oben „aufrollen“, durch die stufige Form des Gleichungssystems. Dabei ist x_3 eine „freie“ Variable, da sie nicht am Anfang einer solchen Gleichung vorkommt. Das ist kein gänzlich präziser Begriff, aber der Gedanke ist der folgende: Wir können die restlichen Gleichungen nach x_3 auflösen und für x_3 dann eine beliebige reelle Zahl einsetzen. Im Allgemeinen kann ein Gleichungssystem mehrere freie Variablen haben. Also:

$$\begin{aligned} & \begin{cases} x_1 - 2x_2 + 3x_3 = -2, \\ x_2 + 4x_3 = 2 \quad | - 4x_3, \\ 0 = 0 \end{cases} \\ \Leftrightarrow & \begin{cases} x_1 - 2x_2 + 3x_3 = -2 \quad | + 2x_2 - 3x_3, \\ x_2 = 2 - 4x_3, \\ 0 = 0 \end{cases} \\ \Leftrightarrow & \begin{cases} x_1 = 2x_2 - 3x_3 - 2, \\ x_2 = 2 - 4x_3, \\ 0 = 0 \end{cases} \\ \Leftrightarrow & \begin{cases} x_1 = 2(2 - 4x_3) - 3x_3 - 2, \\ x_2 = 2 - 4x_3, \\ 0 = 0 \end{cases} \\ \Leftrightarrow & \begin{cases} x_1 = 2 - 11x_3, \\ x_2 = 2 - 4x_3, \\ 0 = 0 \end{cases} \end{aligned}$$

Dabei haben wir im vorletzten Schritt die zweite Gleichung in die erste eingesetzt und im letzten Schritt die erste Gleichung vereinfacht. Hierbei haben wir das Ziel verfolgt, die Variablen weitgehend zu eliminieren: Wir haben erreicht, dass x_1 nur in der ersten Gleichung vorkommt, und danach auch x_2 aus der dritten Gleichung eliminiert. Dabei ist gleichzeitig auch x_3 aus der dritten Gleichung eliminiert worden.

Nun gilt es, die Lösungsmenge in möglichst einfacher Form aufzuschreiben. Wir haben soeben gezeigt, dass die Lösungen des Gleichungssystems genau die Tripel der Form $(2 - 11x_3, 2 - 4x_3, x_3)$ sind, wobei x_3 eine beliebige reelle Zahl ist. Also hat das lineare Gleichungssystem, mit dem wir gestartet sind, unendlich viele Lösungen, die wir nun in der parametrisierten Form angeben haben. Die Lösungsmenge des ursprünglichen linearen Gleichungssystems ist die Menge

$$\{(2 - 11\lambda, 2 - 4\lambda, \lambda) \mid \lambda \in \mathbb{R}\}.$$

(Hierbei haben wir den Parameter nur aus Bequemlichkeit umbenannt.)

Diese parametrisierte Form ermöglicht es uns, eine beliebige Anzahl von Lösungen ohne Aufwand zu produzieren, indem wir konkrete Werte für λ auswählen.

Bei den Umformungen konnte man beobachten, dass während der Anwendungen der Zeilentransformationen die Variablennamen eigentlich keine Rolle gespielt haben und nur als Platzhalter dienten. Wir wollen diese in Zukunft tatsächlich weglassen. Dazu haben wir zunächst die folgende Definition.

Definition 1.4. Eine $k \times n$ -**Matrix** mit reellen Einträgen ist eine rechteckige Anordnung von reellen Zahlen mit k Zeilen und n Spalten.

Im Allgemeinen ist eine Matrix ein relativ einfaches Objekt.

Beispiel. Ein Beispiel für eine 2×3 -Matrix ist gegeben durch

$$\begin{pmatrix} \frac{7}{17} & -12 & 0,12 \\ \pi^2 & -3 & 1 \end{pmatrix}$$

Wir wollen die Nummerierung der Einträge festlegen. Der Eintrag an der Stelle (i, j) soll in der i -ten Zeile und in der j -ten Spalte sein, ähnlich wie bei den linearen Gleichungssystemen. So wäre in dieser Matrix der Eintrag $(2, 1)$ durch π^2 gegeben, während der Eintrag $(1, 2)$ die Zahl -12 ist. Als Merkspruch kann beispielsweise „Zeilen zuerst, Spalten später“ verwendet werden.

In allgemeiner Form würde eine 2×3 -Matrix wie folgt aussehen:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix},$$

wobei $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}$ reelle Zahlen sind.

Ein weiteres Beispiel einer Matrix ist ein ausgefülltes Sudoku; hierbei handelt es sich um eine 9×9 -Matrix.

Eine Matrix ähnelt einem Array von Arrays in vielen Programmiersprachen, wobei wir die Zusatzbedingung haben, dass alle „internen“ Arrays gleich lang sind.

Als nächstes wollen wir nochmal die Verwendung von Matrizen im Zusammenhang mit linearen Gleichungssystemen formalisieren. Die Idee dabei ist, dass wir die Koeffizienten in eine Matrix schreiben und die Variablen dabei weglassen.

Definition 1.5. Sei ein lineares Gleichungssystem in Standardform,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2, \\ \vdots \\ a_{k1}x_1 + a_{k2}x_2 + a_{k3}x_3 + \dots + a_{kn}x_n = b_k, \end{cases}$$

vorgegeben. Die **erweiterte Koeffizientenmatrix** dieses linearen Gleichungssystems ist gegeben durch

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \dots & a_{kn} & b_k \end{array} \right).$$

Die erweiterte Koeffizientenmatrix hat also so viele Zeilen, wie das Gleichungssystem Gleichungen hatte. Die Anzahl der Spalten ist um eins größer als die Anzahl der Variablen. Dabei bezieht sich „erweitert“ darauf, dass wir nicht nur die Koeffizienten a_{ij} in der Matrix vermerken, sondern auch die Konstanten b_1, \dots, b_k .

Bemerkung. Die elementaren Zeilentransformationen können mit der erweiterten Koeffizientenmatrix eines linearen Gleichungssystems anstatt mit dem linearen Gleichungssystem selbst durchgeführt werden. Dabei werden Zeilen spaltenweise addiert bzw. mit Konstanten multipliziert.

Beispiel. Wir notieren nochmal zum Vergleich die Lösung des vorherigen linearen Gleichungssystems in Termen der erweiterten Koeffizientenmatrix. Während wir bei Gleichungssystemen sowohl die Äquivalenzpfeile als auch die geschwungenen Pfeile verwenden können, ergeben Äquivalenzpfeile bei der Arbeit mit Matrizen keinen Sinn, da Matrizen weder wahr noch falsch sein können und somit keine Aussagen sind. (Hingegen ist jede Gleichung in Variablen x_1, \dots, x_n eine Aussage, die bei jedem Tupel $(x_1, \dots, x_n) \in \mathbb{R}^n$ entweder wahr oder falsch ist, und ein Gleichungssystem ist die Und-Verknüpfung dieser einzelnen Aussagen.)

Das lineare Gleichungssystem

$$\begin{cases} 2x_1 - 3x_2 + 10x_3 = -2, \\ x_1 - 2x_2 + 3x_3 = -2, \\ -x_1 + 3x_2 + x_3 = 4 \end{cases}$$

hat die erweiterte Koeffizientenmatrix

$$\left(\begin{array}{ccc|c} 2 & -3 & 10 & -2 \\ 1 & -2 & 3 & -2 \\ -1 & 3 & 1 & 4 \end{array} \right).$$

Wir können nun damit dieselben Operationen durchführen wie mit dem linearen Gleichungssystem. Wir notieren einige dieser Schritte:

$$\begin{array}{c} \left(\begin{array}{ccc|c} 2 & -3 & 10 & -2 \\ 1 & -2 & 3 & -2 \\ -1 & 3 & 1 & 4 \end{array} \right) \xrightarrow{(Z1)I \leftrightarrow II} \left(\begin{array}{ccc|c} 1 & -2 & 3 & -2 \\ 2 & -3 & 10 & -2 \\ -1 & 3 & 1 & 4 \end{array} \right) \\ \xrightarrow{(Z3)II - 2 \cdot I, III + I} \left(\begin{array}{ccc|c} 1 & -2 & 3 & -2 \\ 0 & 1 & 4 & 2 \\ 0 & 1 & 4 & 2 \end{array} \right). \end{array}$$

Wir gehen Teile vom letzten Schritt in etwas mehr Detail durch: Multipliziert man die erste Zeile mit 2, so erhält man die Zeile $(2 \ -4 \ 6 \ | \ -4)$. Subtrahiert man diese nun von der zweiten Zeile, so erhält man

$$(2 - 2 \quad -3 + 4 \quad 10 - 6 \quad | \quad -2 + 4),$$

was genau die neue Zeile

$$(0 \ 1 \ 4 \ | \ 2)$$

ergibt.

Wir bemerken ferner, dass wir die neue erweiterte Koeffizientenmatrix zurück in ein lineares Gleichungssystem übersetzen können und dabei genau das lineare Gleichungssystem bekommen, dass wir nach den entsprechenden Zeilentransformationen aus dem ursprünglichen bekommen haben, nämlich

$$\begin{cases} x_1 - 2x_2 + 3x_3 = -2, \\ \quad \quad x_2 + 4x_3 = 2, \\ \quad \quad x_2 + 4x_3 = 2. \end{cases}$$

Es sei ferner noch angemerkt, dass die erweiterten Koeffizientenmatrizen nur für lineare Gleichungssysteme in Standardform einen angemessenen Ersatz bieten; sobald wir mit dem Rückeinsetzen beginnen möchten, müssen wir zurück zu dem linearen Gleichungssystem selbst wechseln. Das wollen wir auch tun, nachdem wir das lineare Gleichungssystem, wie vorhin schon erwähnt, in eine einfache Form zum Rückeinsetzen gebracht haben.

Wie diese einfache Form auszusehen hat, lässt sich vielleicht bereits anhand der Beispiele erahnen. Am liebsten würden wir die erste Zeile mit einer 1 anfangen lassen, und dann alle Vorkommnisse von der Variable x_1 in Zeilen darunter eliminieren - in der erweiterten Koeffizientenmatrix wollen wir also

lauter 0-Einträge unter der 1 in der ersten Spalte haben. In der zweiten Zeile wollen wir in der zweiten Spalte mit dem x_2 (also dem Eintrag 1) anfangen, und danach in darauffolgenden Gleichungen x_2 eliminieren, und so weiter. Dabei würden wir gerne eine stufige Form erreichen, die wie folgt aussieht:

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

Dabei stehen die *-Symbole für beliebige, eventuell unterschiedliche reelle Zahlen, und die Größe der Matrix ist nur zur besseren Veranschaulichung fest gewählt.

Nun wird sich eine solche Form nicht für jedes Gleichungssystem erreichen lassen; allerdings können wir immer etwas ähnliches erreichen; die „stufige“ Form sieht in etwa wie folgt aus:

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

Wir wollen nun eine präzise Definition von solcher „einfacher Form“ geben.

Definition 1.6. Eine Matrix ist in **Zeilenstufenform**, falls die folgenden zwei Bedingungen in dieser Matrix erfüllt sind:

- (1) In jeder Zeile ist die erste Zahl, die ungleich Null ist, eine 1 und befindet sich rechts von der ersten Zahl $\neq 0$ in der Zeile darüber,
- (2) Alle Nullzeilen befinden sich am unteren Ende der Matrix.

Die Forderung, dass die erste Zahl ungleich Null in jeder Zeile 1 ist, wird in der Literatur häufig weggelassen. Sie ist jedoch zum Lösen linearer Gleichungssysteme „per Hand“ recht praktisch. Mit der „ersten Zahl“ meinen wir hier die erste Zahl von links.

Beispiel. Wir bezeichnen hier der Einfachheit halber die Einträge der Matrix, deren Wert keine Rolle für unsere Betrachtungen spielt, mit *. Dabei kann sich natürlich hinter jedem * eine andere reelle Zahl befinden.

- Damit die Matrix

$$\left(\begin{array}{cccc} 0 & 1 & 2 & 3 \\ * & * & * & * \end{array} \right)$$

in Zeilenstufenform ist, müssen die ersten beiden Einträge der letzten Zeile auf jeden Fall 0 sein, damit der erste Eintrag $\neq 0$ in der zweiten

Zeile rechts von der ersten 1 in der ersten Zeile ist. Danach gibt es allerdings unterschiedliche Optionen, wie diese Matrix aussehen kann, um in Zeilenstufenform zu sein:

$$\begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & * \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{oder} \quad \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- Damit die Matrix

$$\begin{pmatrix} 1 & * & * \\ 0 & 1 & * \\ 0 & 0 & 1 \\ * & * & * \end{pmatrix}$$

in Zeilenstufenform ist, muss die letzte Zeile eine Nullzeile sein. Denn darin muss laut Definition der erste Eintrag $\neq 0$ rechts von der 1 in der Zeile darüber stehen; da es rechts davon allerdings keine Einträge gibt, darf also in der letzten Zeile keine Zahl $\neq 0$ vorkommen.

- In dem vorher betrachteten linearen Gleichungssystem haben wir die elementaren Zeilentransformationen verwendet, um die erweiterte Koeffizientenmatrix auf die Form

$$\left(\begin{array}{ccc|c} 1 & -2 & 3 & -2 \\ 0 & 1 & 4 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

zu bringen, und danach mit Rückeinsetzen begonnen. Diese Matrix ist in Zeilenstufenform.

Das letzte Beispiel illustriert unser allgemeines Vorgehen zum Lösen eines linearen Gleichungssystems, gewissermaßen die erste Version des Gauß'schen Eliminationsverfahrens: Wir wenden so lange elementare Zeilentransformationen auf die erweiterte Koeffizientenmatrix eines linearen Gleichungssystems an, bis diese in Zeilenstufenform ist. Danach übersetzen wir zurück in ein lineares Gleichungssystem und bestimmen durch Rückeinsetzen die Lösungsmenge.

Wir demonstrieren dieses Vorgehen an einem weiteren Beispiel.

Beispiel. Wir suchen die Lösungen des linearen Gleichungssystems

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1, \\ 4x_1 + 5x_2 + 6x_3 = 2, \\ 7x_1 + 8x_2 + 9x_3 = 3, \\ 5x_1 + 7x_2 + 9x_3 = 3. \end{cases}$$

Dieses lineare Gleichungssystem hat als erweiterte Koeffizientenmatrix die Matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 2 \\ 7 & 8 & 9 & 3 \\ 5 & 7 & 9 & 3 \end{array} \right).$$

Wir wenden nun elementare Zeilentransformationen auf diese Matrix an, um sie in Zeilenstufenform zu bringen.

$$\begin{array}{ccc} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 4 & 5 & 6 & 2 \\ 7 & 8 & 9 & 3 \\ 5 & 7 & 9 & 3 \end{array} \right) & \xrightarrow{\text{II}-4\cdot\text{I, III}-7\cdot\text{I, IV}-5\cdot\text{I}} & \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -2 \\ 0 & -6 & -12 & -4 \\ 0 & -3 & -6 & -2 \end{array} \right) \\ \xrightarrow{\text{III}-2\cdot\text{II, IV}-\text{II}} & \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & -3 & -6 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) & \xrightarrow{\text{II}/(-3)} & \left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & \frac{2}{3} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \end{array}$$

Diese Matrix ist nun in Zeilenstufenform. Die Nullzeilen liefern die Gleichung $0 = 0$, die ja für jedes Tupel (x_1, x_2, x_3) erfüllt ist. Wir lassen diese also Weg. Durch Rückeinsetzen brauchen wir also nur noch das folgende lineare Gleichungssystem zu lösen:

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1, \\ x_2 + 2x_3 = \frac{2}{3}. \end{cases}$$

Hierbei sehen wir bereits, dass die Variable x_3 wieder frei wählbar sein wird und dass wir die Variablen x_1, x_2 durch die Variable x_3 ausdrücken können. Insbesondere sehen wir schon, dass die Lösungsmenge unendlich sein wird. Wir bestimmen diese nun explizit:

$$\begin{aligned} & \begin{cases} x_1 + 2x_2 + 3x_3 = 1, \\ x_2 + 2x_3 = \frac{2}{3}. \end{cases} \\ \Leftrightarrow & \begin{cases} x_1 + 2x_2 + 3x_3 = 1, \\ x_2 = \frac{2}{3} - 2x_3 \end{cases} \\ \Leftrightarrow & \begin{cases} x_1 = 1 - 2\left(\frac{2}{3} - 2x_3\right) - 3x_3, \\ x_2 = \frac{2}{3} - 2x_3 \end{cases} \\ \Leftrightarrow & \begin{cases} x_1 = -\frac{1}{3} + x_3, \\ x_2 = \frac{2}{3} - 2x_3. \end{cases} \end{aligned}$$

Hierbei haben wir im ersten Schritt die zweite Gleichung nach x_2 aufgelöst, dann als nächstes die erste Gleichung nach x_1 aufgelöst und den bereits ermittelten Ausdruck für x_2 in Termen von x_3 in die erste Gleichung eingesetzt, und zu guter Letzt haben wir diese Gleichung vereinfacht.

Nun können wir also die Lösungsmenge des ursprünglichen Gleichungssystems in parametrisierter Form angeben. Diese Menge ist

$$\left\{ \left(-\frac{1}{3} + \lambda, \frac{2}{3} - 2\lambda, \lambda \right) \mid \lambda \in \mathbb{R} \right\}.$$

Die Beispiele legen nahe, dass, sobald die linke Seite der erweiterten Koeffizientenmatrix in der Zeilenstufenform ist, tatsächlich ein guter Punkt erreicht ist, um zum Rückeinsetzen zu übergehen. Als nächstes beschäftigen wir uns mit der Frage, ob jede Matrix in Zeilenstufenform gebracht werden kann. Diese wird durch den folgenden Satz beantwortet.

Satz 1.7. *Jede Matrix kann durch endlich viele Zeilentransformationen in Zeilenstufenform gebracht werden.*

Wir wollen keinen präzisen Beweis von diesem Satz geben; allerdings kann man den folgenden Algorithmus als das Herzstück des Beweises betrachten. Dieser Algorithmus liefert eine systematische Methode, um eine Matrix in Zeilenstufenform zu bringen. Wir geben zuerst die Idee von dem Algorithmus und danach einen Pseudo-Code, der den Algorithmus genauer beschreibt. Wir werden allerdings weder beweisen, dass der Algorithmus stets terminiert, noch, dass er das gewünschte Ergebnis liefert. Allerdings sind beide Punkte recht plausibel.

Hat man also eine Matrix, etwa von der Form

$$\begin{pmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{pmatrix}$$

vorgegeben und möchte diese in Zeilenstufenform bringen, so würde man mit dem Eintrag in der linken oberen Ecke (also an der Stelle $(1, 1)$) anfangen, und versuchen diesen Eintrag zu einer 1 zu machen, indem wir die erste Zeile durch den Eintrag an der Stelle $(1, 1)$ teilen. Das kann allerdings misslingen, und zwar dann, wenn der Eintrag an der Stelle $(1, 1)$ Null sein sollte. In diesem Fall würde man versuchen, um eine 1 an der Stelle $(1, 1)$ zu erhalten, zunächst die oberste Zeile mit einer Zeile zu vertauschen, in der der erste Eintrag nicht Null ist. Das erklärt, warum die elementare Zeilentransformation (Z1) überhaupt vonnöten ist. Allerdings kann es auch passieren, dass wir einen solchen Tausch nicht durchführen können, und zwar genau in dem Fall, dass alle Einträge in der ersten Spalte Null sind. Das wäre bei einer erweiterten Koeffizientenmatrix ungewöhnlich (wenn auch nicht verboten), da

die erste Variable in keiner Gleichung vorkommen würde. Wäre es die einzige Situation, in der das Problem auftritt, so könnten wir das vernachlässigen. Allerdings wollen wir das Verfahren gleich rekursiv anwenden, und in weiteren Rekursionsschritten wird das Problem auftreten können, ohne dass man es in der ursprünglichen Form der Matrix hätte sehen können. Daher müssen wir den Fall einer Nullspalte in der ersten Spalte zulassen. In diesem Fall würde man die Nullspalte außer Acht lassen und das Gauß-Verfahren nun auf die Matrix ohne die erste Spalte anwenden, also auf den im Bild blau markierten Teil der Matrix:

$$\begin{pmatrix} 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \end{pmatrix}.$$

Nun kommen wir zurück zu dem „normalen“ Fall, in dem wir in der linken oberen Ecke eine 1 erzeugen konnten. Weiter müssen wir, um die Zeilenstufenform zu erreichen, alle anderen Einträge der ersten Spalte eliminieren, also erreichen, dass diese Einträge 0 sind. Das gewährleisten wir, indem wir jeweils Vielfache der ersten Zeile von der jeweiligen Zeile abziehen; der Faktor ist dabei durch den ersten Eintrag dieser jeweiligen Zeile gegeben. Dieser Schritt ist unproblematisch, und wir erreichen dann die Form

$$\begin{pmatrix} 1 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \end{pmatrix}.$$

Um jetzt weiterzumachen, konzentriert man sich auf die Matrix, die sich durch entfernen der ersten Spalte und ersten Zeile ergibt und die in der folgenden schematischen Darstellung rot markiert ist, und fängt da von vorne an:

$$\begin{pmatrix} 1 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \\ 0 & * & * & * & * \end{pmatrix}.$$

Nun wollen wir, um dies etwas genauer festzuhalten, einen Pseudo-Code angeben, der das Gauß'sche Eliminationsverfahren (oder vielmehr eine Version hiervon) beschreibt. Wie bereits erwähnt, wird das Verfahren rekursiv sein. Wir werden zwei Parameter i und j verwenden, wobei i für die gerade betrachtete Zeilennummer und j für die Spaltennummer stehen wird. Zu Beginn werden beide auf 1 gesetzt, da wir in der linken oberen Ecke anfangen. Wir nehmen an, dass unsere Matrix k Zeilen und n Spalten hat. Das Verfahren soll stoppen, wenn das Ende der Matrix erreicht ist. Den Eintrag der Matrix an der Stelle (i, j) bezeichnen wir mit a_{ij} .

Algorithmus. $i = 1$ $j = 1$ *Gauss*(i, j):Ist $i = k$ oder $j = n + 1$: EndeWenn $a_{ij} = 0$ ist: Suche ein $r > i$ mit $a_{rj} \neq 0$.

- Falls ein solches r existiert, tausche die Zeilen r und i .
- Sonst: Führe *Gauss*($i, j+1$) durch.

Dann:

- Teile die i -te Zeile durch a_{ij} .
- Für jedes $l > i$, ziehe von der Zeile l das a_{lj} -Vielfache der Zeile i ab.
- Führe *Gauss*($i+1, j+1$) durch.

(Die Form des Verfahrens ist an [1] angelehnt.)

Dieses Verfahren ist insofern wichtig, dass es eine allgemeingültige Methode zum Erreichen der Zeilenstufenform liefert. Allerdings ist es in mehreren Aspekten nicht optimal. Für eine Durchführung durch einen Menschen kann es durchaus sinnvoller sein, sich nicht an die feste Reihenfolge zu halten, sondern die elementare Transformationen so durchzuführen, dass die Zahlen möglichst einfach zu handhaben sind, und beispielsweise unnötige Entstehung von Nennern vermeidet.

Bei der Implementierung entstehen allerdings andere Probleme, die mit der Form zusammenhängen, in der die Zahlen im Computer üblicherweise gespeichert werden, worauf wir zu einem späteren Zeitpunkt etwas genauer eingehen. Jedoch sei schon jetzt darauf hingewiesen, dass eines der Probleme in der „bösen“ Abfrage der Form $\text{Zahl} = 0$ besteht. Ein weiteres Problem ist der Genauigkeitsverlust bei Divisionen. Es gibt mehrere Methoden, um dieses grundlegende Verfahren für numerische Anwendungen zu verbessern. Eine recht allgemeine Methode sieht vor, dass man (fast) immer vor den Divisionen einen Zeilentausch durchführt, um nur durch gewisse Zahlen teilen zu müssen. Ferner gibt es Methoden, das Verfahren zu verbessern, falls man schon zu Anfang gewisse Annahmen über die Matrix trifft.

Wir wollen nun das Verfahren nun nochmal an einem Beispiel sehen, bei dem einige der Ausnahmen vorkommen, die in der obigen Behandlung erwähnt wurden. Wir verfolgen hierbei insbesondere das Ziel, zunächst den linken Teil der erweiterten Koeffizientenmatrix auf die Zeilenstufenform zu bringen, auch wenn in dem vorliegenden Fall schon vorher auf die Lösungen geschlossen werden können.

Beispiel. Wir suchen die Lösungsmenge von dem linearen Gleichungssystem, das zu der folgenden erweiterten Koeffizientenmatrix gehört:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 2 \\ -1 & -2 & -2 & 2 & 1 & 1 \\ 2 & 4 & 3 & -1 & 0 & -1 \\ 1 & 2 & 2 & -2 & 1 & 1 \end{array} \right)$$

Wir führen nun die elementaren Zeilentransformationen durch und notieren uns dabei die Schritte:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 2 \\ -1 & -2 & -2 & 2 & 1 & 1 \\ 2 & 4 & 3 & -1 & 0 & -1 \\ 1 & 2 & 2 & -2 & 1 & 1 \end{array} \right) \xrightarrow{\text{II+I, III-2I, IV-I}} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 2 \\ 0 & 0 & -1 & 3 & 2 & 3 \\ 0 & 0 & 1 & -3 & -2 & -5 \\ 0 & 0 & 1 & -3 & 0 & -1 \end{array} \right)$$

Hier haben wir genau die „Ausnahmesituation“ vorliegen, die wir vorher diskutiert haben: In der zweiten Ausführung der Rekursion wenden wir den Algorithmus auf den rot markierten Teil dieser Matrix:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 2 \\ 0 & 0 & -1 & 3 & 2 & 3 \\ 0 & 0 & 1 & -3 & -2 & -5 \\ 0 & 0 & 1 & -3 & 0 & -1 \end{array} \right).$$

In dieser Matrix stehen nun in der ersten Spalte nur Nullen. Also machen wir mit dem blau markierten Teil dieser Matrix weiter:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 2 \\ 0 & 0 & -1 & 3 & 2 & 3 \\ 0 & 0 & 1 & -3 & -2 & -5 \\ 0 & 0 & 1 & -3 & 0 & -1 \end{array} \right).$$

Wir erhalten also:

$$\begin{aligned} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 2 \\ 0 & 0 & -1 & 3 & 2 & 3 \\ 0 & 0 & 1 & -3 & -2 & -5 \\ 0 & 0 & 1 & -3 & 0 & -1 \end{array} \right) & \xrightarrow{\text{II}/(-1)} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 2 \\ 0 & 0 & 1 & -3 & -2 & -3 \\ 0 & 0 & 1 & -3 & -2 & -5 \\ 0 & 0 & 1 & -3 & 0 & -1 \end{array} \right) \\ & \xrightarrow{\text{III-II, IV-II}} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 2 \\ 0 & 0 & 1 & -3 & -2 & -3 \\ 0 & 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 2 & 2 \end{array} \right) \\ & \xrightarrow{\text{III} \leftrightarrow \text{IV}} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 2 \\ 0 & 0 & 1 & -3 & -2 & -3 \\ 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & -2 \end{array} \right) \end{aligned}$$

$$\begin{array}{c} \\ \\ \xrightarrow{\text{III}/2} \end{array} \left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 2 \\ 0 & 0 & 1 & -3 & -2 & -3 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -2 \end{array} \right)$$

Nun ist die linke Seite der Matrix in Zeilenstufenform. Dabei hatten wir nochmal eine Nullspalte bei der Betrachtung der Spalte 4, und eine 0 in der Spalte 5, die allerdings „weggetauscht“ werden konnte.

Nun kommen wir zum Ablesen der Lösungen in diesem Fall. Die letzte Zeile übersetzt sich in die Gleichung

$$0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 = -2,$$

oder, äquivalent dazu, $0 = -2$. Diese Aussage ist allerdings immer falsch, unabhängig von dem Tupel reeller Zahlen $(x_1, x_2, x_3, x_4, x_5)$, das wir in das Gleichungssystem einsetzen. Es kann also kein solches Tupel geben, das alle Gleichungen zu wahren Aussagen macht: Die Aussage $0 = -2$ ist „nicht zu retten“. Die Lösungsmenge des ursprünglichen Gleichungssystems ist also ebenfalls leer. Sie enthält insbesondere 0 Elemente.

Das Argument lässt sich immer anwenden, wenn wir eine Nullzeile auf der linken Seite der Matrix haben, der gegenüber auf der rechten Seite eine Zahl $\neq 0$ steht. An dieser Stelle kann man beim Lösen des linearen Gleichungssystems stets aufhören, da bereits klar ist, dass das Gleichungssystem keine Lösungen hat.

Beispiel. Wir wollen die Rechnungen aus dem letzten Beispiel nochmal ausnutzen, und fragen diesmal nach den Lösungen des homogenen linearen Gleichungssystems

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 0 \\ -1 & -2 & -2 & 2 & 1 & 0 \\ 2 & 4 & 3 & -1 & 0 & 0 \\ 1 & 2 & 2 & -2 & 1 & 0 \end{array} \right).$$

Dieses Gleichungssystem hat also dieselbe linke Seite, die „reine“ Koeffizientenmatrix, wie im Beispiel zuvor, und unterscheidet sich nur in der rechten Seite von dem vorherigen Beispiel. Wir bemerken, dass die Operationen, die man durchführen muss, um die linke Seite der Matrix in Zeilenstufenform zu bringen, nicht von der rechten Seite abhängen. Will man also mehrere unterschiedliche Gleichungssysteme mit derselben linken Seite auf einmal lösen, so kann man die unterschiedlichen rechten Spalten (klar abgetrennt, um Verwirrung zu vermeiden) nebeneinander schreiben und die elementare Umformungen mit all diesen Spalten gleichzeitig durchführen.

Wir müssten uns in unserer Situation auch nur überlegen, welchen Effekt dieselben elementaren Zeilentransformationen wie im vorherigen Beispiel auf der rechten Seite haben werden.

Man stellt allerdings schnell fest, dass elementare Zeilentransformationen in einer Spalte, in der nur Nullen vorkommen, keinerlei Effekt haben: Multipliziert man eine der Nullen mit einer reellen Zahl, so ist sie nach wie vor 0, und genauso gilt, dass alle Summen der Form $0 + c \cdot 0$ ebenfalls 0 sind. Daran merkt man, dass homogene lineare Gleichungssysteme besondere Eigenschaften haben. Wir werden diese zu einem späteren Zeitpunkt genauer untersuchen.

Wenn man die linke Seite also in Zeilenstufenform gebracht hat, so hat man die erweiterte Koeffizientenmatrix vereinfacht zu:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & -3 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Das übersetzt sich in das lineare Gleichungssystem

$$\begin{cases} x_1 + 2x_2 + x_3 + x_4 + x_5 = 0, \\ x_3 - 3x_4 - 2x_5 = 0, \\ x_5 = 0. \end{cases}$$

Hier stellen wir wieder fest, dass wir manche Variablen durch andere ausdrücken können, die wiederum ohne Einschränkungen frei gewählt werden können. Dabei beobachten wir, dass diejenigen Variablen, die zu keiner Stufe „gehören“, frei gewählt werden können, während die Variablen x_1, x_3, x_5 jeweils zu einer Spalte mit einer Stufe gehören und durch die anderen Variablen festgelegt sind.

Wir lösen das Gleichungssystem nun so auf, dass wir x_1 und x_3 durch x_2 und x_4 ausdrücken.

$$\begin{cases} x_1 + 2x_2 + x_3 + x_4 + x_5 = 0, \\ x_3 - 3x_4 - 2x_5 = 0, \\ x_5 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = -2x_2 - x_3 - x_4, \\ x_3 = 3x_4, \\ x_5 = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x_1 = -2x_2 - 4x_4, \\ x_3 = 3x_4, \\ x_5 = 0. \end{cases}$$

Alle Lösungstupel des homogenen linearen Gleichungssystems, mit dem wir gestartet sind, sind also von der Form

$$(-2x_2 - 4x_4, x_2, 3x_4, x_4, 0),$$

wobei die Variablen x_2 und x_4 frei gewählt werden können.

Die Lösungsmenge in der parametrisierten Form kann nun wie folgt angegeben werden:

$$\{(-2\lambda - 4\mu, \lambda, 3\mu, \mu, 0) \in \mathbb{R}^5 \mid \lambda, \mu \in \mathbb{R}\}.$$

Wir bemerken, dass diese Lösungsmenge unendlich ist; ferner ist sie anders als die bisher betrachteten unendlichen Lösungsmengen in parametrisierter Form, denn sie hängt von zwei Parametern ab. Wir werden die Frage, wie viel Parameter man zur Beschreibung der Lösungsmenge eines linearen Gleichungssystems braucht, später nochmal aufgreifen.

Ablesen der Lösungsmenge

Nun haben wir festgestellt, dass wir zu jedem linearen Gleichungssystem seine erweiterte Koeffizientenmatrix betrachten können und die linke Seite dieser in Zeilenstufenform bringen können. Nun wollen wir uns etwas systematischer der Frage widmen, wie die Lösungsmenge, die man durch Rückeinsetzen in präziser Form erhält, allgemeiner beschreiben kann.

Sei also ein lineares Gleichungssystem vorgegeben, in dessen erweiterter Koeffizientenmatrix die linke Seite auf die Zeilenstufenform gebracht wurde. Zunächst beobachten wir, dass es höchstens eine Stufe pro Spalte geben kann, sodass die Anzahl der Stufen kleiner oder gleich ist der Anzahl der Variablen. Wir unterscheiden nun drei Fälle.

Fall 1: Es kommt in der erweiterten Koeffizientenmatrix eine Zeile der Form

$$(0 \ 0 \ \dots \ 0 \ | \ r)$$

vor, wobei r eine reelle Zahl ist, die ungleich Null ist.

Wie wir vorher bereits erörtert haben, entspricht das der Gleichung $0 = r$, die stets falsch ist, da $r \neq 0$ angenommen wurde. Somit hat das Gleichungssystem in diesem Fall keine Lösungen, die Lösungsmenge ist leer.

Fall 2: Nun nehmen wir an, eine Zeile wie im ersten Fall kommt in der erweiterten Koeffizientenmatrix nicht vor. In diesem zweiten Fall wollen wir annehmen, dass die linke Seite der erweiterten Koeffizientenmatrix genauso viele Stufen hat wie wir Variablen haben. Die erweiterte Koeffizientenmatrix sieht schematisch wie folgt aus:

$$\left(\begin{array}{cccc|cc} 1 & * & * & \dots & * & * \\ 0 & 1 & * & \dots & * & * \\ 0 & 0 & 1 & \dots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & * \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 \end{array} \right)$$

Die Matrix besteht also aus einem oberen Teil, der die ersten n Zeilen umfasst, wobei n die Anzahl der Variablen ist, und der auf der linken Seite dementsprechend quadratisch ist (also gleich viele Zeilen wie Spalten hat), und auf der Diagonale lauter Einsen hat, während darunter jeder Eintrag 0 ist. Unter diesem quadratischen Teil dürfen auf der linken Seite nur Nullzeilen stehen, wie man sich leicht anhand der Definition der Zeilenstufenform klarmachen kann. Da wir die Situation des ersten Falls ausgeschlossen haben, stehen also auf der rechten Seite dieser Zeile ebenfalls ausschließlich Nullen, und dieser untere Bereich kann weggelassen werden, da alle zugehörigen Gleichungen von der Form $0 = 0$ sind und stets erfüllt sind.

Widmen wir unsere Aufmerksamkeit also dem oberen Teil dieser Matrix. Die letzte Gleichung hiervon ist besonders einfach; sie hat die Form $x_n = c_n$, wobei c_n die feste reelle Zahl auf der rechten Seite der n -ten Zeile ist. Nun lässt sich die vorletzte Gleichung, die nun von der Form $x_{n-1} + * \cdot c_n = c_{n-1}$ ist, da wir $x_n = c_n$ einsetzen konnten, durch eine eindeutige reelle Zahl lösbar. Wir haben also genau eine Möglichkeit für x_n , daraus können wir den eindeutigen Wert der Variable x_{n-1} in jedem Lösungstupel ermitteln, und man kann sich leicht vorstellen (auch wenn das kein präziser Beweis ist), dass man auf diese Weise für jede Variable einen eindeutigen Wert bekommt.

Insgesamt wird in diesem Fall das lineare Gleichungssystem genau eine Lösung haben. Wir bemerken außerdem, dass dieser Fall nur dann vorkommen kann, wenn wir mindestens so viele Gleichungen wie Variablen haben.

Fall 3: Wir nehmen nun an, dass wir weder im Fall 1 noch im Fall 2 sind. Nach unserer Vorüberlegung hat das Gleichungssystem also weniger Stufen als Variablen. Wir beschreiben die Situation hierbei informell, die wir auch schon in mehreren Beispielen gesehen haben. Die Variablen, die zu den Stufe gehören, können durch die restlichen Variablen ausgedrückt werden. Die Variablen, die zu keiner Stufe gehören, können als freie Parameter gewählt werden. Da wir mindestens eine solche Variable in diesem dritten Fall haben, werden wir mindestens einen freien Parameter haben, der als eine beliebige reelle Zahl gewählt werden kann. Somit wird das lineare Gleichungssystem unendlich viele Lösungen haben.

Für lineare Gleichungssysteme gilt also insbesondere: Ein solches hat entweder 0 Lösungen oder genau eine Lösung oder unendlich viele Lösungen. Ein lineares Gleichungssystem z.B. mit genau 2 Lösungen kann nicht vorkommen. Das ist eine Sache, die durchaus besonders bei den linearen Gleichungssystemen ist; betrachtet man im Vergleich etwa eine quadratische Gleichung, so ist es durchaus ein häufiger Fall, dass diese genau zwei Lösungen hat.

Außerdem wollen wir nochmal unterstreichen, dass, während der erste und der dritte Fall stets auftreten können, der zweite Fall - nämlich eine eindeutige Lösung - nur dann vorkommen kann, wenn es mindestens so viele Gleichungen wie Variablen gibt.

2 Vektorraum \mathbb{R}^n und seine Unterräume

Als nächstes wollen wir uns der Frage zuwenden, wie die Struktur einer solchen Lösungsmenge eines linearen Gleichungssystems aussehen kann. Das ist in den Fällen, dass das Gleichungssystem entweder keine oder genau eine Lösung hat, relativ uninteressant, und bedarf keiner besonderen Untersuchung. Spannender gestaltet sich die Frage in dem Fall, dass die Lösungsmenge unendlich ist. Wie bereits angedeutet, wird ein Unterschied darin bestehen, dass die Anzahl der benötigten Parameter unterschiedlich sein kann. Außerdem wollen wir die Regelmäßigkeiten der Lösungsmenge untersuchen. Es stellt sich dabei heraus, dass der dabei entstehende mathematische Begriff auch von unabhängiger Relevanz ist.

Bevor wir darauf genauer eingehen, wollen wir uns nochmal mit der Menge aller potentiellen Lösungen, dem \mathbb{R}^n , beschäftigen. Zunächst ändern wir die Schreibweise, die wir für Tupel von \mathbb{R}^n verwenden: Von nun an schreiben wir diese in einer Spalte, also etwa

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n,$$

und nennen die Elemente von \mathbb{R}^n Vektoren oder manchmal auch Spaltenvektoren. Das ist vor allem historisch bedingt; für uns soll es aber auch unterstreichen, dass Vektoren nicht nur Ansammlungen von Zahlen sind, sondern auch Objekte, mit denen gerechnet werden kann.

Darauf wollen wir nun etwas genauer eingehen. Zunächst definieren wir Addition von Vektoren. Für Vektoren in \mathbb{R}^2 und \mathbb{R}^3 , wo Vektoren eine anschauliche Bedeutung haben, etwa als Koordinaten von Punkten oder als Ortsvektoren von Punkten, wird häufig schon Addition in der Schule behandelt, etwa in der Physik, wenn die Gesamtwirkung bei gleichzeitiger Wirkung zweier Kräfte in unterschiedliche Richtungen ermittelt werden soll (dabei verwendet man das Stichwort „Kräfteparallelogramm“). Die Vektoren werden komponentenweise addiert; das heißt, die erste Zahl vom ersten Vektor wird zu der ersten Zahl des zweiten Vektors addiert, die zweite zu der zweiten, und so weiter. Insbesondere können nun Vektoren addiert werden, die im selben

\mathbb{R}^n liegen. Nochmal in Formeln: Für Vektoren

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$$

definieren wir

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Insbesondere ist die Summe zweier Vektoren in \mathbb{R}^n wieder ein Vektor in \mathbb{R}^n . Zum Beispiel:

$$\begin{pmatrix} -1 \\ 0 \\ -2,7 \\ \frac{3}{7} \end{pmatrix} + \begin{pmatrix} 2 \\ \pi \\ 3,3 \\ 0,25 \end{pmatrix} = \begin{pmatrix} 1 \\ \pi \\ 0,6 \\ \frac{19}{28} \end{pmatrix}.$$

Hingegen ist

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 2 \\ 7 \end{pmatrix}$$

nicht definiert.

Die zweite Operation für Vektoren, die wir definieren, ist die Skalarmultiplikation. Hierbei kann ein Vektor mit einer festen reellen Zahl multipliziert werden, indem wir jede Komponente des Vektors mit dieser Zahl multiplizieren. Wenn man mit Vektoren zu tun hat, werden die reellen Zahlen häufig zur Unterscheidung „Skalare“ genannt, daher auch der Name. In Formeln heißt das für ein $\lambda \in \mathbb{R}$ und ein

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n,$$

dass

$$\lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_1 \\ \lambda \cdot x_2 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}$$

gesetzt wird. In \mathbb{R}^2 bzw. \mathbb{R}^3 ist die Intuition dahinter, dass der Vektor zwar seine Richtung behält, aber um den Faktor λ gestreckt wird.

Wir fassen nun die wichtigsten Eigenschaften der Addition von Vektoren und der Skalarmultiplikation zusammen. Diese besagen, dass viele Rechenregeln, die für reelle Zahlen gelten, auch für Vektoren gelten.

- **Assoziativität der Addition:** Sind $x, y, z \in \mathbb{R}^n$ Vektoren, so gilt:

$$(x + y) + z = x + (y + z).$$

Das heißt, wir können zunächst die Summe von x und y bilden, und dann z dazu addieren, oder zunächst die Summe von y und z bilden und diese zu x addieren, und das Ergebnis wird gleich sein.

- **Kommutativität der Addition:** Sind $x, y \in \mathbb{R}^n$ Vektoren, so gilt:

$$x + y = y + x$$

Die beiden ersten Regeln zusammen besagen, dass eine Summe von mehreren Vektoren beliebig umsortiert und geklammert werden kann, ohne das Ergebnis zu verändern.

- **Existenz eines Nullelements:** Der Vektor $\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n$ hat eine

besondere Bedeutung, ähnlich wie 0 bei den reellen Zahlen, und zwar gilt für jeden Vektor $x \in \mathbb{R}^n$:

$$\vec{0} + x = x = x + \vec{0}$$

Den Nullvektor zu einem Vektor zu addieren, ändert den Vektor nicht. (Zur Notation: Hier bezeichnen wir den Nullvektor mit $\vec{0}$, um die Unterscheidung zu $0 \in \mathbb{R}$ hervorzuheben. Später werden wir zur Vereinfachung der Notation auch einfach 0 schreiben.)

- **Existenz von Inversen:** Zu jedem Vektor $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ kann man

das „Negative“ davon bilden, also einen Vektor $-x = \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix} \in \mathbb{R}^n$,

und dieser hat die Eigenschaft

$$x + (-x) = \vec{0} = (-x) + x.$$

- **Distributivgesetze für die Skalarmultiplikation:** Das Distributivgesetz für Addition und Multiplikation der reellen Zahlen ist diejenige Rechenregel, die uns erlaubt, Klammern aufzulösen. In diesem Fall müssen wir zwei Fälle unterscheiden: Es können nämlich sowohl Skalare, mit denen wir multiplizieren, als auch Vektoren addiert werden, und da können jeweils im Grunde Klammern wie gewohnt aufgelöst werden. Um das nun in Formeln anzugeben, seien λ, μ reelle Zahlen und $x, y \in \mathbb{R}^n$ Vektoren. Dann gilt:

$$(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x \text{ und } \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y.$$

- **Assoziativgesetz für Skalarmultiplikation:** Sind $\lambda, \mu \in \mathbb{R}$ und $x \in \mathbb{R}^n$, so gilt

$$(\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x).$$

Es ist also egal, ob wir zunächst das Produkt der reellen Zahlen λ und μ bilden und dann Skalarmultiplikation mit x durchführen, oder ob wir zunächst $\mu \cdot x$ bilden und dann diesen Vektor mit dem Skalar λ multiplizieren.

- $1 \in \mathbb{R}$: Diese Rechenregel hat keinen besonderen Namen; sie soll allerdings verdeutlichen, dass dem Skalar 1 eine besondere Bedeutung bei der Skalarmultiplikation zukommt, nämlich die, dass ein Vektor durch die Skalarmultiplikation mit 1 nicht verändert wird. In Formeln heißt das: Für jedes $x \in \mathbb{R}^n$ ist

$$1 \cdot x = x.$$

Wir führen diese Rechenregeln auch deshalb auf, weil es auch andere Objekte als \mathbb{R}^n gibt, genannt **Vektorräume**, die denselben Rechenregeln genügen und ebenfalls in der linearen Algebra studiert werden können. Bei uns werden allerdings die (spezielleren) Vektorräume \mathbb{R}^n die zentrale Rolle spielen.

Es soll noch angemerkt werden, dass es kein „vernünftiges“ Produkt für Vektoren existiert. Zwar kann man da verschiedene Operationen definieren, jedoch hat keine alle gewünschten Eigenschaften für ein Produkt. Das Skalarprodukt zweier Vektoren, das häufig aus der Schule bekannt ist, hat einen grundsätzlichen Unterschied zu der Summe zweier Vektoren: Bildet man das Skalarprodukt zweier Vektoren, so ist das Ergebnis eine reelle Zahl, also ein Skalar, und kein Vektor. Das Skalarprodukt ist ein wichtiger Begriff, den wir später nochmal diskutieren werden. Für den momentan kann man sich allerdings merken: Der Produktvektor zweier Vektoren ist nicht definiert.

Unterräume von \mathbb{R}^n

Unser nächstes Ziel ist es, die Struktur der Lösungsmengen von homogenen linearen Gleichungssystemen zu studieren. Diese wird spezieller sein als für beliebige lineare Gleichungssysteme. Zum einen haben wir bereits festgestellt, dass sich die rechte Seite der erweiterten Koeffizientenmatrix nicht ändert, wenn wir auf diese Matrix die elementaren Zeilentransformationen anwenden: Dabei werden die Nullen in der rechten Spalte mit reellen Zahlen multipliziert und dann addiert, dabei bleiben auf der rechten Seite der erweiterten Koeffizientenmatrix also weiterhin nur Nullen. Ferner hat ein homogenes lineares Gleichungssystem stets mindestens eine Lösung, und zwar,

als Lösungsvektor aufgeschrieben, den Nullvektor $\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$: Setzt man alle Varia-

blen auf 0, so erhält man als Summe auf der linken Seite jeder Gleichung im linearen Gleichungssystem 0, was im homogenen linearen Gleichungssystem stets gleich der rechten Seite ist. Es gibt noch weitere Besonderheiten, die die Struktur der Lösungsmenge von einem homogenen linearen Gleichungssystem betreffen, insbesondere dann, wenn es unendlich ist.

Wir wollen zunächst den dazugehörigen Begriff kennenlernen und dann später einsehen, was dieser mit den Lösungsmengen homogener linearer Gleichungssysteme zu tun hat.

Definition 2.1. Eine Teilmenge $U \subseteq \mathbb{R}^n$ heißt **(linearer) Unterraum** von \mathbb{R}^n , falls folgende drei Bedingungen für U erfüllt sind:

(U0) Die Teilmenge U enthält $0 \in \mathbb{R}^n$ als Element, in Formeln: $0 \in U$.

(U1) Zu jedem Element $u \in U$ und jeder reellen Zahl $\lambda \in \mathbb{R}$ ist das Ergebnis der Skalarmultiplikation $\lambda \cdot u$ wieder ein Element von U , also

$$(u \in U \wedge \lambda \in \mathbb{R}) \Rightarrow \lambda u \in U.$$

(U2) Für je zwei Elemente $u, w \in U$ von U ist deren Summe ebenfalls Element von U , also

$$(u \in U \wedge w \in U) \Rightarrow u + w \in U.$$

Wir betrachten zunächst einige Beispiele.

Beispiel. • Die Teilmenge

$$V = \left\{ \left(\begin{array}{c} b \\ -b \\ 0 \end{array} \right) \in \mathbb{R}^3 \mid b \in \mathbb{R} \right\} \subseteq \mathbb{R}^3$$

ist ein Unterraum von \mathbb{R}^3 . Bevor wir das beweisen, wollen wir uns nochmal vergegenwärtigen, was diese Definition bedeutet, indem wir einige Elemente von V angeben. Um ein Element anzugeben, brauchen wir uns nur einen Wert für die reelle Zahl b auszusuchen. So liefert etwa $b = 2$ das Element $\begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$ der Menge V , und $b = -3,2$ liefert das

Element $\begin{pmatrix} -3,2 \\ 3,2 \\ 0 \end{pmatrix}$ von V . Der Vektor $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ ist hingegen kein Element von V , da es keine Wahl vom Parameter b gibt, sodass

$$\begin{pmatrix} b \\ -b \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

gilt.

Nun wollen wir beweisen, dass V ein Unterraum von \mathbb{R}^3 ist, indem wir nachprüfen, dass jede der drei obigen Eigenschaften für V erfüllt ist.

zu (U0): Wir müssen überprüfen, ob der Nullvektor 0 ein Element von V ist. Setzt man in der Definition von V für den Parameter b den Wert $b = 0$ ein, so erhält man genau den Nullvektor in \mathbb{R}^3 , also $0 \in V$ und (U0) ist also für V erfüllt.

zu (U1): Seien $\begin{pmatrix} c \\ -c \\ 0 \end{pmatrix} \in V$ und $\lambda \in \mathbb{R}$ vorgegeben. Wir müssen überprüfen, ob das λ -Vielfache von diesem Vektor wieder ein Element in V ist. Nach Definition der Skalarmultiplikation ist

$$\lambda \cdot \begin{pmatrix} c \\ -c \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda c \\ -\lambda c \\ 0 \end{pmatrix}.$$

Wir sehen, dass dieser Vektor in V ist, indem wir $b = \lambda c$ wählen.

zu (U2): Wir müssen zeigen, dass die Summe je zweier Elemente von V wiederum in V ist. Seien also

$$\begin{pmatrix} c \\ -c \\ 0 \end{pmatrix}, \begin{pmatrix} d \\ -d \\ 0 \end{pmatrix} \in V$$

vorgegeben. Deren Summe ist der Vektor

$$\begin{pmatrix} c \\ -c \\ 0 \end{pmatrix} + \begin{pmatrix} d \\ -d \\ 0 \end{pmatrix} = \begin{pmatrix} c+d \\ -c-d \\ 0 \end{pmatrix} = \begin{pmatrix} c+d \\ -(c+d) \\ 0 \end{pmatrix}.$$

Dass dieser Vektor in V liegt, sehen wir, indem wir den Parameter $b = c + d$ setzen.

Insgesamt haben wir also nachgewiesen, dass für die Menge V die drei Eigenschaften (U0), (U1), (U2) erfüllt sind, und V somit ein Unterraum von \mathbb{R}^3 ist.

- Sei $W \subseteq \mathbb{R}^3$ gegeben durch

$$W = \left\{ \begin{pmatrix} x - y \\ x + y \\ 2y \end{pmatrix} \in \mathbb{R}^3 \mid x, y \in \mathbb{R} \right\}.$$

Auch diese Teilmenge von \mathbb{R}^3 ist ein Unterraum von \mathbb{R}^3 , wie wir nun beweisen wollen. Dabei gehen wir wie eben vor.

- (U0)** Setzt man in der Definition von W die Parameter $x = y = 0$ ein, so erhält man

$$\begin{pmatrix} 0 - 0 \\ 0 + 0 \\ 2 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

also ist $0 \in W$ erfüllt. Dabei kann man in diesem Fall diese Werte „raten“. Würde man sich systematischer überlegen wollen, ob der Nullvektor in W enthalten ist, so setzt man

$$\begin{pmatrix} x - y \\ x + y \\ 2y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

an und sieht sofort, dass diese Gleichheit zum linearen Gleichungssystem

$$\begin{cases} x - y = 0, \\ x + y = 0, \\ 2y = 0 \end{cases}$$

äquivalent ist. Dieses kann man mit einer beliebigen Methode schnell lösen und man sieht, dass $x = y = 0$ die einzige Lösung liefert.

- (U1)** Sei hier $\begin{pmatrix} z - w \\ z + w \\ 2w \end{pmatrix}$ für gewisse reelle Zahlen $z, w \in \mathbb{R}$ ein Element von W und sei $\lambda \in \mathbb{R}$ vorgegeben. Wir schauen uns nun das λ -fache Vielfache von diesem Vektor an:

$$\lambda \cdot \begin{pmatrix} z - w \\ z + w \\ 2w \end{pmatrix} = \begin{pmatrix} \lambda \cdot (z - w) \\ \lambda \cdot (z + w) \\ \lambda \cdot 2w \end{pmatrix} = \begin{pmatrix} \lambda z - \lambda w \\ \lambda z + \lambda w \\ 2\lambda w \end{pmatrix}.$$

Setzt man nun in der Definition von W die Parameter $x = \lambda z$ und $y = \lambda w$ ein, so erhält man gerade das für uns relevante Ergebnis der Skalarmultiplikation. Es gilt also für jedes Element $\begin{pmatrix} z - w \\ z + w \\ 2w \end{pmatrix}$

von W und jedes $\lambda \in \mathbb{R}$, dass auch $\lambda \cdot \begin{pmatrix} z - w \\ z + w \\ 2w \end{pmatrix}$ in W liegt.

(U2) Nun seien zwei Vektoren in W vorgegeben, etwa

$$\begin{pmatrix} x_1 - y_1 \\ x_1 + y_1 \\ 2y_1 \end{pmatrix}, \begin{pmatrix} x_2 - y_2 \\ x_2 + y_2 \\ 2y_2 \end{pmatrix} \in W.$$

Wir müssen prüfen, ob die Summe dieser beiden Vektoren ebenfalls in W liegt. Diese Summe ist gegeben durch

$$\begin{pmatrix} x_1 - y_1 \\ x_1 + y_1 \\ 2y_1 \end{pmatrix} + \begin{pmatrix} x_2 - y_2 \\ x_2 + y_2 \\ 2y_2 \end{pmatrix} = \begin{pmatrix} (x_1 - y_1) + (x_2 - y_2) \\ (x_1 + y_1) + (x_2 + y_2) \\ 2y_1 + 2y_2 \end{pmatrix}.$$

Wir sortieren nun die Summanden in der jeweiligen Komponente um und klammern, wenn nötig, Faktoren aus, um zu sehen, dass diese Summe wieder in die Form gebracht werden kann, die die Elemente von W haben.

$$\begin{pmatrix} (x_1 - y_1) + (x_2 - y_2) \\ (x_1 + y_1) + (x_2 + y_2) \\ 2y_1 + 2y_2 \end{pmatrix} = \begin{pmatrix} (x_1 + x_2) - (y_1 + y_2) \\ (x_1 + x_2) + (y_1 + y_2) \\ 2(y_1 + y_2) \end{pmatrix}.$$

Nun sieht man, indem man $x = x_1 + x_2$ setzt und $y = y_1 + y_2$ setzt, dass auch diese Summe ein Element der Teilmenge W von \mathbb{R}^3 ist.

Also sehen wir auch hier, dass für W alle drei Bedingungen aus der Definition eines Unterraums erfüllt sind und W somit ein Unterraum von \mathbb{R}^3 ist.

- Als nächstes fragen wir uns, was der kleinste Unterraum von \mathbb{R}^n ist. Als erstes untersuchen wir die leere Teilmenge $\emptyset \subseteq \mathbb{R}^n$, die Teilmenge, die keine Elemente enthält. Es stellt sich heraus, dass dies kein Unterraum von \mathbb{R}^n ist. Um zu beweisen, dass dies wahr ist, reicht es zu zeigen, dass eine der Bedingungen aus der Definition für diese Teilmenge nicht erfüllt ist. In diesem Fall ist schon die Bedingung (U0) nicht erfüllt. Die leere Menge enthält keine Elemente; insbesondere ist der Nullvektor von \mathbb{R}^n kein Element der leeren Menge. Also ist die leere Menge kein Unterraum von \mathbb{R}^n .

- Nun haben wir uns also vor Augen geführt, dass 0 dank der Eigenschaft (U0) auf jeden Fall in jedem Unterraum sein muss. Nun kann man sich fragen, ob die Teilmenge $\{0\}$ von \mathbb{R}^n , die nur den Nullvektor als Element enthält, ein Unterraum von \mathbb{R}^n ist. Das ist nun tatsächlich der Fall. Die Bedingung (U0) ist für $\{0\}$ offensichtlich erfüllt.

Als nächstes prüfen wir, ob die Bedingung (U1) für die Teilmenge $\{0\} \subseteq \mathbb{R}^n$ erfüllt ist. Ist nun ein $x \in \{0\}$ und eine reelle Zahl $\lambda \in \mathbb{R}$ vorgegeben, so stellen wir als erstes fest, dass $x = 0$ ist, denn das ist das einzige Element der Menge $\{0\}$. Multipliziert man mit dem Nullvektor mit einem beliebigen Skalar, so bekommt man wiederum den Nullvektor, also gilt:

$$\lambda \cdot x = \lambda \cdot 0 = 0 \in \{0\},$$

und die Bedingung (U1) ist erfüllt.

Als letztes müssen wir untersuchen, ob die Bedingung (U2) für die Teilmenge $\{0\}$ von \mathbb{R}^n erfüllt ist. Seien also $u, w \in \{0\}$, wir müssen prüfen, ob die Summe $u + w$ auch in dieser Teilmenge liegt. Da die Menge $\{0\}$ aber nur ein einziges Element hat, muss $u = w = 0$ gelten. Wir bemerken bei dieser Gelegenheit, dass wir in der Bedingung (U2) sowohl den Fall betrachten müssen, dass u und w verschiedene Vektoren sind, als auch den Fall, dass $u = w$ ist. Letzterer ist allerdings schon durch die Bedingung (U1) abgedeckt, da $u + u = 2 \cdot u$ ein skalares Vielfaches von u ist. In diesem Fall ist die Lage noch einfacher: Wir haben bereits geschlossen, dass $u = w = 0$ sein müssen und somit $u + w = 0$, also ein Element von $\{0\}$. Insgesamt sehen wir also, dass die Teilmenge $\{0\}$ von \mathbb{R}^n ein Unterraum von \mathbb{R}^n ist, und das ist somit auch der kleinste Unterraum von \mathbb{R}^n .

- Nun kommen wir zu dem anderen Extrem: Was ist der größte Unterraum von \mathbb{R}^n ? Auch das ergibt einerseits ein „langweiliges“ Beispiel, das trotzdem auch lehrreich ist. Wir erinnern uns, dass jede Menge eine Teilmenge von sich selbst ist, und betrachten die Teilmenge aller Vektoren \mathbb{R}^n in \mathbb{R}^n . Diese enthält also insbesondere den Nullvektor, und somit ist die Bedingung (U0) für die Teilmenge \mathbb{R}^n von \mathbb{R}^n erfüllt ist. Die Bedingung (U1) liest sich in diesem Fall so: Nimmt man einen beliebigen Vektor in \mathbb{R}^n und multipliziert diesen mit einer beliebigen reellen Zahl, so erhält man wieder einen Vektor in \mathbb{R}^n . Das ist nach Definition der Skalarmultiplikation erfüllt. Ähnlich einfach gestaltet sich der Nachweis der Eigenschaft (U2): Diese besagt in dem Spezialfall von $\mathbb{R}^n \subseteq \mathbb{R}^n$, dass die Summe zweier Vektoren in \mathbb{R}^n wieder ein Vektor in \mathbb{R}^n ist, was wiederum nach Definition der Addition von Vektoren wahr ist. Die Teilmenge $\mathbb{R}^n \subseteq \mathbb{R}^n$ ist also ein Unterraum von \mathbb{R}^n .

- Nun haben wir einige Beispiele von Unterräumen gesehen; allerdings war bislang die einzige Teilmenge von \mathbb{R}^n , von der wir gesehen haben, dass sie kein Unterraum von \mathbb{R}^n ist, die leere Menge. Nun wollen wir zu einem weniger „exotischen“ Beispiel eines nicht-Unterraums kommen. Wir betrachten hierbei die Menge

$$X = \left\{ \begin{pmatrix} 3 \\ 0 \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid b \in \mathbb{R} \right\}.$$

Die Teilmenge X von \mathbb{R}^2 ist kein Unterraum. Um das einzusehen, reicht es, eine der drei Eigenschaften (U0), (U1) oder (U2) für X zu widerlegen (tatsächlich ist aber für diese spezielle Wahl von X keine der drei Bedingungen erfüllt, allerdings gibt es Teilmengen von \mathbb{R}^n , wo nur eine oder zwei der Bedingungen schiefehen). Wir zeigen, dass der Nullvektor $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ kein Element der Menge X ist. Dafür müssen wir zeigen, dass es keine Wahl des Parameters $b \in \mathbb{R}$ gibt, sodass

$$\begin{pmatrix} 3 \\ 0 \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

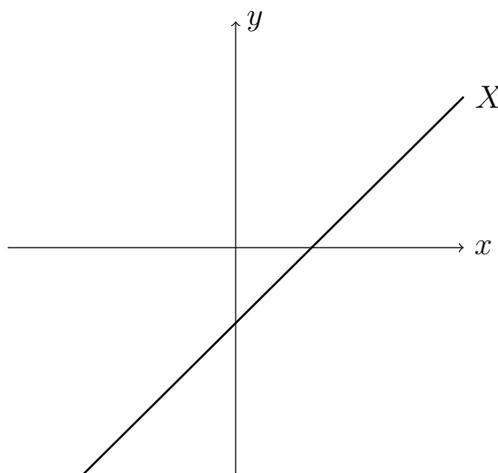
gilt. Wir wandeln diese Gleichung in ein lineares Gleichungssystem um:

$$\begin{pmatrix} 3 \\ 0 \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 3 + b = 0, \\ b = 0. \end{cases}$$

Hier sieht man unmittelbar, dass dieses lineare Gleichungssystem keine Lösungen hat, und somit für keine Wahl des Parameters $b \in \mathbb{R}$ der Nullvektor als Element von X vorkommt. Also ist (U0) für X nicht erfüllt, und X ist somit kein Unterraum.

Nachdem wir nun einige Beispiele von Unterräumen kennengelernt haben, wollen wir auch über eine anschauliche Interpretation von Unterräumen sprechen. Aus dem Schulunterricht kennt man meistens bereits, dass man die Vektorräume \mathbb{R}^2 bzw. \mathbb{R}^3 mit der Ebene bzw. dem Raum identifizieren kann. Das ist eine sehr gute und wichtige Intuition; allerdings muss man bei einigen Analogien vorsichtig sein. So ist es etwa nie festgelegt, wo der Ursprung sein sollte und wo die Achsen hinzeigen sollten. Häufig gibt es Wahlen von Koordinatensystemen, die bequemer sind als andere, aber diese Wahlen sind fast nie zwingend und häufig kann es auch hier hilfreich sein, die Perspektive zu wechseln. Ein weiterer Aspekt, der mit Vorsicht betrachtet werden sollte, ist die Tatsache, dass „Ebenen“, die wir betrachten - etwa das Blatt Papier, auf dem wir zeichnen, oder die Tafel im Vorlesungssaal - stets endliche Gebilde sind, die Grenzen haben. Das trifft auf die Vektorräume \mathbb{R}^2 bzw. \mathbb{R}^3 nicht zu, wie man bei den Analogien stets im Hinterkopf behalten sollte.

Nach dieser Vorwarnung wollen wir uns nun (ohne Beweis) veranschaulichen, was Unterräume von \mathbb{R}^2 bzw. \mathbb{R}^3 sind. Im Falle von \mathbb{R}^2 hat man zwei Unterräume, die eine Sonderrolle spielen, nämlich $\{0\}$ und \mathbb{R}^2 , über die wir bereits gesprochen haben. Alle andere Unterräume von \mathbb{R}^2 lassen sich sehr einfach beschreiben: Es sind die Ursprungsgeraden in der Ebene. Wir haben übrigens bereits ein Beispiel dafür gesehen, dass beliebige Geraden in der Ebene im Allgemeinen keine Unterräume sind, haben wir tatsächlich eben schon gesehen. Die Menge $X \subseteq \mathbb{R}^2$, die wir soeben untersucht haben, ist eine Gerade in der Ebene, die in der Parameterform angegeben ist, und sieht in etwa so im Bild aus:



Im \mathbb{R}^3 ist die Lage ähnlich, jedoch etwas komplizierter: Auch hier kommt den Unterräumen $\{0\}$ und \mathbb{R}^3 eine Sonderrolle zu. Auch hier sind alle Geraden, die durch den Ursprung gehen, Unterräume von \mathbb{R}^3 , allerdings hat man weitere Unterräume: Die Ebenen, die durch den Ursprung gehen, sind ebenfalls Unterräume von \mathbb{R}^3 . Den Unterschied, den wir intuitiv bereits als „Dimension“ kennen, werden wir später genauer untersuchen.

Es ist auch angemessen, eine weitere Heuristik für Unterräume zu bestimmen: Es ist im Allgemeinen hilfreich zu wissen, dass Unterräume „ebene“ und nicht „gekrümmte“ Objekte sein sollten, also etwa keine Parabeln, Kreise, Hyperbeln, sondern etwas, was wie Ebenen und Geraden aussieht, und zusätzlich noch durch den Ursprung geht. Die *Heuristik* ist auch, dass in Formeln, die einen Unterraum definieren, keine Quadrate, fünfte Potenzen, Logarithmen, Sinusfunktionen, ja nichtmal Konstanten vorkommen sollen, sondern nur Summen von Vielfachen von Variablen. Man sollte jedoch bedenken, dass dies nur eine *Faustregel* ist und im Einzelfall nicht stimmen muss; es ist allerdings trotzdem hilfreich, das im Kopf zu behalten.

Nun können wir zu dem Zusammenhang zwischen linearen Gleichungssystemen und Unterräumen präziser machen.

Satz 2.2. *Die Lösungsmenge L eines homogenen linearen Gleichungssystems in n Variablen ist ein Unterraum von \mathbb{R}^n .*

Beweis. Sei das homogene lineare Gleichungssystem

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = 0, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = 0, \\ \vdots \\ a_{k1}x_1 + a_{k2}x_2 + a_{k3}x_3 + \dots + a_{kn}x_n = 0, \end{cases}$$

vorgegeben. (Wir erinnern uns, dass „homogen“ gerade hieß, dass auf der rechten Seite nur Nullen und keine anderen reellen Zahlen stehen dürfen.) Wir prüfen nun, dass für die Lösungsmenge L dieses Gleichungssystems die Bedingungen (U0), (U1) und (U2) erfüllt sind.

zu (U0): Zunächst müssen wir zeigen, dass der Nullvektor in L liegt, also eine Lösung des vorgegebenen linearen Gleichungssystems ist. Das haben wir bereits zu Anfang dieses Abschnitts überlegt: Setzt man $x_1 = x_2 = \dots = x_n = 0$ in die linke Seite des linearen Gleichungssystems ein, so erhält man als Ergebnis Null und die Gleichung wird für diese Belegung der Variablen somit zu einer wahren Aussage. Also ist $0 \in L$, und die Bedingung (U0) ist erfüllt für L .

zu (U1): Im nächsten Schritt müssen wir zeigen, dass eine beliebige Lösung des homogenen linearen Gleichungssystems, wenn sie mit einem Skalar, also einer festen reellen Zahl, multipliziert wird, wieder eine Lösung

des Gleichungssystems liefert. Sei also $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ ein Element von L , also

eine Lösung des vorgegebenen Gleichungssystems, und $\lambda \in \mathbb{R}$ sei eine

beliebige reelle Zahl. Wir wollen also prüfen, ob der Vektor $\begin{pmatrix} \lambda y_1 \\ \lambda y_2 \\ \vdots \\ \lambda y_n \end{pmatrix}$

wieder eine Lösung des Gleichungssystems ist. Dafür setzen wir diesen in die linke Seite jeder der k Gleichungen ein und prüfen, dass als Ergebnis 0 rauskommt, also die Gleichung zu einer wahren Aussage wird. Wir setzen in die i -te Gleichung ein, $1 \leq i \leq k$, und klammern den gemeinsamen Faktor λ aus:

$$a_{i1} \cdot (\lambda y_1) + a_{i2} \cdot (\lambda y_2) + \dots + a_{in} \cdot (\lambda y_n) = \lambda \cdot (a_{i1}y_1 + a_{i2}y_2 + \dots + a_{in}y_n).$$

Nun ist die Zahl in Klammern genau die, die wir erhalten, wenn wir in

die i -te Gleichung den Vektor $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ einsetzen. Da dieser nach Voraus-

setzung ein Element der Lösungsmenge L ist, muss also als Ergebnis 0 herauskommen. Somit erhalten wir

$$a_{i1} \cdot (\lambda y_1) + a_{i2} \cdot (\lambda y_2) + \dots + a_{in} \cdot (\lambda y_n) = \lambda \cdot 0 = 0.$$

Also ist der Vektor $\begin{pmatrix} \lambda y_1 \\ \lambda y_2 \\ \vdots \\ \lambda y_n \end{pmatrix}$ tatsächlich eine Lösung der Gleichung und somit ein Element von L . Die Bedingung (U1) ist somit erfüllt.

zu (U2): Der letzte Schritt verläuft ganz ähnlich wie der vorherige. Wir müssen zeigen, dass die Summe zweier Lösungsvektoren eines homogenen linearen Gleichungssystems wiederum eine Lösung dieses Gleichungssystems ist. Dafür betrachten wir Vektoren $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ und $\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$ in L und

wollen zeigen, dass ihr Summenvektor $\begin{pmatrix} y_1 + z_1 \\ y_2 + z_2 \\ \vdots \\ y_n + z_n \end{pmatrix}$ wieder ein Element

von L ist. Dafür setzen wir diesen erneut in die i -te Gleichung des homogenen linearen Gleichungssystems ein, das wir betrachten, und lösen die Klammern auf:

$$\begin{aligned} & a_{i1} \cdot (y_1 + z_1) + a_{i2} \cdot (y_2 + z_2) + \dots + a_{in} \cdot (y_n + z_n) \\ &= a_{i1}y_1 + a_{i1}z_1 + a_{i2}y_2 + a_{i2}z_2 + \dots + a_{in}y_n + a_{in}z_n. \end{aligned}$$

Nun nutzen wir - nachdem wir soeben ja bereits das Distributivgesetz für die Multiplikation und Addition der reellen Zahlen genutzt haben - auch noch die Kommutativität und Assoziativität der Addition der reellen Zahlen aus, also die Tatsache, dass wir die Summen reeller Zahlen beliebig umordnen können, und sortieren alle Summanden mit y 's in eine und alle mit z 's in eine andere Klammer:

$$\begin{aligned} & a_{i1}y_1 + a_{i1}z_1 + a_{i2}y_2 + a_{i2}z_2 + \dots + a_{in}y_n + a_{in}z_n \\ &= (a_{i1}y_1 + a_{i2}y_2 + \dots + a_{in}y_n) + (a_{i1}z_1 + a_{i2}z_2 + \dots + a_{in}z_n). \end{aligned}$$

Nun steht in der ersten Klammer gerade die linke Seite der i -ten Gleichung des betrachteten Gleichungssystems, in die der Vektor $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ eingesetzt wurde; ähnlich sieht es mit der zweiten Klammer aus, nur dass

da der Vektor $\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$ eingesetzt wurde. Da diese beiden Vektoren jeweils Elemente der Lösungsmenge L sind, ergibt jede der Klammern 0 - die erste, weil der erste der beiden Vektoren in L ist, und die zweite Klammer, da der zweite Vektor in L ist. Insgesamt erhalten wir also:

$$\begin{aligned} & a_{i1} \cdot (y_1 + z_1) + a_{i2} \cdot (y_2 + z_2) + \dots + a_{in} \cdot (y_n + z_n) \\ &= (a_{i1}y_1 + a_{i2}y_2 + \dots + a_{in}y_n) + (a_{i1}z_1 + a_{i2}z_2 + \dots + a_{in}z_n) \\ &= 0 + 0 = 0. \end{aligned}$$

Also ist die Summe zweier Elemente von L wieder selbst ein Element von L , und somit ist die Bedingung (U2) für L erfüllt.

Insgesamt haben wir nun gezeigt, dass die Lösungsmenge L eines jeden homogenen linearen Gleichungssystems in n Variablen ein Unterraum von \mathbb{R}^n ist. \square

Insbesondere haben wir dabei zwei ganz besondere Eigenschaften von homogenen linearen Gleichungssystemen gezeigt, die uns etwas über die Struktur der Lösungsmenge verraten, und die homogene lineare Gleichungssystemen so speziell machen: Addiert man zwei Lösungen oder multipliziert man eine Lösung mit einem festen Faktor, so erhält man erneut eine Lösung des homogenen linearen Gleichungssystems.

Wir wollen nun weitere Eigenschaften von Unterräumen erforschen. Insbesondere wollen wir eine einfache Methode haben, um Unterräume mit vorgegebenen Eigenschaften zu bauen.

Dazu beschäftigen wir uns mit der Frage, wie der kleinste Unterraum aussieht, der vorgegebene Vektoren enthält. Um das konkreter zu gestalten,

betrachten wir die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ in \mathbb{R}^3 . Sicherlich ist \mathbb{R}^3 selbst ein Unterraum von \mathbb{R}^3 , der diese Vektoren enthält, allerdings wollen wir uns überlegen, ob es auch andere gibt.

Zunächst können wir uns fragen, ob die Menge aus diesen beiden Vektoren, $\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$, ein Unterraum von \mathbb{R}^3 ist. Diese Frage ist schnell verneint: Die Bedingung (U0) ist hier bereits nicht erfüllt, denn diese Menge enthält ja den Nullvektor nicht.

Also machen wir den nächsten Versuch, indem wir nun den Nullvektor hinzufügen und uns fragen, ob die Menge $\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$ ein Unterraum

von \mathbb{R}^3 ist. Nun ist die Bedingung (U0) für die Menge, die wir betrachten, erfüllt. Allerdings merkt man auch hier schnell, dass die Bedingung (U1) nicht erfüllt ist. Etwa das Zweifache des Vektors $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, also der Vektor $\begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$, ist nicht in der Menge enthalten. Und das Problem ist nicht nur für das Vielfache mit dem Faktor 2 vorhanden, es fehlt auch das Dreifache, das 3,2-fache, das -4 -fache dieses Vektors, und so weiter. Genauso fehlen die Vielfachen des Vektors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Wir machen also einen nächsten, bereits etwas komplizierteren Versuch, und betrachten die Menge

$$\left\{ \lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \cup \left\{ \mu \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mid \mu \in \mathbb{R} \right\}.$$

(Zur Erinnerung: Die Vereinigung $X \cup Y$ zweier Mengen ist als diejenige Menge definiert, die genau alle Elemente von X und alle Elemente von Y enthält.) Das ist also die Menge aller Vielfachen von dem Vektor $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und

von dem Vektor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Man kann sich also fragen, ob das ein Unterraum von \mathbb{R}^3 ist. Zunächst bemerken wir, dass der Nullvektor in dieser Menge enthalten ist, da zum Beispiel

$$0 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

gilt. Also ist (U0) für unsere neue Menge erfüllt. Außerdem erhalten wir, wenn wir von einem Vielfachen λv eines Vektors v , das μ -fache Vielfache bilden, wiederum ein Vielfaches von v , dank der Assoziativität der Skalarmultiplikation:

$$\mu \cdot (\lambda \cdot v) = (\mu \cdot \lambda) \cdot v,$$

also das $(\mu \cdot \lambda)$ -Vielfache von v . Somit ist die Bedingung (U1) für unsere neue Menge erfüllt. Es bleibt also, die Bedingung (U2) dafür zu untersuchen. Wieder kommt man schnell zu dem Schluss, dass diese nicht erfüllt ist: etwa ist der Summenvektor

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}$$

weder ein Vielfaches von $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ noch von $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Wir müssen also dafür sorgen, dass auch Summen von Vielfachen der vorgegebenen Vektoren in der Menge enthalten sind, damit sie eine Chance hat, ein Unterraum zu sein. Um das nochmal zusammenzufassen: Ist $U \subseteq \mathbb{R}^3$ ein Unterraum, der die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ enthält, so muss er auch jeden Vektor der Form

$$\lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \mu \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

für beliebige reelle Zahlen $\lambda, \mu \in \mathbb{R}$ enthalten.

Um diese Erkenntnisse nun präziser zu fassen, führen wir einige Begriffe ein.

Definition 2.3. Seien $v_1, \dots, v_k \in \mathbb{R}^n$ Vektoren. Eine **Linearkombination** der Vektoren v_1, v_2, \dots, v_k mit Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}$ ist die Summe

$$\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2 + \dots + \lambda_k \cdot v_k \in \mathbb{R}^n.$$

Die **lineare Hülle** (auch **Spann**, englisch „span“) der Vektoren v_1, v_2, \dots, v_k ist die Menge

$$\text{Span}(v_1, v_2, \dots, v_k) = \{ \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k \mid \lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R} \} \subseteq \mathbb{R}^n.$$

Die lineare Hülle der Vektoren v_1, v_2, \dots, v_k ist also die Menge aller Linearkombinationen der Vektoren v_1, v_2, \dots, v_k mit allen möglichen Koeffizienten.

Um diesen abstrakten Begriff etwas besser zu verstehen, wollen wir einige Beispiele betrachten. Danach werden wir zeigen, dass die lineare Hülle einer Ansammlung von Vektoren in \mathbb{R}^n stets ein Unterraum von \mathbb{R}^n ist.

Beispiel. • In manchen Fällen kann man nicht viel mehr zu der Beschreibung der linearen Hülle sagen. In dem Fall, den wir als Beispiel betrachtet haben, ist

$$\text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = \left\{ \lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \mu \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \right\} \subseteq \mathbb{R}^3$$

die lineare Hülle der Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Man kann zeigen, dass das der kleinste Unterraum von \mathbb{R}^3 ist, der diese zwei Vektoren enthält.

- Betrachtet man die lineare Hülle der Vektoren $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$, so sehen wir, dass wir diesen Unterraum von \mathbb{R}^3 bereits untersucht haben, indem wir die Definition etwas umschreiben:

$$\begin{aligned} \text{Span} \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \right) &= \left\{ x \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \mid x, y \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} x \\ x \\ 0 \end{pmatrix} + \begin{pmatrix} -y \\ y \\ 2y \end{pmatrix} \mid x, y \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} x - y \\ x + y \\ 2y \end{pmatrix} \mid x, y \in \mathbb{R} \right\}, \end{aligned}$$

und das war gerade ein Beispiel für Unterräume, die wir betrachtet haben.

Wir kommen nun wieder zu allgemeineren Aussagen und halten nochmal das Ergebnis unserer Vorüberlegungen mit Hilfe der neuen Begriffe fest.

Bemerkung. Sei $U \subseteq \mathbb{R}^n$ ein Unterraum von \mathbb{R}^n und w_1, \dots, w_k eine Ansammlung von Vektoren in U . Dann ist jede Linearkombination der Vektoren w_1, \dots, w_k , etwa $\mu_1 w_1 + \dots + \mu_k w_k$ mit $\mu_1, \dots, \mu_k \in \mathbb{R}$, ein Element von U , denn eine Linearkombination entsteht gerade als Summe von Vielfachen von diesen Vektoren, und sowohl alle Vielfache von diesen Vektoren sind nach Eigenschaft (U1) der Unterräume wieder in U enthalten, als auch ihre Summe, dank der Eigenschaft (U2).

Da jede Linearkombination der Vektoren w_1, \dots, w_k in U enthalten ist, ist auch die Ansammlung all solcher Linearkombinationen eine Teilmenge von U . Wir fassen zusammen: Ist $U \subseteq \mathbb{R}^n$ ein Unterraum von \mathbb{R}^n und w_1, \dots, w_k eine Ansammlung von Vektoren in U , so gilt: $\text{Span}(w_1, \dots, w_k) \subseteq U$.

Im Allgemeinen ist die lineare Hülle mehrerer Vektoren aus einem Unterraum zwar eine Teilmenge von diesem Unterraum, muss allerdings nicht unbedingt der ganze Unterraum sein. Ist es doch der Fall, so hat man einen speziellen Namen dafür.

Definition 2.4. Sei $U \subseteq \mathbb{R}^n$ ein Unterraum und $w_1, w_2, \dots, w_k \in U$ Vektoren in U . Gilt für diese Vektoren, dass

$$\text{Span}(w_1, \dots, w_k) = U$$

ist, so nennt man diese Ansammlung von Vektoren ein **Erzeugendensystem** von U .

Wir betrachten nun weitere Beispiele für lineare Hüllen.

Beispiel. • Bildet man die lineare Hülle von einem einzelnen Vektor, so ist diese lineare Hülle einfach die Menge der Vielfachen von diesem Vektor, etwa wie in diesem Beispiel:

$$\text{Span} \left(\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right) = \left\{ \lambda \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \subseteq \mathbb{R}^3.$$

• Wir betrachten ein weiteres Beispiel, nämlich

$$\text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} \right).$$

Das ist nach Definition die Menge

$$\left\{ \lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \mu \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + \nu \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} \mid \lambda, \mu, \nu \in \mathbb{R} \right\} \subseteq \mathbb{R}^4.$$

An diesem Beispiel wollen wir uns vergegenwärtigen, dass jeder der Vektoren, deren lineare Hülle wir bilden, in dieser linearen Hülle ent-

halten ist. So ist etwa $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$ ein Element von dieser linearen Hülle, wie

man sehen kann, indem man die Parameter $\lambda = 1, \mu = 0, \nu = 0$ wählt:

$$1 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}.$$

Ähnlich sieht man, dass $\begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$ ein Element der linearen Hülle ist, indem

man die Parameter $\lambda = 0, \mu = 1, \nu = 0$ wählt, und um $\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ zu erhalten,

setzt man $\lambda = 0, \mu = 0, \nu = 1$.

Nun wollen wir wie versprochen zeigen, dass die lineare Hülle einer Ansammlung von Vektoren tatsächlich ein Unterraum ist. Wir haben bereits in einem Beispiel uns klargemacht, dass der kleinste Unterraum, der eine Ansammlung von Vektoren enthält, alle Linearkombinationen dieser Vektoren enthalten muss; nun sehen wir, dass dies tatsächlich auch ausreicht.

Satz 2.5. *Seien $v_1, v_2, \dots, v_k \in \mathbb{R}^n$ Vektoren. Dann ist die lineare Hülle $\text{Span}(v_1, v_2, \dots, v_k)$ dieser Vektoren ein Unterraum von \mathbb{R}^n .*

Beweis. Wie bereits mehrfach gesehen, müssen wir zeigen, dass die drei Bedingungen (U0), (U1), (U2) für die Menge $\text{Span}(v_1, v_2, \dots, v_k)$ erfüllt sind.

(U0) Zunächst müssen wir zeigen, dass der Nullvektor in der linearen Hülle der Vektoren $v_1, v_2, \dots, v_k \in \mathbb{R}^n$ enthalten ist. Also müssen wir den Nullvektor als Linearkombination der Vektoren v_1, v_2, \dots, v_k darstellen. Das ist einfach zu bewerkstelligen, denn wir können die Linearkombination mit Koeffizienten $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$ betrachten und erhalten

$$0 \cdot v_1 + 0 \cdot v_2 + \dots + 0 \cdot v_k = 0.$$

Der Nullvektor ist also in linearer Hülle einer beliebigen Ansammlung von Vektoren enthalten, in Zeichen $0 \in \text{Span}(v_1, v_2, \dots, v_k)$. Somit ist die Bedingung (U0) für die Menge $\text{Span}(v_1, v_2, \dots, v_k)$ erfüllt.

(U1) Nun müssen wir beweisen, dass das Vielfache einer Linearkombination der Vektoren $v_1, v_2, \dots, v_k \in \mathbb{R}^n$ wieder eine Linearkombination dieser Vektoren ist. Sei also ein beliebiges Element der Menge $\text{Span}(v_1, v_2, \dots, v_k)$ vorgegeben, d.h. eine Linearkombination $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k$ der Vektoren $v_1, v_2, \dots, v_k \in \mathbb{R}^n$ mit Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_k$. Ferner sei ein Skalar $\lambda \in \mathbb{R}$ vorgegeben. Bilden wir das λ -fache des vorgegebenen Vektors, so erhalten wir (durch Ausmultiplizieren und Zusammenfassen der Vorfaktoren)

$$\begin{aligned} \lambda(\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k) &= \lambda \cdot \lambda_1 v_1 + \lambda \cdot \lambda_2 v_2 + \dots + \lambda \cdot \lambda_k v_k \\ &= (\lambda \lambda_1) \cdot v_1 + (\lambda \lambda_2) \cdot v_2 + \dots + (\lambda \lambda_k) \cdot v_k. \end{aligned}$$

Wir erkennen, dass dies wieder eine Linearkombination der Vektoren v_1, v_2, \dots, v_k ist, und zwar mit Koeffizienten $\lambda \lambda_1, \lambda \lambda_2, \dots, \lambda \lambda_k$. Also ist das Vielfache eines Elements von $\text{Span}(v_1, v_2, \dots, v_k)$ stets wieder ein Element dieser Menge, und die Bedingung (U1) ist für diese Menge erfüllt. Wir bemerken, dass wir dabei mehrere Rechenregeln für Vektoren benutzt haben, die wir früher aufgestellt hatten, insbesondere das Distributivgesetz und die Assoziativität der Skalarmultiplikation.

(U2) Der letzte Teil des Beweises ist wie der dem vorherigen recht ähnlich. Nun haben wir zwei (ob unterschiedliche oder gleiche, wissen wir nicht) Linearkombinationen der Vektoren v_1, v_2, \dots, v_k vorgegeben, mit Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_k$ bzw. $\mu_1, \mu_2, \dots, \mu_k$. Wir müssen zeigen, dass die Summe dieser Vektoren $\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k$ und $\mu_1 v_1 + \mu_2 v_2 + \dots + \mu_k v_k$ wieder ein Element von $\text{Span}(v_1, v_2, \dots, v_k)$ ist, also wieder eine Linearkombination der Vektoren v_1, v_2, \dots, v_k . Wir bilden wieder die Summe und benutzen die Rechenregeln für Vektoren, um die Summanden geeignet zu sortieren und zusammenzufassen:

$$\begin{aligned} & (\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k) + (\mu_1 v_1 + \mu_2 v_2 + \dots + \mu_k v_k) \\ &= (\lambda_1 v_1 + \mu_1 v_1) + (\lambda_2 v_2 + \mu_2 v_2) + \dots + (\lambda_k v_k + \mu_k v_k) \\ &= (\lambda_1 + \mu_1) v_1 + (\lambda_2 + \mu_2) v_2 + \dots + (\lambda_k + \mu_k) v_k. \end{aligned}$$

Das ist wieder eine Linearkombination der Vektoren v_1, v_2, \dots, v_k , und zwar mit Koeffizienten $\lambda_1 + \mu_1, \lambda_2 + \mu_2, \dots, \lambda_k + \mu_k$, also selbst ein Element der Menge $\text{Span}(v_1, v_2, \dots, v_k)$. Die Summe zweier Vektoren in dieser Menge ist also stets wieder in $\text{Span}(v_1, v_2, \dots, v_k)$ enthalten, und somit ist die Bedingung (U2) für die Menge $\text{Span}(v_1, v_2, \dots, v_k)$ erfüllt.

Insgesamt sehen wir also, dass die lineare Hülle $\text{Span}(v_1, v_2, \dots, v_k)$ der Vektoren $v_1, v_2, \dots, v_k \in \mathbb{R}^n$ stets ein Unterraum von \mathbb{R}^n ist. \square

3 Lineare Unabhängigkeit und Basen

Wir haben uns bereits mit den Begriffen der linearen Hülle und des Erzeugendensystems ein wenig vertraut gemacht. Wir wollen nochmal auf ein Spezialfall eingehen, in dem die lineare Hülle einiger Vektoren sich besonders einfach beschreiben lässt. Die lineare Hülle der Vektoren $v_1, v_2, \dots, v_k \in \mathbb{R}^n$ ist stets im \mathbb{R}^n enthalten; allerdings ist das in manchen Fällen keine echte Teilmenge, sondern sogar der ganze Vektorraum \mathbb{R}^n . In diesem Fall ist die Ansammlung von Vektoren v_1, v_2, \dots, v_k nach Definition ein Erzeugendensystem des \mathbb{R}^n . Mit anderen Worten: In diesem Fall lässt sich jeder Vektor in \mathbb{R}^n als eine Linearkombination der Vektoren v_1, v_2, \dots, v_k darstellen. Das kann man den Vektoren nicht immer gleich ansehen. Wir werden später Methoden kennenlernen, um das zu prüfen. Im Moment wollen wir uns zwei Beispiele nur exemplarisch anschauen.

Beispiel. Wir wollen als erstes zeigen, dass die Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$

ein Erzeugendensystem von \mathbb{R}^4 sind, also dass

$$\text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right) = \mathbb{R}^4$$

ist. Wieder anders ausgedrückt, müssen wir jeden Vektor in \mathbb{R}^4 als Linearkombination der vier obigen Vektoren schreiben. Im Allgemeinen würd man dafür ein mehr oder minder kompliziertes Verfahren verwenden, das auf Lösen gewisser linearer Gleichungssysteme hinauslaufen wird. In diesem Fall ist

es jedoch einfach, Koeffizienten zu finden, die jeden Vektor $\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$ zur Linear-

kombination der obigen Vektoren machen. Für die Vorüberlegung bezeichnen wir diese Koeffizienten mit ?; dabei steht ? möglicherweise für mehrere unterschiedliche Zahlen. Wir suchen reelle Zahlen, die wir für ? einsetzen können, sodass die Gleichheit

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = ? \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + ? \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + ? \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + ? \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

stimmt. Nun stellen wir fest, dass nur der erste Koeffizient zu der ersten Komponente der Summe auf der rechten Seite beitragen wird, da die restlichen Vektoren

$\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ in der ersten Komponente 0 haben und dies

durch Skalarmultiplikation nicht ändern wird. Also muss der erste Koeffizient notwendigerweise x sein, sonst kann die Gleichheit in der ersten Komponente nicht erfüllt sein. Ähnlich stellt man fest, dass nur der zweite Koeffizient zu der zweiten Komponente beiträgt und somit y sein muss. Die letzten beiden Koeffizienten müssen mit der analogen Überlegung z bzw. w sein, und insgesamt erhalten wir

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = x \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + w \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Dass die Koeffizienten in diesem Fall so einfach sind, ist eine Besonderheit von diesem Erzeugendensystem. Wir werden später nochmal über die Eigenschaften von diesem Erzeugendensystem, das manchmal auch *Standardbasis* von \mathbb{R}^4 genannt wird, sprechen.

An der Zahl der Komponenten 4 ist hingegen nichts besonders. Auch in Dimension 5 kann man die ähnlich gebildeten Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

betrachten, und diese werden wieder ein Erzeugendensystem des \mathbb{R}^5 bilden; allgemeiner kann man in jedem \mathbb{R}^n die Standardbasis betrachten, die aus n Vektoren besteht, jeweils mit genau einer 1 in einer Komponente und Nullen in allen anderen Komponenten. Diese Vektoren in \mathbb{R}^n bilden ein Erzeugendensystem des \mathbb{R}^n .

Wir betrachten ein zweites Beispiel von etwas anderer Natur.

Beispiel. In diesem Beispiel wollen wir zeigen, dass $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ ein Erzeugendensystem von \mathbb{R}^2 ist. Wir müssen also zu jedem Vektor $\begin{pmatrix} x \\ y \end{pmatrix}$ in \mathbb{R}^2 eine Linearkombination der Vektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, die genau den Vektor $\begin{pmatrix} x \\ y \end{pmatrix}$ ergibt. Dafür gibt es auch systematischere Verfahren; allerdings wäre der erste Ansatz, die Werte x und y als Parameter zu betrachten und nach den reellen Zahlen λ, μ zu suchen, die der Gleichung

$$\lambda \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

genügen. Das ist äquivalent zu

$$\begin{cases} \lambda + \mu = x, \\ \lambda + 2\mu = y, \end{cases}$$

also zu einem linearen Gleichungssystem in Variablen λ und μ mit Parametern x und y . Löst man dieses, so sieht man, dass die eindeutige Lösung dieses Gleichungssystems für beliebige Werte von $x, y \in \mathbb{R}$ gegeben ist durch $\lambda = 2x - y, \mu = y - x$. Tatsächlich prüfen wir nach, dass

$$(2x - y) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (y - x) \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

gilt. Also liefert die Linearkombination der Vektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ mit Koeffizienten $2x - y, y - x$ genau den Vektor $\begin{pmatrix} x \\ y \end{pmatrix}$. Da jeder Vektor von \mathbb{R}^2 sich als

Linearkombination von den Vektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ schreiben lässt, ist diese Ansammlung von Vektoren ein Erzeugendensystem von \mathbb{R}^2 , und umformuliert gilt:

$$\text{Span} \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right) = \mathbb{R}^2.$$

Nun kann man sich fragen, ob es Zufall ist, dass das Erzeugendensystem vom \mathbb{R}^4 vier Elemente hatte, während das Erzeugendensystem von \mathbb{R}^2 nur zwei Elemente hatte. Tatsächlich werden wir später sehen, dass wir im \mathbb{R}^n stets mindestens n Vektoren für jedes Erzeugendensystem brauchen, allerdings auch mehr als n Vektoren möglich sind.

Damit eng verknüpft ist die Frage der Redundanz bei einem Erzeugendensystem. Hat man ein Erzeugendensystem von einem Unterraum U vorgegeben, also eine Ansammlung von Vektoren in U , sodass sich jeder Vektor in U als Linearkombination dieser Ansammlung von Vektoren schreiben lässt, so kann man sich fragen, ob manche der Vektoren vielleicht „überflüssig“ sind und ohne U zu verändern gestrichen werden können. Das kann tatsächlich passieren, und ob dieser Fall eintritt, ist nicht immer ganz einfach zu sehen. Wir erläutern das an zwei Beispielen etwas genauer.

Beispiel. Wir betrachten den Unterraum $U = \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right)$ von \mathbb{R}^3 .

Nach Definition von Erzeugendensystemen ist $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ ein Erzeugendensystem von U . Wir wollen nun zeigen, dass der zweite Vektor hierbei redundant

ist, also dass sich jede Linearkombinationen von Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ auch

als Linearkombination, d.h. als Vielfaches, von dem Vektor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ schreiben

lässt, und umgekehrt jede Linearkombination vom Vektor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ eine Linear-

kombination der Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ ist.

Zunächst zu der zweiten Aussage. Wir haben bereits gezeigt, dass der Vektor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ in der linearen Hülle $U = \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right)$ enthalten ist. Ferner hatten wir uns bereits überlegt, dass lineare Hülle von einer Ansammlung

von Vektoren in einem Unterraum U eine Teilmenge von U sein wird, folglich also

$$\text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) \subseteq \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right).$$

Das können wir allerdings auch direkt sehen. Hat man nämlich eine Linearkombination des Vektors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ mit dem Koeffizienten $\lambda \in \mathbb{R}$, also ein λ -Vielfaches von diesem, so lässt sich dieses Vielfache wie folgt als Linearkombination der Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ schreiben:

$$\lambda \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \lambda \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix},$$

also mit Koeffizienten $\lambda, 0$. (Es gäbe auch andere Möglichkeiten, etwa auch, Koeffizienten $0, \frac{\lambda}{2}$ zu wählen; allerdings ist an dieser Stelle nur entscheidend, dass es überhaupt geht, ob es mehrere Möglichkeiten dafür gibt, ist für unsere Fragestellung irrelevant.)

Nun fragen wir uns also, ob umgekehrt auch jedes Element der Menge $\text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right)$ auch ein Element von $\text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right)$ ist, also ob man eine Inklusion

$$\text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) \supseteq \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right).$$

hat. Sei also eine Linearkombination der Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ mit Koeffizienten λ, μ vorgegeben. Wir müssen zeigen, dass dieser Vektor eine Linearkombination, also ein Vielfaches, des Vektors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ist. Das ist nicht weiter schwierig:

$$\lambda \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mu \cdot \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda + 2\mu \\ 0 \\ 0 \end{pmatrix} = (\lambda + 2\mu) \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

der vorgegebene Vektor ist also ein $(\lambda + 2\mu)$ -faches des Vektors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ und insbesondere ein Element von $\text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right)$.

Insgesamt haben wir also gezeigt, dass der Unterraum U von \mathbb{R}^3 auch die lineare Hülle des Vektors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ist. Also sind sowohl $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ als auch die Ansammlung von Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ jeweils ein Erzeugendensystem von U , und in dem zweiten Erzeugendensystem haben wir eine Redundanz: Wir können den Vektor $\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ entfernen und erhalten ein neues, kleineres Erzeugendensystem.

In dem vorherigen Beispiel war die Redundanz leicht zu erahnen, da die Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ Vielfache voneinander sind. Das ist jedoch nicht der einzige Fall, in dem eine Redundanz vorliegt. Im nächsten Beispiel beschäftigen wir uns mit einem viel weniger offensichtlichen Beispiel für Redundanz in einem Erzeugendensystem.

Beispiel. Wir betrachten wieder den Unterraum

$$W = \text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} \right)$$

von \mathbb{R}^4 . Das ist nach Definition die Menge

$$\left\{ \lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \mu \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + \nu \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} \mid \lambda, \mu, \nu \in \mathbb{R} \right\} \subseteq \mathbb{R}^4.$$

Wir wollen nun zeigen, dass man den Vektor $\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ weglassen kann und genau

dieselbe lineare Hülle erhält, also dass

$$\text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} \right) = \text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} \right)$$

gilt.

Wie wir uns bereits überlegt haben, sind die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$

Elemente von W , also muss deren lineare Hülle eine Teilmenge von W sein, in Zeichen

$$\text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} \right) \subseteq W.$$

Wir können diese Inklusion auch alternativ explizit sehen: Hat man eine

Linearkombination der Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$ mit Koeffizienten λ, μ , so wird

diese mit $\nu = 0$ zu einer Linearkombination der Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$,

also

$$\lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \mu \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} = \lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \mu \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$$

Hierbei haben wir übrigens keine Eigenschaften dieser drei speziellen Vektoren genutzt, die analoge Inklusion erhält man für drei beliebige Vektoren.

Nun müssen wir umgekehrt einsehen, dass jede Linearkombination der

Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ sich bereits als Linearkombination der ersten bei-

den Vektoren darstellen lässt. Das ist eine Besonderheit dieser konkreten Vektoren, und sie geht insbesondere auf die folgende Beobachtung zurück:

Der dritte Vektor $\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ lässt sich als Linearkombination der beiden anderen Vektoren darstellen. Tatsächlich gilt:

$$\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} = (-1) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}.$$

Nun kann man diese Tatsache für ein recht abstraktes Argument ausnutzen.

Wir haben gesehen, dass alle drei Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ sich als Linearkombinationen der Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$ darstellen lassen, also Elemente

des Unterraums $\text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} \right)$ von \mathbb{R}^4 sind. Wir haben gesehen, dass

dann auch ihre lineare Hülle W in diesem Unterraum $\text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} \right)$ enthalten sein muss, also

$$W \subseteq \text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} \right)$$

gelten muss.

Wir wollen allerdings auch ein alternatives, direktes Argument anführen.

Wir wollen zeigen, dass jede Linearkombination der Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$

sich als Linearkombinationen der Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$ schreiben lässt. Seien

also Koeffizienten $\lambda, \mu, \nu \in \mathbb{R}$ vorgegeben. Wir betrachten die Linearkombination

$$\lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \mu \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + \nu \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}.$$

In diese setzen wir den Ausdruck für den Vektor $\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ in Termen der beiden

anderen Vektoren ein:

$$\lambda \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + \nu \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + \nu \cdot \left((-1) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} \right)$$

Wir nutzen nun die Rechenregeln für Vektoren, insbesondere das Distributivgesetz für Skalarmultiplikation und Addition und Kommutativität und Assoziativgesetz der Addition, um diesen Ausdruck umzuformen. Einfacher gesagt, lösen wir die Klammern auf, stellen die Terme um und Klammern die

Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$ wieder aus. Dabei erhalten wir:

$$\begin{aligned} & \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + \nu \cdot \left((-1) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} \right) \\ &= \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} - \nu \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + 2\nu \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} \\ &= \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} - \nu \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + 2\nu \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} \\ &= (\lambda - \nu) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + (\mu + 2\nu) \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}. \end{aligned}$$

Also sehen wir auch auf diesem Wege, dass

$$W \subseteq \text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} \right)$$

gilt.

Insgesamt sehen wir also, dass sowohl die Ansammlung von Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$ als auch $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ jeweils ein Erzeugendensystem von W ist. Bei dem zweiten Erzeugendensystem haben wir also wieder eine Redundanz, die allerdings nicht daher kommt, dass dieser Vektor ein Vielfaches eines der anderen Vektoren ist.

Im letzten Beispiel beruhte der explizite Beweis der „Redundanz“ darauf, dass wir einen der drei Vektoren als Linearkombination der beiden anderen darstellen konnten. Es wird auch allgemein ein recht ähnliches Kriterium für Redundanz geben. Da es etwas einfacher ist, das Kriterium dann auf konkrete Vektoren anzuwenden, formulieren wir das Kriterium leicht anders. Es wird in Kürze, wenn auch nicht sofort, klar werden, wie es mit der Vorüberlegung und mit der Redundanz bei Erzeugendensystemen zusammenhängt.

Wir fangen gleich mit einer Definition an. Der Hintergrund dieser Definition ist der folgende: Wir können aus jeder Ansammlung von Vektoren durch eine Linearkombination 0 erzeugen, wie wir bereits gesehen haben: Wir brauchen nur für alle Koeffizienten 0 einsetzen. Allerdings kann man bei manchen Ansammlungen von Vektoren 0 auch durch andere Koeffizienten erreichen, es liegt also eine Redundanz vor. Es stellt sich heraus, dass diese Redundanz schon alle Redundanzen erfasst.

Definition 3.1. Seien $v_1, \dots, v_k \in \mathbb{R}^n$ Vektoren. Die Vektoren v_1, \dots, v_k heißen **linear unabhängig**, falls die einzige Linearkombination dieser Vektoren, die Null ergibt, nur Nullen als Koeffizienten hat. In Formeln: Die Vektoren v_1, \dots, v_k heißen **linear unabhängig**, falls für alle $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ aus

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = 0$$

bereits $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$ folgt.

Sind Vektoren v_1, \dots, v_k nicht linear unabhängig, so nennt man sie **linear abhängig**.

Dies ist ein recht abstrakter Begriff, der zunächst einiger Gewöhnung bedarf. Wir fangen mit einigen konkreten Beispielen an.

Beispiel. • Es ist etwas einfacher zu zeigen, dass gewisse Vektoren linear abhängig sind, denn da reicht es, eine Auswahl von Koeffizienten zu finden, die nicht alle 0 sind und derart, dass die Linearkombination mit diesen Koeffizienten 0 ergibt. Das kann unter Umständen einfacher sein, als zu beweisen, dass es solche Koeffizienten nicht gibt. Wir wollen nun

zeigen, dass die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ linear abhängig sind. Dafür

geben wir ein Beispiel von einer Linearkombination mit Koeffizienten an, die nicht sämtlich 0 sind, und sodass diese Linearkombination 0 ergibt. Tatsächlich gilt

$$(-1) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wir bemerken, dass diese Linearkombination aus der Gleichung

$$\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} = (-1) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}.$$

entsteht, indem wir den Vektor $\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ von beiden Seiten abziehen.

Das ist nicht die einzige Linearkombination der Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$,

die Null ergibt, aber nicht nur Nullen als Koeffizienten hat. Eine weitere ist etwa gegeben durch

$$1 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + (-2) \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Es gibt sogar unendlich viele solche Linearkombinationen; für die Definition der linearen Abhängigkeit ist jedoch entscheidend, dass es mindestens eine solche Linearkombination gibt (deren Koeffizienten nicht nur Nullen sind).

Die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ sind also linear abhängig. Diese Erkenntnis geht damit einher, dass wir bereits eine Redundanz bei diesen Vektoren als Erzeugendensystem von $\text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} \right)$ festgestellt haben.

- Die Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ sind ebenfalls linear abhängig. Eine Ansammlung von Vektoren, unter denen Vielfache von anderen Vektoren dieser Ansammlung vorkommen, ist stets linear abhängig. Das ist gewissermaßen der einfachste Fall von linearer Abhängigkeit; es gibt, wie wir ja auch gesehen haben, auch andere Fälle.

Hier wollen wir aber die Behauptung, die Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ seien linear abhängig, wieder explizit beweisen, indem wir wieder eine Linearkombination dieser Vektoren angeben, die Null ergibt, allerdings wieder nicht nur Nullen als Koeffizienten hat.

Wieder wird es unendlich viele solche Linearkombinationen geben; es ist am einfachsten, wenn wir die Linearkombination

$$2 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

wählen. Für den Nachweis der linearen Abhängigkeit hätten wir allerdings auch genauso gut die Linearkombination

$$16 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + (-8) \cdot \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

wählen können.

Wieder geht hier die lineare Abhängigkeit mit der Redundanz bei den Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$ als Erzeugendensystem des Unterraums $\text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right)$ von \mathbb{R}^3 einher.

- In diesem Beispiel wollen wir nun endlich ein explizites Beispiel für lineare unabhängige Vektoren sehen. Wir behaupten, dass die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ linear unabhängig sind. Dafür müssen wir zeigen, dass jede

Linearkombination dieser Vektoren, die Null ergibt, bereits nur Nullen als Koeffizienten hat. Seien also $\lambda, \mu \in \mathbb{R}$ beliebige reelle Zahlen. Wir wollen zeigen, dass aus der Gleichheit

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

bereits folgt, dass $\lambda = 0$ und $\mu = 0$ folgt. Wir sehen, dass die obige Gleichung zu dem System aus den drei Gleichungen äquivalent ist, die sich in den Zeilen ergeben:

$$\begin{aligned} \lambda \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \mu \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ \Leftrightarrow \begin{cases} \lambda + \mu = 0, \\ 2\lambda = 0, \\ 3\lambda = 0. \end{cases} \end{aligned}$$

Das ist nun ein lineares Gleichungssystem, die wir ja schon zu lösen gelernt haben. Dieses lineare Gleichungssystem ist besonders einfach, und man sieht sofort, dass $\lambda = 0, \mu = 0$ die einzige Lösung dieses Gleichungssystems liefert. Aus der Annahme, dass

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

gilt, haben wir also $\lambda = 0, \mu = 0$ gefolgert, und somit sind die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ linear unabhängig.

Nun wollen wir zu dem allgemeinen Verfahren kommen, wie man lineare Unabhängigkeit von vorgegebenen Vektoren testet. Im Grunde ergibt sich dieses Verfahren aus der Überlegung, die wir im letzten Beispiel geführt haben. Wir erhalten aus der hypothetischen Gleichheit

$$\lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_k v_k = 0$$

für vorgegebene Vektoren $v_1, \dots, v_k \in \mathbb{R}^n$ ein lineares Gleichungssystem für die Koeffizienten $\lambda_1, \lambda_2, \dots, \lambda_k$ und können dieses lösen, um zu entscheiden, ob die vorgegebenen Vektoren linear unabhängig sind. Das führen wir zunächst in einem Beispiel aus, und kommen dann zum allgemeinen Verfahren.

Beispiel. Wir wollen wissen, ob die Vektoren $\begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \\ 7 \\ 0 \end{pmatrix}$ in

\mathbb{R}^5 linear unabhängig sind. Wir müssen also entscheiden, ob es reelle Zahlen $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in \mathbb{R}$ gibt, die nicht alle gleichzeitig 0 sind und sodass die Gleichung

$$\lambda_1 \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 3 \\ 2 \\ 0 \\ 7 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

erfüllt ist. Die beiden Seiten sind genau dann gleich, wenn sie komponentenweise gleich sind. Wir erhalten also aus jeder Zeile eine Gleichung, und diese zusammen ergeben ein homogenes lineares Gleichungssystem in den Variablen $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ mit 5 Gleichungen:

$$\begin{aligned} & \lambda_1 \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_4 \begin{pmatrix} 3 \\ 2 \\ 0 \\ 7 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ \Leftrightarrow & \begin{cases} 1 \cdot \lambda_1 + 1 \cdot \lambda_2 + 0 \cdot \lambda_3 + 3 \cdot \lambda_4 = 0 \\ 2 \cdot \lambda_1 + 0 \cdot \lambda_2 + 0 \cdot \lambda_3 + 2 \cdot \lambda_4 = 0 \\ (-1) \cdot \lambda_1 + 1 \cdot \lambda_2 + 0 \cdot \lambda_3 + 0 \cdot \lambda_4 = 0 \\ 1 \cdot \lambda_1 + 0 \cdot \lambda_2 + 1 \cdot \lambda_3 + 7 \cdot \lambda_4 = 0 \\ 0 \cdot \lambda_1 + 0 \cdot \lambda_2 + 0 \cdot \lambda_3 + 0 \cdot \lambda_4 = 0 \end{cases} \end{aligned}$$

Wir wissen bereits, wie wir ein lineares Gleichungssystem mehr oder minder systematisch lösen können. Wir stellen zunächst die erweiterte Koeffizientenmatrix dieses homogenen linearen Gleichungssystems auf, um die Notation zu vereinfachen. Diese sieht wie folgt aus:

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 3 & 0 \\ 2 & 0 & 0 & 2 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Wir bemerken gleich, dass für die späteren Rechnungen die Aufstellung des expliziten Gleichungssystems überflüssig sein wird, denn wir können gleich zu der erweiterten Koeffizientenmatrix davon übergehen, die als Spalten auf

der linken Seite genau die Einträge der Vektoren haben wird, die auf lineare Unabhängigkeit untersucht werden.

Nun bringen wir die linke Seite der erweiterten Koeffizientenmatrix in Zeilenstufenform durch elementare Zeilentransformationen. Wir benutzen dabei nicht das systematische Gauß-Verfahren, sondern vertauschen zunächst Bequemlichkeit halber die Zeilen.

$$\begin{array}{ccc}
 \left(\begin{array}{cccc|c} 1 & 1 & 0 & 3 & 0 \\ 2 & 0 & 0 & 2 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 7 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) & \xrightarrow{\text{I} \leftrightarrow \text{II}, \text{II}_{\text{neu}} \leftrightarrow \text{III}_{\text{alt}}, \text{III}_{\text{neu}} \leftrightarrow \text{IV}_{\text{alt}}} & \left(\begin{array}{cccc|c} 2 & 0 & 0 & 2 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 7 & 0 \\ 1 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \\
 & & \xrightarrow{\text{I}/2, \text{II} + \text{I}_{\text{neu}}, \text{III} - \text{I}_{\text{neu}}, \text{IV} - \text{I}_{\text{neu}}} & \left(\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 6 & 0 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \\
 & & & \xrightarrow{\text{IV} - \text{II}} & \left(\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)
 \end{array}$$

Nun ist die Matrix in Zeilenstufenform, und wir bemerken, dass wir genauso viele Stufen wie Variablen haben, nämlich jeweils 4. Für diesen Fall haben wir uns bereits überlegt, dass das zugehörige lineare Gleichungssystem eindeutig lösbar ist. Wir kennen auch schon die eindeutige Lösung dieses Gleichungssystems, da es homogen ist: Es ist die Lösung $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. Also

haben wir für die Vektoren $\begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \\ 7 \\ 0 \end{pmatrix}$ gezeigt, dass die einzige

Linearkombination dieser Vektoren, die 0 ergibt, nur Nullen als Koeffizienten hat. Folglich sind diese Vektoren nach Definition linear unabhängig.

Es ist leicht einzusehen, dass dieses Verfahren allgemein funktioniert. Wir halten dieses Verfahren nochmal genauer fest.

Allgemeines Verfahren. Wir geben hier ein allgemeines Verfahren an, um zu entscheiden, ob Vektoren v_1, v_2, \dots, v_k in \mathbb{R}^n linear unabhängig sind.

1. Man bilde die Matrix mit den (Einträgen der) Vektoren v_1, v_2, \dots, v_k als Spalten.

2. Man bringe diese Matrix durch elementare Zeilentransformationen in Zeilenstufenform.
3. Hat die Matrix in Zeilenstufenform genauso viele Stufen wie die Anzahl der vorgegebenen Vektoren (also k Stufen), so sind die Vektoren v_1, v_2, \dots, v_k linear unabhängig. Hat die Matrix in Zeilenstufenform weniger Stufen als die Anzahl der vorgegebenen Vektoren, so sind die Vektoren v_1, v_2, \dots, v_k linear abhängig.

Wir bemerken nochmal, dass es pro Spalte höchstens eine Stufe geben kann, und eine beliebige Matrix in Zeilenstufenform höchstens so viele Stufen haben kann wie Spalten. Deswegen war es in der obigen Fallunterscheidung nicht nötig, den Fall zu betrachten, dass die Matrix mehr als k Stufen hat.

Nun ist es ganz ähnlich so, dass man auch höchstens eine Stufe pro Zeile haben kann. Hat man also mehr als n Vektoren in \mathbb{R}^n vorgegeben, so wird man im obigen Verfahren stets höchstens n Stufen erhalten und somit weniger Stufen als Vektoren. Das ist die wesentliche Idee im Beweis der folgenden Proposition.

Proposition 3.2. *Seien Vektoren v_1, v_2, \dots, v_k in \mathbb{R}^n vorgegeben. Ist $k > n$, so sind v_1, v_2, \dots, v_k linear abhängig.*

Hat man also eine weniger als n oder genau n Vektoren in \mathbb{R}^n vorgegeben, so kann man nicht ohne weiteres sagen, ob diese linear unabhängig sind und muss zum Beispiel das obige Verfahren anwenden. Hat man allerdings mehr als n Vektoren in \mathbb{R}^n vorgegeben, so weiß man mit Sicherheit, dass diese Vektoren linear abhängig sind.

Hat man festgestellt, dass gewisse Vektoren v_1, v_2, \dots, v_k in \mathbb{R}^n linear abhängig sind, so weiß man, dass eine Redundanz bei diesen Vektoren vorliegt. Wir haben die Form der Redundanz noch nicht genauer diskutiert. Man kann sich allerdings fragen, ob man beliebige Vektoren weglassen kann, um ein linear unabhängiges System zu erreichen, oder etwa, um eine Ansammlung von Vektoren zu erreichen, die denselben Unterraum aufspannen. Es stellt sich heraus, dass man stets ein linear unabhängiges System von Vektoren aus v_1, v_2, \dots, v_k auswählen kann, das immer noch ein Erzeugendensystem von $\text{Span}(v_1, v_2, \dots, v_k)$ ist, allerdings geht das nicht mit einer beliebigen Auswahl von Vektoren (selbst wenn man eine Anzahl von weggelassenen Elementen festlegt). Das sehen wir im folgenden Beispiel, in dem wir allerdings die Beweise teilweise weglassen.

Beispiel. Wir betrachten die Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ in \mathbb{R}^3 . Es ist

nicht schwer, nachzuprüfen, dass diese ein Erzeugendensystem von \mathbb{R}^3 bilden, und nach der obigen Proposition wissen wir, dass sie linear abhängig sind. Tatsächlich kann man leicht einsehen, dass man den ersten oder den zweiten

Vektor weglassen kann und nach wie vor ein Erzeugendensystem von \mathbb{R}^3 hat (für die Ansammlung von Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ haben wir das bereits nachgeprüft), und dass die restlichen drei Vektoren linear unabhängig sind. Lässt man hingegen den Vektor $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ oder den Vektor $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ weg, oder sogar beide, so sind die verbleibenden Vektoren nach wie vor linear abhängig, denn in der Linearkombination

$$2 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - 1 \cdot \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

haben die letzten beiden Vektoren Nullen als Koeffizienten, es macht also keinen Unterschied, ob man sie dabei hat oder nicht. (Hier sehen wir übrigens ein Beispiel von einer Linearkombination, die Null ergibt und in der einige, aber nicht alle Koeffizienten Null sind. Das reicht jedoch nach Definition der linearen Unabhängigkeit, um zu sehen, dass die Vektoren linear abhängig sind.)

Man stellt auch fest, dass weder der Unterraum $\text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right)$ von \mathbb{R}^3 noch $\text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$ der ganze Vektorraum \mathbb{R}^3 ist.

Wir werden später nochmal genauer auf die Frage, wie man die Redundanz in einem linear abhängigen System von Vektoren entfernt, eingehen. Im Moment wollen wir aber nochmal unterstreichen, dass lineare Unabhängigkeit eines Erzeugendensystems eine hilfreiche Eigenschaft ist. Für solche Erzeugendensysteme verwenden wir den Begriff *Basis*, wie in der folgenden Definition erläutert. Dabei bildet der Unterraum $\{0\}$ von \mathbb{R}^n im Zusammenhang mit Basen und Erzeugendensystem stets eine Ausnahme, die wir nicht weiter berücksichtigen wollen. Deswegen werden wir meist diesen Unterraum ausschließen.

Definition 3.3. Sei $U \neq \{0\}$ ein Unterraum von \mathbb{R}^n und seien w_1, w_2, \dots, w_k Vektoren in U . Man nennt die Vektoren w_1, w_2, \dots, w_k eine **Basis** von U , falls diese Ansammlung von Vektoren ein Erzeugendensystem von U ist und gleichzeitig die Vektoren w_1, w_2, \dots, w_k linear unabhängig sind.

Wir betrachten zunächst ein Beispiel.

Beispiel. Sei U der Unterraum $\text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} \right)$ von \mathbb{R}^4 , den wir

bereits mehrfach betrachtet haben. Nach Definition ist $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$ ein

Erzeugendensystem von U , allerdings haben wir bereits gesehen, dass diese Vektoren nicht linear unabhängig sind, und somit keine Basis von U bilden.

Wir haben außerdem gesehen, dass die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$ ebenfalls ein

Erzeugendensystem von U bilden. Man kann zum Beispiel mit unserem allgemeinen Verfahren leicht nachrechnen, dass diese Vektoren linear unabhängig sind und somit eine Basis von U bilden.

Wir kommen nun zu einem der zentralen Sätze der linearen Algebra. Wir werden diesen leider nicht beweisen, da der Beweis einiges an Zeit in Anspruch nehmen würde. Dieser Satz macht die Verbindung zwischen Redundanz und linearer Abhängigkeit etwas präziser, und gibt zwei Charakterisierungen von Basen.

Satz 3.4. *Sei $U \neq \{0\}$ ein Unterraum von \mathbb{R}^n . Dann gilt:*

- 1) *Jede Basis ist ein Erzeugendensystem mit der minimalen Anzahl der Elemente unter allen Erzeugendensystemen von U . Umgekehrt ist jedes Erzeugendensystem von U , das diese minimale Anzahl von Elementen (unter Erzeugendensystemen von U) besitzt, automatisch eine Basis von U .*
- 2) *Jede Basis ist ein linear unabhängiges System von Vektoren in U , das die maximale Anzahl der Elemente unter allen linear unabhängigen Systemen von Vektoren in U besitzt. Umgekehrt ist jedes linear unabhängige System von Vektoren in U , das die maximal mögliche Anzahl von Elementen unter linear unabhängigen Systemen in U besitzt, automatisch eine Basis von U .*

Wir bemerken, dass es in dem Unterraum $U \neq \{0\}$ von \mathbb{R}^n jede Ansammlung von mehr als n Vektoren automatisch linear abhängig ist. Daher ist es im zweiten Teil des obigen Satzes überhaupt sinnvoll, von der maximalen möglichen Anzahl der Elemente in einem linear unabhängigen System von Vektoren in U zu sprechen. Häufig wird diese maximal mögliche Anzahl kleiner sein als n .

Wir bemerken außerdem, dass der erste Teil des obigen Satzes in einer gewissen Weise sagt, dass eine Basis ein Erzeugendensystem ohne Redundanz ist, denn wir können nach dem ersten Teil des obigen Satzes keinen Vektor weglassen und immer noch ein Erzeugendensystem von U erhalten. Mehr als das - wir können auch nicht unabhängig von der vorgegebenen Basis ein anderes Erzeugendensystem finden, das weniger Vektoren hat (dann wäre die Basis nämlich nicht minimal).

Ferner stellen wir fest, dass diese maximale Anzahl von Elementen für einen festen Unterraum U stets dieselbe ist. Gäbe es nämlich zwei Basen von dem Unterraum U mit unterschiedlicher Anzahl von Elementen, so hätte ja eine dieser Basen mehr Elemente als die andere Basis, und somit wäre die Basis mit der geringeren Anzahl von Elementen nicht maximal unter linear unabhängigen Systemen in U . Daraus ergibt sich der erste Teil des folgenden Korollars.

Korollar 3.5. *Sei $U \neq \{0\}$ ein Unterraum von \mathbb{R}^n .*

- 1) *Alle Basen von U haben dieselbe Anzahl von Elementen.*
- 2) *Jede Basis von \mathbb{R}^n hat genau n Elemente.*
- 3) *Der Unterraum U besitzt unendlich viele Basen.*

Wir werden keinen detaillierten Beweis des Korollars liefern. Die Beweis-idee für den ersten Teil haben wir eben bereits erläutert. Zu dem zweiten Teil bemerken wir, dass wir bereits eine n -elementige Basis des \mathbb{R}^n kennengelernt haben. Tatsächlich haben wir bereits gesehen, dass die n Vektoren in \mathbb{R}^n , die jeweils alle bis auf einen die Einträge 0 haben und an dieser ausgezeichneten Stelle den Eintrag 1 haben, also Vektoren von der Form

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Wir hatten bereits darüber gesprochen (und im Spezialfall $n = 5$ gezeigt), dass diese Vektoren ein Erzeugendensystem des \mathbb{R}^n bilden. Will man einen

beliebigen Vektor $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ als Linearkombination der Standardbasis darstellen,

so lässt sich besonders einfach sagen, wie die Koeffizienten in der Linearkombination aussehen müssen: Es sind genau die entsprechenden Komponenten

des Vektors; etwas genauer:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Um zu zeigen, dass dies tatsächlich eine Basis ist, müssen wir also noch prüfen, ob die Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

linear unabhängig sind.

Dafür kann man unterschiedliche Argumentationen wählen. Eine Begründung erhält man dadurch, dass man das allgemeine Verfahren zum Prüfen auf lineare Unabhängigkeit auf diese Vektoren anwendet. Schreibt man diese Vektoren in eine Matrix, so ergibt sich eine Matrix mit n Zeilen und n Spalten mit nur Einsen auf der Diagonale und Nullen überall sonst. Diese Matrix ist insbesondere schon in Zeilenstufenform und hat genau n Stufen, also genauso viele Stufen wie Spalten. Somit sind die Vektoren linear unabhängig.

Eine andere Möglichkeit zu argumentieren ist etwas direkter: Ergibt eine Linearkombination von den obigen Vektoren mit den Koeffizienten x_1, x_2, \dots, x_n den Nullvektor, so gilt

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

also sind alle Koeffizienten x_1, x_2, \dots, x_n bereits Null, wenn die Linearkombination Null ergibt.

In jedem Fall sehen wir dann, dass \mathbb{R}^n eine Basis mit genau n Elementen hat, nämlich die Standardbasis von \mathbb{R}^n . Nach dem ersten Teil des Korollars hat also jede Basis vom \mathbb{R}^n genau n Elemente.

Zu dem letzten Punkt wollen wir noch eine (nicht praktikable, aber theoretisch gültige) Methode erläutern, eine (und dann unendlich viele) Basen von einem Unterraum $U \neq \{0\}$ in \mathbb{R}^n zu finden. Man nimmt zunächst einen beliebigen Vektor $u_1 \in U$ mit der Eigenschaft $u_1 \neq 0$. Damit ist entweder

die Maximalanzahl 1 des linear unabhängigen Systems von Vektoren in U erreicht, oder man kann einen Vektor $u_2 \neq 0$ in U finden, sodass u_1, u_2 linear unabhängig sind. (Das ist nicht gänzlich offensichtlich, ist allerdings nicht schwer zu beweisen.) Im ersten Fall ist u_1 nach Satz 3.4 eine Basis von U , und wir wären fertig. Im zweiten Fall haben wir nun eine ähnliche Fallunterscheidung: Entweder ist u_1, u_2 bereits von der maximalen Größe und insbesondere eine Basis von U , oder wir finden noch ein $u_3 \in U$, sodass u_1, u_2, u_3 linear unabhängig sind. So können wir weitermachen; da wir höchstens n Vektoren in einem linear unabhängigen System von Vektoren in $U \subseteq \mathbb{R}^n$ haben können, wird das Verfahren irgendwann zuende sein, und dabei werden wir eine Basis von U ermittelt haben.

Es ist etwas einfacher zu sagen, wie man aus einer Basis vom Unterraum $U \neq \{0\}$ von \mathbb{R}^n unendlich viele Basen von U bekommt. Ist nämlich u_1, u_2, \dots, u_k eine Basis von U , so ist auch für jede reelle Zahl $\lambda \neq 0$ auch $\lambda \cdot u_1, u_2, \dots, u_k$ eine Basis von U , wie man unschwer nachprüfen kann. Tatsächlich gibt es im allgemeinen noch viele weitere Basen von U ; allerdings reicht unsere Konstruktion aus, um unendlich viele Basen von U zu bekommen.

Das Korollar führt uns zu der folgenden Definition, die der wichtigen Kenngröße eines Unterraums, nämlich der Anzahl der Elemente einer beliebigen Basis von diesem Unterraum, nun einen Namen gibt.

Definition 3.6. Sei $U \neq \{0\}$ ein Unterraum von \mathbb{R}^n . Die Anzahl der Elemente einer beliebigen Basis von U nennt man die **Dimension** von U . Wir schreiben $\dim U$ für Dimension von U .

Ferner wollen wir die Konvention $\dim\{0\} = 0$ verwenden.

Beispiel. • Zunächst sehen wir, dass die zweite Aussage des Korollars sich nun zu $\dim \mathbb{R}^n = n$ umformulieren lässt. Das entspricht auch unserer Intuition von Dimension, was darauf hinweist, dass die obige Definition der Dimension zumindest zum Teil unserer Vorstellung entspricht.

• Wir betrachten wieder den Unterraum $U = \text{Span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} \right)$

von \mathbb{R}^4 . Wir haben bereits gesehen, dass $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}$ eine Basis von U

ist. Damit folgt aus dem obigen Satz, dass jedes Erzeugendensystem vom Unterraum U mindestens zwei Elemente besitzen muss.

Ferner folgt aus dem Satz 3.4, dass je drei oder mehr Vektoren in U automatisch linear abhängig sind (a priori wussten wir, dass je fünf

oder mehr Vektoren in U linear unabhängig sind, da U eine Teilmenge von \mathbb{R}^4 ist).

- Man kann zeigen, dass auch allgemeiner der obige Dimensionsbegriff mit dem intuitiven übereinstimmt. Etwa für Ursprungsgeraden im \mathbb{R}^2 und \mathbb{R}^3 kann man zeigen, dass diese alle die Dimension 1 haben, während die Ursprungsebenen in \mathbb{R}^3 die Dimension 2 haben.

Allgemeiner kann man sagen, dass Dimension anschaulich der Anzahl der „gänzlich unabhängigen“ Richtungen in dem untersuchten Unterraum entspricht. Befindet man sich etwa in einer Ebene, so reicht ja die Angabe, wie weit man von einem festen Punkt, den wir uns als Koordinatenursprung vorstellen können, man nach rechts bzw. links zu gehen hat, und die Angabe, wie weit geradeaus bzw. in die exakt umgekehrte Richtung man gehen muss, aus, um jeden Punkt der Ebene zu beschreiben. Wir wissen bereits aus der Schule, dass man, nachdem man in der Ebene ein Koordinatensystem festgelegt hat, jeden Punkt der Ebene durch zwei Koordinaten beschreiben kann, während man im \mathbb{R}^3 drei Koordinaten braucht. Die Anzahl der Koordinaten, die man braucht, um einen Punkt des Unterraums festzulegen, ist eine anschauliche Interpretation der Dimension.

Als nächstes wollen wir weiter die Eigenschaften des Dimensionsbegriffs untersuchen. Die nächste Eigenschaft ist recht anschaulich, bedarf aber nicht-destotrotz eines Beweises.

Proposition 3.7. *Seien $U, V \subseteq \mathbb{R}^n$ Unterräume von \mathbb{R}^n mit der Eigenschaft, dass $U \subseteq V$ ist. Dann gilt: $\dim U \leq \dim V$. Ist ferner U eine echte Teilmenge von V , also $U \neq V$, so gilt $\dim U < \dim V$.*

Beweis der ersten Aussage. Wir beschränken uns auf den Beweis der ersten Aussage und werden die strikte Ungleichung nicht beweisen.

Zunächst wollen wir uns um die Sonderfälle kümmern. Ist $U = \{0\}$ der Null-Unterraum, so ist $\dim U = 0$. Diese ist kleiner oder gleich der Dimension des Unterraums V , da die letztere entweder selbst Null ist oder die Anzahl von Vektoren in einer Basis von V ist, also eine positive ganze Zahl. Insgesamt gilt in diesem Fall $0 = \dim U \leq \dim V$.

Von nun an nehmen wir an, dass $U \neq \{0\}$ nicht der Null-Unterraum von \mathbb{R}^n ist. Nach Korollar 3.5 wissen wir, dass U eine Basis hat. Wir betrachten eine Basis u_1, \dots, u_k von U . Nach Definition der Dimension ist $k = \dim U$, mit anderen Worten: Die Anzahl der Vektoren in jeder Basis von U ist genau die Dimension von U . Nun sind die Vektoren u_1, \dots, u_k ja insbesondere linear unabhängig, da sie eine Basis von U bilden. Außerdem impliziert $U \subseteq V$, dass jedes Element von U insbesondere auch ein Element von V ist. Also sind die Vektoren u_1, \dots, u_k linear unabhängig und liegen in V . Das heißt aber, dass ein maximales linear unabhängiges System von Vektoren in V mindestens k

Vektoren haben muss, denn wir haben mit u_1, \dots, u_k ein Beispiel von einem linear unabhängigen System in V gefunden, das k Elemente hat. Da jede Basis von V nach Satz 3.4 ein linear unabhängiges System von Vektoren in V mit maximal möglicher Anzahl von Elementen ist, und die Anzahl der Elemente in einer Basis von V gerade $\dim V$ ist, folgt: $\dim V \geq k = \dim U$. Das ist aber gerade die Behauptung. \square

Als nächstes wollen wir uns noch vergewissern, dass eine Basis eines Unterraums $U \neq \{0\}$ von \mathbb{R}^n in einer weiteren Hinsicht keine Redundanz enthält. Da eine Basis eines Unterraums $U \neq \{0\}$ von \mathbb{R}^n insbesondere ein Erzeugendensystem von U ist, lässt sich jedes Element von U als Linearkombination der Vektoren der Basis schreiben. A priori ist es jedoch denkbar, dass es unterschiedliche Darstellungen von ein und demselben Vektor als eine solche Linearkombination geben kann. Hat man ein Erzeugendensystem vorgegeben, das keine Basis ist, so werden solche Darstellungen als Linearkombinationen nicht eindeutig sein. Bei einer Basis gibt es eine solche Redundanz hingegen nicht, was der Aussage des folgenden Satzes entspricht.

Satz 3.8. *Seien w_1, \dots, w_l linear unabhängige Vektoren in \mathbb{R}^n und sei v ein Vektor in der linearen Hülle $\text{Span}(w_1, \dots, w_l)$ der Vektoren w_1, \dots, w_l . Sind $\lambda_1, \dots, \lambda_l \in \mathbb{R}$ und $\mu_1, \dots, \mu_l \in \mathbb{R}$ reelle Zahlen, so dass die Linearkombinationen $\lambda_1 w_1 + \dots + \lambda_l w_l$ und $\mu_1 w_1 + \dots + \mu_l w_l$ beide v ergeben, so müssen die Koeffizienten der Vektoren jeweils übereinstimmen, d.h. es muss gelten: $\lambda_1 = \mu_1, \lambda_2 = \mu_2, \dots, \lambda_l = \mu_l$.*

Um die Aussage nochmal umzuformulieren: Sind w_1, \dots, w_l linear unabhängige Vektoren in \mathbb{R}^n , so ist die Darstellung jedes Vektors der linearen Hülle $\text{Span}(w_1, \dots, w_l)$ der Vektoren w_1, \dots, w_l als Linearkombination eben dieser Vektoren eindeutig. erinnert man sich an die Definition der linearen Unabhängigkeit, so stellt man fest, dass diese genau diese Form von Eindeutigkeit für einen einzigen Vektor in $\text{Span}(w_1, \dots, w_l)$, nämlich für den Nullvektor fordert. Aus dieser Nicht-Redundanz der Darstellungen vom Nullvektor kann man also, so die Aussage des Satz, auf die Nicht-Redundanz der Darstellungen beliebiger Vektoren der linearen Hülle schließen.

Wir wollen diesen Satz nun beweisen.

Beweis. Seien w_1, \dots, w_l linear unabhängige Vektoren in \mathbb{R}^n und sei v ein Vektor in der linearen Hülle $\text{Span}(w_1, \dots, w_l)$ der Vektoren w_1, \dots, w_l , für den es die Koeffizienten $\lambda_1, \dots, \lambda_l \in \mathbb{R}$ und $\mu_1, \dots, \mu_l \in \mathbb{R}$ gibt, sodass

$$v = \lambda_1 w_1 + \dots + \lambda_l w_l = \mu_1 w_1 + \dots + \mu_l w_l$$

gilt. Unser Ziel ist es, zu beweisen, dass die jeweiligen Koeffizienten gleich sind. Dabei müssen wir ausnutzen, dass wir diese Art Eindeutigkeit der Darstellung für den Nullvektor bereits haben, da w_1, \dots, w_l linear unabhängig sind. Wir starten also mit der einfachen Identität $0 = v - v$ und setzen

für v an den beiden Stellen die beiden Linearkombinationen der Vektoren w_1, \dots, w_l ein. Danach sortieren wir die Summanden um, wie wir bereits mehrfach gemacht haben:

$$\begin{aligned} 0 &= v - v \\ &= (\lambda_1 w_1 + \dots + \lambda_l w_l) - (\mu_1 w_1 + \dots + \mu_l w_l) \\ &= \lambda_1 w_1 + \dots + \lambda_l w_l - \mu_1 w_1 - \dots - \mu_l w_l \\ &= (\lambda_1 - \mu_1) w_1 + \dots + (\lambda_l - \mu_l) w_l. \end{aligned}$$

Dies ist nun eine Linearkombination der Vektoren w_1, \dots, w_l , die den Nullvektor liefert, also müssen alle Koeffizienten Null sein, da die Vektoren w_1, \dots, w_l linear unabhängig sind. Folglich gilt $\lambda_1 - \mu_1 = 0$, $\lambda_2 - \mu_2 = 0$ und so weiter, schließlich auch $\lambda_l - \mu_l = 0$. Bringt man nun in jeder Gleichung das entsprechende μ_i auf die andere Seite, so erhalten wir genau die Gleichheiten, die wir zeigen wollten, also $\lambda_i = \mu_i$ für alle $1 \leq i \leq l$. \square

Daraus folgt unmittelbar, wie wir im Wesentlichen bereits angemerkt haben, die folgende Beobachtung.

Bemerkung. Sei $U \neq \{0\}$ ein Unterraum von \mathbb{R}^n und sei w_1, \dots, w_l eine Basis von U . Dann lässt sich jeder Vektor $u \in U$ eindeutig als Linearkombination der Vektoren w_1, \dots, w_l schreiben. Zu jedem $u \in U$ gibt es also eindeutige reelle Zahlen $\lambda_1, \dots, \lambda_l$, sodass

$$u = \lambda_1 w_1 + \dots + \lambda_l w_l$$

gilt. Man nennt diese reelle Zahlen $\lambda_1, \dots, \lambda_l$ die **Koordinaten von u bezüglich der Basis w_1, \dots, w_l** .

Diese Bezeichnung ist mit unserem „üblichen“ Begriff von Koordinaten konsistent, denn die Koordinaten eines Vektors $v \in \mathbb{R}^n$ bezüglich der Standardbasis sind gerade die Komponenten von v , wie wir uns bereits überlegt haben.

Wie diese Bezeichnung schon andeutet, bekommen wir bei einer anderen Wahl der Basis andere Koordinaten von demselben Vektor. Ein einfaches Beispiel davon erhält man, wenn man etwa bei den üblichen Koordinaten im Raum zunächst die y - und dann die x -Koordinate angeben würde. Dann beschreibt dasselbe Zahlentripel auf einmal einen anderen Punkt im Raum.

Die Koordinaten bezüglich einer Basis werden vor allem dann relevant, wenn man „den Blickwinkel wechseln“ möchte, etwa einen Kameraschwenk macht, und dadurch etwa „rechts vor mir“ oder „geradeaus vor mir“ eine andere Bedeutung haben als vorher. Später beschäftigen wir uns ausführlicher damit, wie man die Koordinaten von einem Vektor bezüglich einer Basis bestimmt, und wie sich diese Koordinaten für verschiedene Basen zueinander verhalten.

Jetzt wollen wir uns zunächst der Frage widmen, wie man eine Basis von einem Unterraum $U \neq \{0\}$ von \mathbb{R}^n finden kann. Ein ganz allgemeines Verfahren lässt sich dafür nicht so recht angeben, da die Methode davon abhängig sein dürfte, in welcher Form U vorgegeben ist. Wir wollen uns mit dem Spezialfall beschäftigen, dass U bereits als lineare Hülle einer Anzahl von Vektoren angegeben ist, d.h. mit dem Fall, dass wir bereits ein Erzeugendensystem von U kennen. Ist dieses Erzeugendensystem bereits linear unabhängig, so ist nicht viel zu tun, da dieses Erzeugendensystem bereits eine Basis von U ist. Wir wollen nun dem Fall widmen, dass das vorgegebene Erzeugendensystem linear abhängig ist. In diesem Fall stellt es sich heraus - auch wenn wir das nicht beweisen wollen - dass man stets eine Basis von U aus diesem Erzeugendensystem von U auswählen kann, indem man einen Teil der Vektoren weglässt.

Unser Verfahren wird allerdings nur sagen, welche Vektoren auf jeden Fall weggelassen werden dürfen. Es wird nicht präzise aussagen, ob auch andere Vektoren hätten weggelassen werden können, um eine Basis zu erhalten. Das Verfahren gibt uns also *eine* Möglichkeit (oder unter Umständen mehrere Möglichkeiten), um eine Basis von U aus einem Erzeugendensystem von U zu erhalten; es sagt *nicht* aus, wie *alle* solche mögliche Basen aussehen.

Bevor wir nun das Verfahren allgemeiner behandeln, betrachten wir ein Beispiel.

Beispiel. Wir betrachten den Unterraum

$$U = \text{Span} \left(\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 10 \\ 6 \end{pmatrix}, \begin{pmatrix} 4 \\ 13 \\ 8 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 5 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$$

von \mathbb{R}^3 . Wir wissen bereits, dass das Erzeugendensystem

$$\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 10 \\ 6 \end{pmatrix}, \begin{pmatrix} 4 \\ 13 \\ 8 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 5 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

von U keine Basis von U ist, da sechs Vektoren in \mathbb{R}^3 automatisch linear abhängig sind.

Wir haben soeben behauptet, dass man aus diesen sechs allerdings einige Vektoren streichen kann, sodass die restlichen eine Basis von U bilden. Dazu führen wir zunächst die ersten Schritte des Verfahrens aus, um zu untersuchen, ob all diese Vektoren linear unabhängig sind - auch wenn wir die Antwort auf diese Frage schon wissen.

Dabei schreiben wir die Einträge der Vektoren in eine Matrix und erhalten so die Matrix

$$\begin{pmatrix} 1 & 3 & 4 & 2 & -1 & 0 \\ 3 & 10 & 13 & 6 & -3 & 1 \\ 2 & 6 & 8 & 5 & -1 & 1 \end{pmatrix}.$$

Nun bringen wir diese Matrix durch elementare Zeilentransformationen in Zeilenstufenform:

$$\begin{pmatrix} 1 & 3 & 4 & 2 & -1 & 0 \\ 3 & 10 & 13 & 6 & -3 & 1 \\ 2 & 6 & 8 & 5 & -1 & 1 \end{pmatrix} \xrightarrow{II-3I, III-2I} \begin{pmatrix} 1 & 3 & 4 & 2 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Nun bemerken wir, dass wenn wir bestimmte Vektoren aus den ursprünglichen auswählen und mit diesen genau dieselben elementaren Zeilentransformationen durchführen, die zugehörige Matrix automatisch in Zeilenstufenform sein wird, und wir können hier sogar 3 Vektoren auswählen, sodass die zugehörige Matrix dann genau 3 Stufen hat. Am einfachsten ist es, wenn man jeweils den ersten Vektor jeder Stufe nimmt. Allgemeiner kann man nach dem Motto „ein Vektor pro Stufe“ vorgehen, wobei beachtet werden muss, dass der gewählte Vektor einer Stufe in der entsprechenden Zeile keine Null haben darf. In diesem Fall können wir die rot markierten Vektoren wählen und deren Veränderung durch die elementaren Zeilentransformationen beobachten:

$$\begin{pmatrix} 1 & 3 & 4 & 2 & -1 & 0 \\ 3 & 10 & 13 & 6 & -3 & 1 \\ 2 & 6 & 8 & 5 & -1 & 1 \end{pmatrix} \xrightarrow{II-3I, III-2I} \begin{pmatrix} 1 & 3 & 4 & 2 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

Um das nochmal zu verdeutlichen, schreiben wir nun nur diese Vektoren hin:

$$\begin{pmatrix} 1 & 4 & -1 \\ 3 & 13 & -3 \\ 2 & 8 & -1 \end{pmatrix} \xrightarrow{II-3I, III-2I} \begin{pmatrix} 1 & 4 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Also sehen wir so bereits, dass die gewählten Vektoren linear unabhängig sind. Tatsächlich erhält man auf diese Weise stets auch eine Basis von U , wenn man zu jeder Stufe einen Vektor (mit der erwähnten Einschränkung) gewählt hatte. Insbesondere ist der Unterraum U von \mathbb{R}^3 der ganze \mathbb{R}^3 , da drei linear unabhängige Vektoren in \mathbb{R}^3 stets schon eine Basis von \mathbb{R}^3 und somit insbesondere ein Erzeugendensystem von \mathbb{R}^3 bilden, wie wir aus dem Satz 3.4 und dem Korollar 3.5 schließen können.

So erhält man in jedem Fall eine Basis von U als Auswahl aus dem ursprünglichen Erzeugendensystem. Das Verfahren kann allerdings nicht beantworten, ob Vektoren, die zur selben Stufe gehören, linear unabhängig sind oder nicht, oder noch allgemeiner - ob man, wenn man nicht je einen Vektor pro Stufe wählt und die Matrix, die nur aus diesen Vektoren besteht, nicht durch die bereits durchgeführten Transformationen in Zeilenstufenform gebracht, ob man dann eine Basis von U erhält oder nicht.

Das wollen wir in diesem Beispiel nochmal illustrieren. Wählt man die ersten drei Vektoren, also

$$\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 10 \\ 6 \end{pmatrix}, \begin{pmatrix} 4 \\ 13 \\ 8 \end{pmatrix},$$

so gehören diese nicht zu drei unterschiedlichen Stufen in der obigen Matrix, und tatsächlich kann man in diesem konkreten Fall sogar mit den obigen Zeilentransformationen sehen, dass die Vektoren linear abhängig sind, denn die Zeilenstufenform der zugehörigen Matrix hat nur 2 Stufen:

$$\begin{pmatrix} 1 & 3 & 4 \\ 3 & 10 & 13 \\ 2 & 6 & 8 \end{pmatrix} \xrightarrow{\text{II}-3\text{I, III}-2\text{I}} \begin{pmatrix} 1 & 3 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

und die Vektoren $\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 10 \\ 6 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 13 \\ 8 \end{pmatrix}$ sind linear abhängig, also insbesondere keine Basis von U .

Wählt man hingegen die letzten drei Vektoren, die alle zu einer Stufe gehören,

$$\begin{pmatrix} 2 \\ 6 \\ 5 \end{pmatrix}, \begin{pmatrix} -1 \\ -3 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

so sieht man, dass die zugehörige Matrix durch obige Zeilentransformationen nicht in Zeilenstufenform gebracht wird, also

$$\begin{pmatrix} 2 & -1 & 0 \\ 6 & -3 & 1 \\ 5 & -1 & 1 \end{pmatrix} \xrightarrow{\text{II}-3\text{I, III}-2\text{I}} \begin{pmatrix} 2 & -1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Diese Matrix ist nicht in Zeilenstufenform. Wir führen weitere elementare Zeilentransformationen durch, um zu prüfen, ob die Vektoren $\begin{pmatrix} 2 \\ 6 \\ 5 \end{pmatrix}$, $\begin{pmatrix} -1 \\ -3 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ linear unabhängig sind.

$$\begin{pmatrix} 2 & -1 & 0 \\ 6 & -3 & 1 \\ 5 & -1 & 1 \end{pmatrix} \xrightarrow{\text{II}-3\text{I, III}-2\text{I}} \begin{pmatrix} 2 & -1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \xrightarrow{\text{III} \leftrightarrow \text{II, III}_{\text{neu}} \leftrightarrow \text{I}} \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{\text{II}-2\text{I, III}_{\text{neu}}/(-3)} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Die Vektoren $\begin{pmatrix} 2 \\ 6 \\ 5 \end{pmatrix}$, $\begin{pmatrix} -1 \\ -3 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, die in der alten Matrix zu einer Stufe gehört haben, stellen sich als linear unabhängig und somit als Basis von U heraus.

Jedoch ist das ein „Zufall“, und im allgemeinen können wir bei den Vektoren, die zur selben Stufe in der ursprünglichen Matrix gehören, keine Aussage über lineare Unabhängigkeit treffen. Will man aber nur irgendeine Basis finden und ist nicht an allen möglichen Auswahlen von Basen aus dem gegebenen Erzeugendensystem interessiert, so ist diese Frage irrelevant.

Wir wollen das Verfahren aus diesem Beispiel nun etwas systematischer beschreiben.

Allgemeines Verfahren. Wir wollen ein allgemeines Verfahren angeben, wie man eine Basis von dem Unterraum $\text{Span}(w_1, \dots, w_l)$ von \mathbb{R}^n bestimmt, die durch Weglassen von einigen Vektoren aus w_1, \dots, w_l entsteht.

1. Man schreibe die Einträge der Vektoren w_1, \dots, w_l als Spalten in eine Matrix und bringe diese Matrix durch elementare Zeilentransformationen in Zeilenstufenform.
2. Fängt in der k -ten Spalte eine Stufe in der obigen Zeilenstufenform an, so nehme man den Vektor w_k zur zukünftigen Basis hinzu. Das muss für alle Stufen durchgeführt werden.

Wir bemerken, dass die Dimension des Unterraums $\text{Span}(w_1, \dots, w_l)$ von \mathbb{R}^n der Anzahl der Stufen in der Matrix ist, die man aus der Matrix mit Spalten w_1, \dots, w_l erhält und die in Zeilenstufenform ist. Es ist eine nicht gänzlich offensichtliche Aussage, dass die Auswahl von Vektoren, die wir in dem allgemeinen Verfahren immer noch ein Erzeugendensystem von $\text{Span}(w_1, \dots, w_l)$, die wir jedoch nicht beweisen wollen. Hingegen ist es relativ klar, dass die Auswahl von Vektoren, die wir getroffen haben, linear unabhängig ist: Führt man mit der Matrix aus diesen Vektoren genau dieselben Zeilentransformationen durch, die wir mit der Matrix aus den Vektoren w_1, \dots, w_l durchgeführt haben, so erhalten wir ein Teil der zu w_1, \dots, w_l gehörenden Matrix in Zeilenstufenform, und zwar die erste Spalte jeder Stufe. Diese Matrix ist also auch in Zeilenstufenform und hat genauso viele Stufen wie Spalten (und genauso viele Stufen wie die Matrix, die wir erhalten haben, als wir mit allen Vektoren w_1, \dots, w_l gestartet sind).

Die Auswahl von Vektoren, die wir dabei erhalten, ist im Allgemeinen keineswegs die einzig mögliche; dieses Verfahren gibt keinerlei Auskunft darüber, welche Auswahlen sonst möglicherweise Basen liefern würden. Eine Teilaussage dazu kann man ohne weiteren Aufwand allerdings tätigen. Ersetzt man den Vektor in einer Stufe durch einen anderen in derselben Stufe, der in der relevanten Zeile (also in der Zeile, in der die besagte Stufe anfängt) keine 0 hat, so funktioniert das Verfahren ganz genauso. Wir sehen sofort, dass die Matrix, die zu ausgewählten Vektoren gehört, nach einer eventuellen Multiplikation einer Zeile mit einer reellen Zahl in Zeilenstufenform ist und genauso viele Stufen wie Spalten hat - und genauso viele Stufen wie die Matrix, die wir erhalten haben, als wir mit allen Vektoren w_1, \dots, w_l

gestartet sind. Zusammen mit der vorherigen Feststellung, dass diese letzte Stufenanzahl die Dimension von $\text{Span}(w_1, \dots, w_l)$ ist, ist es unmittelbar, dass diese weiteren Auswahlen ebenfalls Basen liefern, wenn wir die folgende Bemerkung heranziehen:

Bemerkung. Sei $U \subseteq \mathbb{R}^n$ ein Unterraum von \mathbb{R}^n mit $U \neq \{0\}$ und sei die Dimension von U sei k . Dann sind beliebige k linear unabhängige Vektoren in U bereits eine Basis von U . Genauso ist jedes Erzeugendensystem von U , das genau k Vektoren hat, bereits eine Basis von U .

Kennt man also die Dimension von einem Unterraum von \mathbb{R}^n , so braucht man von einer Ansammlung der „richtigen“ Anzahl von Vektoren nur die lineare Unabhängigkeit (oder nur das Erzeugendensystem-Sein) zu prüfen, um nachzuweisen, dass man eine Basis dieses Unterraums gefunden hat. Natürlich kann dieses Kriterium nicht angewandt werden, wenn man die Dimension des Unterraums noch nicht kennt.

Wir wollen nochmal an einem Beispiel veranschaulichen, warum die zusätzliche Bedingung bei der Auswahl eines Vektors einer Stufe nötig ist.

Beispiel. Wir betrachten den Unterraum $\text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$ von \mathbb{R}^3 . Dieser Unterraum ist besonders einfach, nämlich der \mathbb{R}^3 selbst, denn als Linearkombinationen der Standardbasis können wir jeden Vektor in \mathbb{R}^3 erhalten. Wir kennen auch schon eine Auswahl aus diesen Vektoren, die eine Basis vom \mathbb{R}^3 liefert, nämlich die Standardbasis. Wir wollen aber sehen, ob man in dem obigen allgemeinen Verfahren hier im zweiten Schritt einen anderen Vektor in der zweiten Stufe hätte wählen können. Die Matrix, die die Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ als Spalten hat, ist bereits in Zeilenstufenform:

$$\begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Wählt man allerdings in der zweiten Stufe anstatt $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ den Vektor $\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$, so ist die Matrix aus dem ersten, dritten und vierten Vektor, d.h.

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

nicht in Zeilenstufenform. Bringt man sie in Zeilenstufenform durch Vertauschen der letzten beiden Zeilen, so sieht man, dass diese weniger Stufen als

Vektoren hat. Insbesondere sind die Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ linear abhängig und damit auch keine Basis von \mathbb{R}^3 .

Informell gesagt kann man also beim obigen Verfahren - bis auf die Ausnahme mit den Nullen - dem Slogan „ein Vektor pro Stufe“ folgen, um eine Basis aus einem Erzeugendensystem eines Unterraums auszuwählen. Wie bereits erläutert, ist es allerdings möglich, dass man andere Auswahl von Vektoren - die dieser Faustregel nicht folgen - treffen könnte, die auch eine Basis liefern. Das Verfahren sichert einem also mindestens eine Basis, gibt aber in der vorgestellten Form keine Auskunft darüber, welche Möglichkeiten man sonst hätte, eine Basis aus dem vorgegebenen Erzeugendensystem auszuwählen.

Nun nehmen wir an, dass wir zu einem Unterraum U bereits eine Basis vorgegeben haben. Dann wissen wir, dass jedes Element von U sich eindeutig als Linearkombination der Vektoren der Basis darstellen lässt. Jetzt wollen wir uns mit der Frage beschäftigen, wie man diese Linearkombination tatsächlich finden kann, also wie man die Koordinaten von einem Vektor bezüglich einer Basis herausfinden kann. Ein etwas allgemeineres Verfahren lernen wir später kennen, nachdem wir einige verwandte theoretische Grundlagen studiert haben. Eine mögliche Motivation dafür ist es, die Koordinaten von einem Vektor in einem Unterraum intrinsisch angeben zu wollen.

Beispiel. Wir betrachten die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, die, wie man leicht nachprüfen kann, drei linear unabhängige Vektoren in \mathbb{R}^3 sind und somit eine Basis vom \mathbb{R}^3 bilden. Dann wissen wir, dass sich jeder Vektor $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ mit beliebigen $x, y, z \in \mathbb{R}$ eindeutig als Linearkombination der Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ schreiben lässt, d.h., dass es eindeutige reelle Zahlen $\lambda_1, \lambda_2, \lambda_3$ gibt, sodass

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda_1 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 0 \\ 5 \\ 5 \end{pmatrix} + \lambda_3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

gilt. Um diese Koordinaten von dem Vektor $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ bezüglich der Basis $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ von \mathbb{R}^3 zu bestimmen, behandeln wir nun $x, y, z \in \mathbb{R}$ als Parameter. Dann haben wir ein lineares Gleichungssystem in Variablen $\lambda_1, \lambda_2, \lambda_3$ zu lösen.

Das machen wir nach der bereits erlernten Methode. Die zugehörige erweiterte Koeffizientenmatrix ist gegeben durch

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & x \\ 2 & 5 & 0 & y \\ 3 & 5 & 1 & z \end{array} \right).$$

Wir bringen diese Matrix durch elementare Zeilentransformationen auf Zeilenstufenform:

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & 0 & 0 & x \\ 2 & 5 & 0 & y \\ 3 & 5 & 1 & z \end{array} \right) &\xrightarrow{II-2\cdot I, III-3\cdot I} \left(\begin{array}{ccc|c} 1 & 0 & 0 & x \\ 0 & 5 & 0 & y-2x \\ 0 & 5 & 1 & z-3x \end{array} \right) \\ &\xrightarrow{III-II} \left(\begin{array}{ccc|c} 1 & 0 & 0 & x \\ 0 & 5 & 0 & y-2x \\ 0 & 0 & 1 & z-x-y \end{array} \right) \\ &\xrightarrow{III-II} \left(\begin{array}{ccc|c} 1 & 0 & 0 & x \\ 0 & 1 & 0 & \frac{y-2x}{5} \\ 0 & 0 & 1 & z-x-y \end{array} \right). \end{aligned}$$

Da die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 5 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ eine Basis von \mathbb{R}^3 bilden, wussten wir bereits, dass das lineare Gleichungssystem genau eine Lösung besitzen wird. Nun haben wir diese Lösung ermittelt: Diese ist durch $\lambda_1 = x$, $\lambda_2 = \frac{y-2x}{5}$, $\lambda_3 = z - x - y$ gegeben. Der Vektor $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ lässt sich schreiben als

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \frac{y-2x}{5} \cdot \begin{pmatrix} 0 \\ 5 \\ 0 \end{pmatrix} + (z-x-y) \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

In diesem Fall war es recht einfach, dieses Gleichungssystem zu lösen; im Allgemeinen wird das Auffinden der Koeffizienten komplizierter sein. Wir werden später ein weiteres Verfahren kennenlernen, mit dem man die Koordinaten eines Vektors bezüglich einer Basis etwas systematischer bestimmen kann. Das ist manchmal auch vom Rechenaufwand her vorteilhaft, meist ist es aber aus theoretischen Gründen nützlich.

4 Lineare Abbildungen und Matrizen

Wir wollen uns nun den linearen Abbildungen widmen. Dabei verfolgen wir mehrere Ziele. Zum einen würden wir gerne in der Lage sein, zwei Unterräume

von \mathbb{R}^n oder gar einen Unterraum von \mathbb{R}^n und einen von \mathbb{R}^k zu vergleichen. Inzwischen kennen wir eine Kenngröße, die wir vergleichen können, nämlich die Dimension des Unterraums. Doch gibt es eine sinnvolle Möglichkeit, Unterräume gleicher Dimension zu vergleichen? Sind sie in gewissem Sinne „dasselbe“? Wie könnte man überhaupt eine Aussage von diesem Typ präzise machen? Wie kann man die Koordinatenänderung beschreiben, die dadurch entsteht, dass man das Koordinatensystem verändert, etwa wie bei einem Kameranachrichten? Und wie kann man ein dreidimensionales Objekt sinnvoll in der Ebene, also etwa auf dem Bildschirm oder auf Papier, abbilden? Diese Fragen wollen wir, zumindest in Ansätzen, in diesem Kapitel untersuchen.

Wir erinnern uns, dass eine Möglichkeit, beliebige Mengen zueinander in Beziehung zu setzen, darin bestand, Abbildungen zwischen diesen Mengen zu finden. Um die Anzahl der Elemente von endlichen (und eigentlich auch von unendlichen) Mengen zu vergleichen, haben wir Bijektionen zwischen den Mengen gesucht. Auch für Unterräume brauchen wir einen Abbildungsbegriff, allerdings ist es meist nicht ratsam, beliebige Abbildungen zwischen Unterräumen zu betrachten. Da Unterräume nämlich Mengen mit besonderen Eigenschaften sind, wollen wir, dass die Abbildungen diese besondere Eigenschaften erhalten. Damit hängt es auch zusammen, dass diese spezielle Abbildungen - die linearen Abbildungen - in ihrer Definition ganz ähnliche Eigenschaften wie (U1) und (U2) für Unterräume haben werden. Obwohl die Definition von linearen Abbildungen auf Anhieb ein wenig kompliziert erscheinen mag, stellt es sich heraus, dass lineare Abbildungen in gewisser Hinsicht die einfachsten Abbildungen überhaupt sind. Das werden wir im Laufe unserer Untersuchungen noch etwas präziser machen können.

Zunächst erinnern wir an die allgemeine Definition von Abbildungen, die aus Mathematischen Grundlagen 1 bekannt ist.

Definition 4.1. Seien X, Y beliebige Mengen. Eine **Abbildung** $f: X \rightarrow Y$ ist eine Zuordnung, die jedem Element von X genau ein Element von Y zuordnet. Die Menge X heißt **Quelle** und die Menge Y das **Ziel** der Abbildung f .

Wir erinnern uns nochmal daran, dass eine Vorschrift ohne Angabe von Quelle und Ziel noch keine Abbildung darstellt, und dieselbe Vorschrift mag je nach Quelle und Ziel eine Abbildung definieren oder auch nicht, und die Eigenschaften der gegebenenfalls definierten Abbildung sich in Abhängigkeit von Quelle und Ziel ändern können. Daher wollen wir uns nochmal vergegenwärtigen, dass die Angabe einer Abbildung stets aus einer Quellmenge, einer Zielmenge und einer Zuordnungsvorschrift besteht.

Es sei ferner noch angemerkt, dass eine Zuordnung, die keine Abbildung definiert, unter Umständen kein wirkliches mathematisches Objekt ist, etwa wenn einem $x \in X$ ein Objekt zugeordnet wird, das gar nicht wirklich definiert ist (was bei uns vor allem passieren kann, wenn man durch 0 teilt oder eine Wurzel einer negativen Zahl zu ziehen versucht, da dadurch keine reelle

Zahlen definiert werden; es kann im Allgemeinen aber auch andere Probleme geben).

Nun wollen wir zu der Definition einer linearen Abbildung kommen.

Definition 4.2. Seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^k$ Unterräume von \mathbb{R}^n bzw. \mathbb{R}^k . Eine Abbildung $\varphi: U \rightarrow V$ heißt **linear**, falls sie die beiden folgenden Bedingungen erfüllt:

(A1) Für je zwei Elemente $u_1, u_2 \in U$ von U ist es egal, ob diese zunächst addiert und dann in φ eingesetzt werden, oder ob die beiden Vektoren zunächst einzeln in die Abbildung φ eingesetzt werden und dann die Ergebnisse addiert werden. In Formeln:

$$\varphi(u_1 + u_2) = \varphi(u_1) + \varphi(u_2).$$

(A2) Für jedes Element $w \in U$ und jede reelle Zahl $\lambda \in \mathbb{R}$ ist es egal, ob man zunächst die Skalarmultiplikation $\lambda \cdot u$ durchführt und dann das Ergebnis in φ einsetzt, oder ob man zunächst w in φ einsetzt und dann das Ergebnis mit $\lambda \in \mathbb{R}$ multipliziert. In Formeln:

$$\varphi(\lambda \cdot w) = \lambda \cdot \varphi(w).$$

Man beachte, dass man nur von einer linearen Abbildung sprechen kann, wenn sowohl Quellmenge als auch Zielmenge jeweils ein Unterraum sind. Wir werden überwiegend den Fall $U = \mathbb{R}^n$ und $V = \mathbb{R}^k$ betrachten, da dieser am einfachsten ist. Zum Teil sieht man die Notwendigkeit, mit Unterräumen zu arbeiten, unmittelbar an der Definition: Wäre etwa $u_1 + u_2$ für Elemente $u_1, u_2 \in U$ nicht wieder ein Element von U , so wüßten wir nicht, wie man $u_1 + u_2$ in die Abbildung φ einsetzen kann. Genauso wenig wäre auch $\varphi(\lambda \cdot w)$ definiert, wenn $\lambda \cdot w$ kein Element von U wäre. Auch die Eigenschaft, dass V ein Unterraum ist, wird benötigt; darauf gehen wir jedoch nicht näher ein.

Man beachte auch, dass man zunächst eine Abbildung braucht, um eine lineare Abbildung zu haben. Das muss also als erstes geprüft werden, wenn man überprüfen will, ob eine Zuordnung eine lineare Abbildung definiert. In den meisten von uns betrachteten Fällen wird das allerdings klar sein. Es sei nochmal angemerkt, dass dann sowohl die Bedingung (A1) als auch die Bedingung (A2) erfüllt sein müssen, damit die Abbildung linear ist.

Ferner wollen wir nochmal die Ähnlichkeit der Bedingungen in dieser Definition zu den Bedingungen (U1) und (U2) aus der Definition der Unterräume hervorheben. Diese Art von Bedingungen erfasst, informell gesagt, das Wesentliche des Begriffs „Linearität“, ob für Mengen oder für Abbildungen. (Das ist natürlich keine präzise Aussage, sondern nur bildlich zu verstehen.)

Nun wollen wir einige Beispiele für diese recht abstrakte Definition sehen.

Beispiel. • Die Abbildung $\psi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die jedem Element $\begin{pmatrix} x \\ y \end{pmatrix}$ der Quelle das (eindeutige) Element $\begin{pmatrix} 2x + y \\ 3x \end{pmatrix}$ des Ziels zuordnet, ist linear. Das wollen wir nachprüfen.

Um sich vorher zu vergegenwärtigen, wie diese Abbildung „funktio- niert“, setzen wir einige Werte in ψ ein. (Das ist mathematisch nicht notwendig, sondern dient nur der Veranschaulichung.) So gilt beispie- lweise:

$$\begin{aligned} \psi\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right) &= \begin{pmatrix} 2 \cdot 1 + 2 \\ 3 \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \\ \psi\left(\begin{pmatrix} z \\ 0 \end{pmatrix}\right) &= \begin{pmatrix} 2 \cdot z + 0 \\ 3 \cdot z \end{pmatrix} = \begin{pmatrix} 2z \\ 3z \end{pmatrix} \end{aligned}$$

zu (A1): Wir müssen nachprüfen, dass für zwei beliebige Elemente der Quelle \mathbb{R}^2 , also für $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ das Einsetzen der Summe dasselbe Ergebnis liefert wie jeweiliges Einsetzen in ψ , gefolgt von Addition. Wir führen beide Rechnungen durch, was sich als recht einfach erweist:

$$\psi\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = \psi\left(\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}\right) = \begin{pmatrix} 2(x_1 + x_2) + (y_1 + y_2) \\ 3(x_1 + x_2) \end{pmatrix}$$

Andererseits ist

$$\begin{aligned} \psi\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right) + \psi\left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) &= \begin{pmatrix} 2x_1 + y_1 \\ 3x_1 \end{pmatrix} + \begin{pmatrix} 2x_2 + y_2 \\ 3x_2 \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 + y_1 + 2x_2 + y_2 \\ 3x_1 + 3x_2 \end{pmatrix}. \end{aligned}$$

Die beiden Ergebnisse sehen zwar nicht genau gleich aus, können aber durch Sortieren von Summanden und Ausklammern gemein- samer Faktoren ineinander überführt werden:

$$\begin{aligned} \begin{pmatrix} 2x_1 + y_1 + 2x_2 + y_2 \\ 3x_1 + 3x_2 \end{pmatrix} &= \begin{pmatrix} (2x_1 + 2x_2) + (y_1 + y_2) \\ 3x_1 + 3x_2 \end{pmatrix} \\ &= \begin{pmatrix} 2(x_1 + x_2) + (y_1 + y_2) \\ 3(x_1 + x_2) \end{pmatrix}. \end{aligned}$$

Für alle $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ gilt also

$$\psi\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = \psi\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right) + \psi\left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right),$$

und somit ist die Bedingung (A1) für ψ erfüllt.

zu (A2): Der Nachweis, dass die Bedingung (A2) für ψ erfüllt ist, ist von ganz ähnlicher Natur. Sei $\begin{pmatrix} x \\ y \end{pmatrix}$ also ein beliebiges Element der Quelle \mathbb{R}^2 , und sei $\lambda \in \mathbb{R}$ eine beliebige reelle Zahl. Wir müssen prüfen, ob das Einsetzen des skalaren Vielfachen $\lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix}$ in ψ dasselbe Ergebnis liefert wie das λ -fache von $\psi \left(\begin{pmatrix} x \\ y \end{pmatrix} \right)$. Das rechnen wir wieder einfach nach:

$$\psi \left(\lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix} \right) = \psi \left(\begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix} \right) = \begin{pmatrix} 2\lambda x + \lambda y \\ 3\lambda x \end{pmatrix},$$

und andererseits ist

$$\lambda \cdot \psi \left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = \lambda \cdot \begin{pmatrix} 2x + y \\ 3x \end{pmatrix} = \begin{pmatrix} \lambda \cdot (2x + y) \\ 3\lambda x \end{pmatrix}.$$

Löst man nun in dem zweiten Fall die Klammer auf, so erhält man auch hier

$$\begin{pmatrix} \lambda \cdot (2x + y) \\ 3\lambda x \end{pmatrix} = \begin{pmatrix} 2\lambda x + \lambda y \\ 3\lambda x \end{pmatrix}$$

und somit gilt

$$\psi \left(\lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix} \right) = \lambda \cdot \psi \left(\begin{pmatrix} x \\ y \end{pmatrix} \right)$$

für alle $\lambda \in \mathbb{R}$ und $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$.

Insgesamt haben wir also gezeigt, dass ψ eine lineare Abbildung ist.

- Nun wollen wir ein Beispiel einer nicht-linearen Abbildung anführen. Dafür wollen wir zeigen, dass die Abbildung $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die durch $g \left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} x^2 \\ y \end{pmatrix}$ gegeben ist, nicht linear ist. Dafür reicht es - anders in dem vorherigen Beispiel - einen einzigen Vektor $\begin{pmatrix} x \\ y \end{pmatrix}$ in der Quelle \mathbb{R}^2 und eine reelle Zahl $\lambda \in \mathbb{R}$ anzugeben, für die die Bedingung (A2) nicht erfüllt ist, denn dann kann g nicht mehr linear sein. Hier wählen wir den Vektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und die Konstante $\lambda = 2 \in \mathbb{R}$: Das ist zwar nicht der einzige Fall, in dem die Bedingung (A2) nicht erfüllt ist, aber es reicht, einen einzigen solchen Fall anzuführen. Nun überprüfen

wir, dass die beiden Seiten der Gleichung in der Bedingung (A2) für g unterschiedliche Ergebnisse liefert, wenn wir den Vektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{R}$ und $\lambda = 2$ einsetzen. Die linke Seite liefert

$$g\left(2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = g\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2^2 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}.$$

Auf der rechten Seite ergibt sich hingegen

$$2 \cdot g\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = 2 \cdot \begin{pmatrix} 1^2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

Die Werte stimmen nicht überein, folglich ist g nicht linear.

Informell kann man sagen, dass gerade die Abbildungen linear sind, die man (in jeder Komponente) als Summe oder Differenz der konstanten Vielfachen der Komponenten des eingesetzten Vektors schreiben kann. Ähnlich wie bei linearen Gleichungssystemen dürfen - als Faustregel - keine Quadrate, keine Kuben, keine höheren Potenzen der Komponenten des eingesetzten Vektors vorkommen, auch keine „komplizierten“ Funktionen wie Sinus oder

Logarithmus; im Grunde sollte das Bild eines Vektors $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ so ähnlich aus-

sehen wie die linke Seite des linearen Gleichungssystems in Standardform in Variablen x_1, \dots, x_n . Diese Heuristik wollen wir im folgenden präziser machen.

Zunächst wollen wir nochmal hervorheben, dass Addition von Konstanten durchaus für die Linearität hinderlich sein kann. Das ist anders, als häufig aus der Schule bekannter Begriff der Linearität. Darauf wollen wir im Rahmen von einem Beispiel etwas genauer eingehen.

Beispiel. Wir betrachten die Abbildung $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$, die durch

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 2x + 5 \\ y \end{pmatrix}$$

gegeben ist. Hier haben wir in der ersten Komponente einen konstanten Summanden 5, den man als Grund dafür ansehen kann, dass f nicht linear ist. Wir wollen allerdings nachrechnen, dass f tatsächlich nicht linear ist. Tatsächlich erfüllt f weder die Bedingung (A1) noch die Bedingung (A2); wir wollen allerdings nur für (A1) ein Gegenbeispiel liefern, da wir dann schon

wissen, dass f nicht linear sein kann. Wir betrachten die Vektoren $u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

und $u_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Das ist wieder nicht das einzige mögliche Gegenbeispiel, allerdings reicht auch in diesem Fall ein Gegenbeispiel aus, und wir haben uns eben für besonders „einfache“ Vektoren u_1, u_2 entschieden. Nun überprüfen wir, dass es nicht dasselbe liefert, die Vektoren u_1 und u_2 zunächst zu addieren und dann f auf die Summe anzuwenden, wie wenn wir zunächst f auf die einzelnen Vektoren anwenden und dann die Summe der Bilder bilden. Die Rechnung liefert:

$$f(u_1 + u_2) = f\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right) = f\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2 \cdot 1 + 5 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \\ 0 \end{pmatrix}$$

Andererseits ist

$$\begin{aligned} f(u_1) + f(u_2) &= f\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) + f\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2 \cdot 1 + 5 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \cdot 0 + 5 \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 7 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 12 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

Da die Ergebnisse nicht gleich sind, ist die Bedingung (A1) für f nicht erfüllt und f somit auch keine lineare Abbildung.

Wir erinnern uns, dass an der Schule eine Funktion der Form $h: \mathbb{R} \rightarrow \mathbb{R}, h(x) = 2x + 5$, die ja eine ähnliche Form wie das hier betrachtete f hat, als „linear“ bezeichnet wurde. Dafür gibt es zwar gute Gründe, nämlich die Tatsache, dass der Graph dieser Funktion eine Gerade ist, allerdings keine Ursprungsgerade; aber wir werden uns an die in der mathematischen Literatur üblichere Konvention halten, nach der dies keine lineare Abbildung ist. Stattdessen würde man üblicherweise sowohl f als auch h , oder allgemein eine Abbildung, die nur durch Addition von Konstanten sich von einer linearen Abbildung unterscheidet, als *affin-linear* bezeichnen. Wir werden später nochmal über affine Unterräume und affin-lineare Abbildungen zumindest kurz sprechen. Es sei noch angemerkt, dass der Graph einer linearen Abbildung $\mathbb{R} \rightarrow \mathbb{R}$, wobei linear in „unserem“ Sinne gemeint ist, auch stets eine Gerade sein wird, allerdings immer eine Ursprungsgerade; und ähnliche höherdimensionalen Aussagen kann man über beliebige lineare Abbildungen treffen.

Als nächstes wollen wir uns überlegen, dass eine lineare Abbildung bereits durch einige wenige Werte festgelegt ist, und sich einfach durch diese Werte beschreiben lässt. Wir fangen damit an, dass wir uns diese Beschreibung an einem Beispiel erarbeiten.

Beispiel. Sei eine *lineare* Abbildung $\varphi: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ vorgegeben. Als erstes nehmen wir an, dass wir einen Wert von φ bereits vorgegeben haben, und zwar sei

$$\varphi \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}.$$

Da φ linear ist und dadurch insbesondere der Bedingung (A2) genügt, ist durch diesen Wert nicht nur der Wert von φ für den Vektor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ festgelegt, sondern auch für alle Vielfache von diesem Vektor, denn nach (A2) (in der zweiten Gleichheit verwendet) gilt :

$$\varphi \left(\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} \right) = \varphi \left(x \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = x \cdot \varphi \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = x \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ 2x \\ 3x \\ 0 \end{pmatrix}$$

für jede reelle Zahl x . Durch den einen vorgegebenen Wert haben wir also bereits Werte auf unendlich vielen Vektoren der Quelle festgelegt, nämlich auf allen Vielfachen des vorgegeben Vektors.

Allerdings kann man weder durch Summen noch durch Vielfache des Vektors $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ den Vektor $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ erhalten. Der Wert von φ auf diesem ist noch nicht festgelegt. Also setzen wir diesen Wert wieder nach Belieben fest, etwa

$$\varphi \left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}.$$

Dadurch sind nun, nach demselben Argument wie vorher, zusätzlich die Werte von φ auf allen Vektoren der Form $\begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix}$ für beliebige $y \in \mathbb{R}$ festgelegt, und zwar durch

$$\varphi \left(\begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 4y \\ 5y \\ 6y \\ 0 \end{pmatrix}.$$

Aber jetzt ist φ auch noch auf weiteren Vektoren bereits fest: Wir haben nämlich noch gar nicht ausgenutzt, dass für die lineare Abbildung φ auch die

Bedingung (A1) erfüllt ist. Diese liefert nämlich, auf die Vektoren $\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix}$ für beliebige $x, y \in \mathbb{R}$ angewandt:

$$\begin{aligned} \varphi \left(\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \right) &= \varphi \left(\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \right) = \varphi \left(\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} \right) + \varphi \left(\begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \right) \\ &= \begin{pmatrix} x \\ 2x \\ 3x \\ 0 \end{pmatrix} + \begin{pmatrix} 4y \\ 5y \\ 6y \\ 0 \end{pmatrix} = \begin{pmatrix} x + 4y \\ 2x + 5y \\ 3x + 6y \\ 0 \end{pmatrix}. \end{aligned}$$

Also ist die Abbildung φ , nachdem wir nur zwei Werte festgelegt haben - nämlich für die beiden Standardbasisvektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ - bereits auf dem ganzen zweidimensionalen Unterraum

$$\left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \mid x, y \in \mathbb{R} \right\}$$

von \mathbb{R}^3 festgelegt. Nun liegt die Vermutung nah, dass man, um die Abbildung φ auf ganz \mathbb{R}^3 festzulegen, nur noch einen weiteren Wert braucht, und man diesen etwa für den letzten Standardbasisvektor auswählen kann; also setzen wir

$$\varphi \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}.$$

Wir hatten uns bereits überlegt, dass man jeden Vektor in \mathbb{R}^3 als Linearkombination der Standardbasisvektoren schreiben kann, und dass die Koeffizienten dabei besonders einfach zu bestimmen sind, da es gerade die entsprechenden Komponenten des Vektors sind. In Formeln heißt das, dass für alle $x, y, z \in \mathbb{R}$ gilt:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Nun können wir diesen Ausdruck in die lineare Abbildung $\varphi: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ einsetzen und erhalten

$$\varphi \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) = \varphi \left(x \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right).$$

Wir machen uns jetzt zunutze, dass φ eine lineare Abbildung ist und man somit (nach (A1)) die Summe aus φ „herausziehen“ kann. Wir erhalten also:

$$\varphi \left(x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) = \varphi \left(x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) + \varphi \left(y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) + \varphi \left(z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right).$$

Im nächsten Schritt können wir in jedem der Summanden die Linearität von φ ausnutzen und dank der Bedingung (A2) den skalaren Faktor „herausziehen“. Wir erhalten also:

$$\begin{aligned} & \varphi \left(x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) + \varphi \left(y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) + \varphi \left(z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \\ &= x\varphi \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) + y\varphi \left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) + z\varphi \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right). \end{aligned}$$

Nun haben wir die Werte $\varphi \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right)$, $\varphi \left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right)$, $\varphi \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$ gerade festgelegt und können diese in unsere Rechnung einsetzen:

$$x \cdot \varphi \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) + y \cdot \varphi \left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) + z \cdot \varphi \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) = x \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}.$$

Zusammenfassend ist also

$$\begin{aligned} \varphi \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) &= x \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1x + 4y + 7z \\ 2x + 5y + 8z \\ 3x + 6y + 9z \\ 0x + 0y + 0z \end{pmatrix}. \end{aligned}$$

Tatsächlich haben wir also dadurch, dass wir die Werte von φ für die Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ der Standardbasis der Quelle gewählt haben, schon alle Werte der linearen Abbildung φ festgelegt. Wir haben auch soeben explizit angegeben, wie aus diesen Werten und den Komponenten des Vektors $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ der Wert von $\varphi\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right)$ ermittelt werden kann.

Man kann sich vorstellen, dass bis auf die Anzahl der Standardbasisvektoren nichts an diesem Verfahren speziell von den gewählten Zahlen abhängt. Da die Werte einer linearen Abbildung auf der Standardbasis der Quelle die lineare Abbildung komplett bestimmen, sind diese Werte eine wichtige Kenngröße der Abbildung. Um sich diese systematisch zu notieren, schreibt man die Bildvektoren als Spalten in eine Matrix.

Wir betrachten eine lineare Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^k$. Die Standardbasis von \mathbb{R}^n genau n Standardbasisvektoren, die wir abkürzend mit

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

bezeichnen. (Wir merken an, dass diese Bezeichnung leicht missverständlich sein kann, da e_1 als Standardbasisvektor in \mathbb{R}^3 eine andere Bedeutung hat als e_1 in \mathbb{R}^4 . Wir werden die Bezeichnung jedoch stets so verwenden, dass aus dem Kontext klar sein sollte, welche Dimension und dementsprechend welcher Vektor e_1 gemeint ist.)

Die Abbildung φ ist also unserer Vorüberlegung zufolge durch die Werte $\varphi(e_1), \varphi(e_2), \dots, \varphi(e_n)$ bereits festgelegt. Man kann auch zeigen, dass beliebige Werte für $\varphi(e_1), \varphi(e_2), \dots, \varphi(e_n)$ tatsächlich zu einer linearen Abbildung φ gehören. Einen Schritt in diese Richtung wollen wir nun unternehmen. Zunächst geben wir der Matrix, die aus diesen Bildvektoren besteht, einen Namen.

Definition 4.3. Sei $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^k$ eine lineare Abbildung. Die $k \times n$ -Matrix, die die Einträge der Vektoren $\varphi(e_1), \varphi(e_2), \dots, \varphi(e_n)$ als Spalten hat, wird die **darstellende Matrix** von φ genannt.

Wir erinnern uns nochmal, dass $k \times n$ -Matrix unserer Konvention zufolge heißt, dass die Matrix k Zeilen und n Spalten besitzt. Tatsächlich muss die

darstellende Matrix einer linearen Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^k$ genau n Spalten haben, da diese gerade von den Bildvektoren $\varphi(e_1), \varphi(e_2), \dots, \varphi(e_n)$ gebildet werden. Jeder dieser Bildvektoren ist also ein Element von \mathbb{R}^k , hat also k Komponenten, also hat die darstellende Matrix von φ auch genau k Zeilen.

Beispiel. Ein Beispiel für eine darstellende Matrix haben wir im Wesentlichen gesehen: Die Abbildung $\varphi: \mathbb{R}^3 \rightarrow \mathbb{R}^4$, die durch die Werte

$$\varphi \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}, \quad \varphi \left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 4 \\ 5 \\ 6 \\ 0 \end{pmatrix}, \quad \varphi \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 7 \\ 8 \\ 9 \\ 0 \end{pmatrix}$$

festgelegt ist und nach unserer Überlegung die Vorschrift

$$\varphi \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) = \begin{pmatrix} 1x + 4y + 7z \\ 2x + 5y + 8z \\ 3x + 6y + 9z \\ 0x + 0y + 0z \end{pmatrix}$$

hat, ist eine lineare Abbildung mit der darstellenden Matrix

$$\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \\ 0 & 0 & 0 \end{pmatrix}.$$

Wir wissen bereits, dass die Spalten dieser Matrix die Abbildung φ ganz festlegen. Insbesondere bemerken wir, dass dieselben Zahlen wie in dieser Matrix in der Abbildungsvorschrift von φ vorkommen.

Will man etwa $\varphi \left(\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \right)$ ermitteln, so rechnet man

$$\begin{aligned} \varphi \left(\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \right) &= \begin{pmatrix} 1 \cdot (-1) + 4 \cdot 2 + 7 \cdot (-1) \\ 2 \cdot (-1) + 5 \cdot 2 + 8 \cdot (-1) \\ 3 \cdot (-1) + 6 \cdot 2 + 9 \cdot (-1) \\ 0 \cdot (-1) + 0 \cdot 2 + 0 \cdot (-1) \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned}$$

Wir bemerken, dass wir gerade eine Linearkombination der Spalten der Matrix ausgerechnet haben, und zwar eine, in der die Koeffizienten die Einträge des eingesetzten Vektors sind.

An dieser Stelle sei noch angemerkt, dass es durchaus vorkommen kann, wie wir soeben gesehen haben, dass eine lineare Abbildung beim Auswerten auf einem Vektor den Nullvektor liefert. Es ist sogar immer der Fall, wenn man den Nullvektor der Quelle einsetzt, wie wir später nochmal sehen werden. Es kann allerdings wie hier auch passieren, dass Auswerten auf anderen Vektoren ebenfalls den Nullvektor liefert.

Nun behaupten wir, aus der darstellenden Matrix einer linearen Abbildung die lineare Abbildung wieder rekonstruieren zu können. Bevor wir das allgemeine Verfahren dazu kennenlernen, die Matrix-Vektor-Multiplikation, leiten wir uns dieses exemplarisch an einem Beispiel her.

Beispiel. Wir betrachten die Matrix

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

Wir fragen uns, ob B die darstellende Matrix einer linearen Abbildung ist. Da B drei Spalten hat, in denen die Werte der linearen Abbildung auf der Standardbasisvektoren stehen müssten, schließen wir, dass eine solche Abbildung, sollte sie denn überhaupt existieren, auf \mathbb{R}^3 definiert werden würde und Werte in \mathbb{R}^2 annehmen würde. Wir bezeichnen diese hypothetische Abbildung mit $\psi: \mathbb{R}^3 \rightarrow \mathbb{R}^2$. Wir wollen herausfinden, was die Abbildungsvorschrift für ψ sein sollte. Wir wissen bereits, da wir B als die darstellende Matrix von ψ haben wollen, dass notwendigerweise

$$\psi \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} -\frac{1}{2} \\ 2 \end{pmatrix}, \quad \psi \left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ -\frac{1}{3} \end{pmatrix}, \quad \psi \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

gelten muss. Nun gehen wir wie eben vor: Wir wissen, dass jeder Vektor in der Quelle \mathbb{R}^3 der hypothetischen Abbildung ψ sich schreiben lässt als Linearkombination der Vektoren der Standardbasis, und zwar als

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Sollte ψ linear sein, dann muss es ja Bedingungen (A1) und (A2) erfüllen, also müssen wir dieselbe Rechnung wie oben durchführen können und würden erhalten:

$$\begin{aligned} \psi \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) &= \psi \left(x \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \\ &= \psi \left(x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right) + \psi \left(y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) + \psi \left(z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \end{aligned}$$

$$= x\psi\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) + y\psi\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) + z\psi\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right).$$

Für diese speziellen Vektoren wurden ja die Werte durch die Matrix festgelegt, sodass wir notwendigerweise für jede lineare Abbildung ψ , deren darstellende Matrix B ist, folgende Identität für alle $x, y, z \in \mathbb{R}$ haben müssen:

$$\begin{aligned} \psi\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) &= x\psi\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) + y\psi\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) + z\psi\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) \\ &= x \cdot \begin{pmatrix} -\frac{1}{2} \\ 2 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ -\frac{1}{3} \end{pmatrix} + z \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{2}x + 0y + 3z \\ 2x - \frac{1}{3}y + 1z \end{pmatrix}. \end{aligned}$$

Tatsächlich kann man leicht nachprüfen, dass diese Vorschrift eine lineare Abbildung $\psi: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definiert, und die darstellende Matrix dieser Abbildung gerade B ist. Wir haben also aus der Matrix B eine lineare Abbildung gebaut, die B als darstellende Matrix hat. Auch hier lässt sich beobachten, dass die Einträge des Vektors $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$, der in ψ eingesetzt wurde, als Koeffizienten für die Spalten der Matrix B fungiert haben. Das motiviert die folgende Operation, die das allgemeine Verfahren zum Bestimmen der Abbildung, die zu einer Matrix gehört, liefern wird.

Matrix-Vektor-Multiplikation. Sei eine $k \times n$ -Matrix C und ein Vektor

$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ in \mathbb{R}^n vorgegeben. (An dieser Stelle ist es wichtig, dass die Anzahl der

Komponenten des Vektors genau mit der Anzahl der Spalten von der Matrix C übereinstimmt. Tut sie das nicht, so sagt man, dass das Produkt von C

mit $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ nicht definiert ist.) Wir verwenden unsere übliche Notation für die

Einträge der Matrix C :

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} & \dots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \dots & c_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & c_{k3} & \dots & c_{kn} \end{pmatrix}.$$

Dann ist das Produkt der Matrix C mit dem Vektor $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ wieder ein Vektor,

diesmal in \mathbb{R}^k (die Anzahl der Komponenten vom Ergebnis ist also gerade die Anzahl der Zeilen der Matrix C), und diesen erhält man als Linearkombination der Spalten von C mit den Koeffizienten x_1, x_2, \dots, x_n . Wir nutzen also gerade die Einträge des Vektors als Koeffizienten, daher muss deren Anzahl mit der Anzahl der Spalten von C übereinstimmen. In Formeln heißt das,

dass das Produkt $C \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ wie folgt definiert ist:

$$\begin{aligned} C \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} &= \begin{pmatrix} c_{11} & c_{12} & c_{13} & \cdots & c_{1n} \\ c_{21} & c_{22} & c_{23} & \cdots & c_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{k1} & c_{k2} & c_{k3} & \cdots & c_{kn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \\ &= x_1 \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{k1} \end{pmatrix} + x_2 \begin{pmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{k2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} c_{1n} \\ c_{2n} \\ \vdots \\ c_{kn} \end{pmatrix} \\ &= \begin{pmatrix} c_{11}x_1 + c_{12}x_2 + c_{13}x_3 + \cdots + c_{1n}x_n \\ c_{21}x_1 + c_{22}x_2 + c_{23}x_3 + \cdots + c_{2n}x_n \\ \vdots \\ c_{k1}x_1 + c_{k2}x_2 + c_{k3}x_3 + \cdots + c_{kn}x_n \end{pmatrix}. \end{aligned}$$

Wir vergleichen die Matrix-Vektor-Multiplikation mit den Ergebnissen unserer Vorüberlegungen an einem Beispiel.

Beispiel. Sei A die 3×2 -Matrix

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

Nach Definition können wir diese Matrix mit Vektoren in \mathbb{R}^2 multiplizieren und erhalten dadurch Vektoren in \mathbb{R}^3 . Nach Definition ist

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

Gleichzeitig haben wir uns überlegt, dass jede lineare Abbildung $g: \mathbb{R}^2 \rightarrow \mathbb{R}^3$, deren darstellende Matrix gerade A ist, folgende Form haben muss:

$$\begin{aligned}
 g\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) &= g\left(x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\
 &= g\left(x \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + g\left(y \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\
 &= xg\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) + yg\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) \\
 &= x \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix} \\
 &= \begin{pmatrix} 1 \cdot x + 4 \cdot y \\ -1 \cdot x + 2 \cdot y \\ 0 \cdot x + 3 \cdot y \end{pmatrix}.
 \end{aligned}$$

Das ist passenderweise genau das Ergebnis der Matrix-Vektor-Multiplikation, die wir soeben ausgeführt haben. In diesem Beispiel sehen wir also, dass die Matrix-Vektor-Multiplikation genau die Abbildung liefert, die die einzig mögliche potentiell lineare Abbildung mit dieser darstellenden Matrix ist. Dass das auch allgemein so ist, ist die Aussage des folgenden Satzes.

Der folgende Satz fasst das Verhältnis von linearen Abbildungen und Matrizen etwas genauer zusammen.

Satz 4.4. 1) Sei A eine $k \times n$ -Matrix. Dann ist die Abbildung $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^k$, die für alle $v \in \mathbb{R}^n$ durch $T_A(v) = A \cdot v$ definiert ist, eine lineare Abbildung mit darstellenden Matrix A .

2) Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^k$ eine lineare Abbildung mit darstellenden Matrix D . Dann ist $f = T_D$, ausgeschrieben: Für alle $v \in \mathbb{R}^n$ stimmt der Wert $f(v)$ mit dem Ergebnis $D \cdot v$ der Matrix-Vektor-Multiplikation von v mit der darstellenden Matrix D überein.

Informell kann man den Satz so zusammenfassen: Klarerweise legt eine lineare Abbildung ihre darstellende Matrix eindeutig fest; aber auch umgekehrt legt die darstellende Matrix auch die lineare Abbildung fest. Ferner erhält man, wenn man zu einer Matrix A die zugehörige lineare Abbildung T_A und davon wiederum die darstellende Matrix bildet, wieder A . Startet man mit einer linearen Abbildung von \mathbb{R}^n nach \mathbb{R}^k , bildet die darstellende Matrix dieser Abbildung und dann die zugehörige lineare Abbildung bildet, so erhält man wieder die ursprüngliche lineare Abbildung. (Die letzten zwei Punkte sind nicht selbstverständlich, obwohl in diesem Fall ziemlich klar. Betrachten wir alle Studierende der Uni Bremen und die Gesamtheit ihrer

persönlicher E-Mail-Adressen. Jeder E-Mail-Adresse kann ihr(e) eindeutig(e) BesitzerIn zugeordnet werden. Umgekehrt kann jedem Studierenden eine eindeutige Uni-E-Mail-Adresse zugeordnet werden. Startet man jedoch mit einer beliebigen E-Mail-Adresse, ermittelt ihren Besitzer und ordnet dem dann die Uni-E-Mail-Adresse zu, so ist es nicht unbedingt dieselbe E-Mail-Adresse, mit der wir angefangen haben. Bei Matrizen und linearen Abbildungen wie oben kann das allerdings nicht passieren.)

Wir geben nun einige Beispiele für Matrix-Vektor-Multiplikation.

Beispiel. Sei B die 4×3 -Matrix

$$B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 4 & -3 & \frac{1}{2} \\ -\frac{1}{2} & 5 & \frac{3}{2} \end{pmatrix},$$

und sei $T_B: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ die zugehörige lineare Abbildung, die durch $T_B(v) = B \cdot v$ für alle $v \in \mathbb{R}^3$ definiert ist.

Wir berechnen einige Beispiele für Matrix-Vektor-Multiplikation, indem wir für v einige Beispielvektoren aus \mathbb{R}^3 einsetzen:

$$\begin{aligned} T_B \left(\begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \right) &= B \cdot \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 4 & -3 & \frac{1}{2} \\ -\frac{1}{2} & 5 & \frac{3}{2} \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot (-1) + 2 \cdot 0 + 3 \cdot 2 \\ 0 \cdot (-1) + (-1) \cdot 0 + (-2) \cdot 2 \\ 4 \cdot (-1) + (-3) \cdot 0 + \frac{1}{2} \cdot 2 \\ -\frac{1}{2} \cdot (-1) + 5 \cdot 0 + \frac{3}{2} \cdot 2 \end{pmatrix} \\ &= \begin{pmatrix} 5 \\ -4 \\ -3 \\ \frac{7}{2} \end{pmatrix}. \end{aligned}$$

Hierbei haben wir nochmal die Einträge der Matrix spaltenweise farblich hervorgehoben, um besser verfolgen zu können, wie die Rechnung funktioniert. Man beachte, dass in jeder Zeile die jeweiligen Produkte addiert werden, sodass am Schluss wirklich ein Spaltenvektor, mit jeweils nur einem Eintrag in jeder Zeile, herauskommt; an diesem Ergebnis sieht man nicht mehr, wo die ursprünglichen Einträge der Matrix oder des eingesetzten Vektors eingegangen sind.

Nun setzen wir als weitere Übung einen besonders einfachen Vektor ein:

$$T_B \left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right) = B \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 4 & -3 & \frac{1}{2} \\ -\frac{1}{2} & 5 & \frac{3}{2} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{aligned}
&= \begin{pmatrix} 1 \cdot 0 + 2 \cdot 1 + 3 \cdot 0 \\ 0 \cdot 0 + (-1) \cdot 1 + (-2) \cdot 0 \\ 4 \cdot 0 + (-3) \cdot 1 + \frac{1}{2} \cdot 0 \\ -\frac{1}{2} \cdot 0 + 5 \cdot 1 + \frac{3}{2} \cdot 0 \end{pmatrix} \\
&= \begin{pmatrix} 2 \\ -1 \\ -3 \\ 5 \end{pmatrix}.
\end{aligned}$$

In diesem Fall erhalten wir genau die zweite Spalte der Matrix B als Produkt. Das stimmt mit der Aussage des obigen Satzes zusammen, dass die darstellende Matrix von T_B gerade B ist, und dementsprechend der zweite Standardbasis von \mathbb{R}^3 von T_B gerade auf den Vektor abgebildet wird, der durch die zweite Spalte von B bestimmt wird.

Setzt man den Nullvektor ein, so erhält man

$$\begin{aligned}
T_B \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right) &= B \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 4 & -3 & \frac{1}{2} \\ -\frac{1}{2} & 5 & \frac{3}{2} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 \cdot 0 + 2 \cdot 0 + 3 \cdot 0 \\ 0 \cdot 0 + (-1) \cdot 0 + (-2) \cdot 0 \\ 4 \cdot 0 + (-3) \cdot 0 + \frac{1}{2} \cdot 0 \\ -\frac{1}{2} \cdot 0 + 5 \cdot 0 + \frac{3}{2} \cdot 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.
\end{aligned}$$

Allgemeiner kann man sich auf eine ähnliche Art und Weise vergewissern, dass für jede $k \times n$ -Matrix A das Produkt von A mit dem n -komponentigem Nullvektor genau den k -komponentigen Nullvektor liefert wird, denn alle vorkommenden Produkte in der Matrix-Vektor-Multiplikation 0 sein werden und somit auch Summen, also alle Komponenten des Ergebnisses. Da jede lineare Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^k$ durch Matrix-Vektor-Multiplikation gegeben ist, bildet jede solche lineare Abbildung den Nullvektor auf den Nullvektor ab. Tatsächlich ist das auch für alle lineare Abbildungen wahr, mit beliebigen Unterräumen als Quelle und Ziel.

Schließlich wollen wir uns nochmal die allgemeine Formel für T_B vergegenwärtigen:

$$T_B \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) = B \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 4 & -3 & \frac{1}{2} \\ -\frac{1}{2} & 5 & \frac{3}{2} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$= \begin{pmatrix} 1 \cdot x + 2 \cdot y + 3 \cdot z \\ 0 \cdot x + (-1) \cdot y + (-2) \cdot z \\ 4 \cdot x + (-3) \cdot y + \frac{1}{2} \cdot z \\ -\frac{1}{2} \cdot x + 5 \cdot y + \frac{3}{2} \cdot z \end{pmatrix}$$

Wir wollen nochmal betonen, dass darstellende Matrizen nur für lineare Abbildungen einen Sinn ergeben.

Bemerkung. Der Ausdruck „darstellende Matrix“ ergibt nur für lineare Abbildungen Sinn und wird deswegen auch nur für solche definiert. Zwar kann man auch für andere Abbildungen von \mathbb{R}^n nach \mathbb{R}^k die Werte der Abbildung auf den Standardbasisvektoren in eine Matrix schreiben, jedoch wird diese Matrix im Falle der nicht-linearen Abbildungen *nicht* die gesamte Abbildung beschreiben; daher haben nicht-lineare Abbildungen keine darstellenden Matrizen.

Das wollen wir nochmal an einem Beispiel veranschaulichen:

Beispiel. Wir wissen bereits, dass die Abbildung $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die durch $g\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x^2 \\ y \end{pmatrix}$ gegeben ist, nicht linear ist. Wir bestimmen die Werte der Abbildung g auf den Standardbasisvektoren in \mathbb{R}^2 :

$$\begin{aligned} g\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) &= \begin{pmatrix} 1^2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ g\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) &= \begin{pmatrix} 0^2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

Die darstellende Matrix *wäre*, wenn sie denn definiert wäre, die Matrix

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

Die zu der Matrix A assoziierte lineare Abbildung ist gegeben durch

$$\begin{aligned} T_A\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot x + 0 \cdot y \\ 0 \cdot x + 1 \cdot y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}. \end{aligned}$$

Insbesondere stellen wir fest, dass g und T_A unterschiedliche Abbildung von \mathbb{R}^2 nach \mathbb{R}^2 definieren; setzt man etwa $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ ein, so erhält man

$$g\left(\begin{pmatrix} 2 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2^2 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix},$$

$$T_A \left(\begin{pmatrix} 2 \\ 0 \end{pmatrix} \right) = \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$

also unterschiedliche Ergebnisse. Die Abbildung g ist also nicht dieselbe wie T_A und insbesondere wird g nicht vollständig durch A festgelegt.

Bevor wir weitere Beispiele für die lineare Abbildungen von \mathbb{R}^n nach \mathbb{R}^k sehen, wollen wir noch kurz auf den allgemeineren Fall eingehen, dass sowohl die Quelle als auch das Ziel der linearen Abbildung ein beliebiger Unterraum ist. Wir erinnern uns, dass eine lineare Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^k$ durch ihre Werte auf der Standardbasis e_1, e_2, \dots, e_n von \mathbb{R}^n schon vollständig bestimmt ist. Für die Rechnungen ist die Tatsache, dass es sich hierbei gerade um die Standardbasis handelt, besonders praktisch; für die Aussage an sich ist jedoch einzig wichtig, dass es eine Basis ist. Tatsächlich ist es auch bei Unterräumen von \mathbb{R}^n der Fall, wie wir im folgenden Satz sehen werden.

Satz 4.5. *Seien $U \neq \{0\} \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^k$ Unterräume von \mathbb{R}^n bzw. \mathbb{R}^k . Sei ferner u_1, u_2, \dots, u_l eine Basis von U . Eine lineare Abbildung $\varphi: U \rightarrow V$ ist durch ihre Werte $\varphi(u_1), \varphi(u_2), \dots, \varphi(u_l)$ vollständig festgelegt. Umgekehrt legt auch jede beliebige Auswahl von Werten $\varphi(u_1), \varphi(u_2), \dots, \varphi(u_l)$ in V eine lineare Abbildung von U nach V fest.*

In der Quelle wählen wir uns eine Basis, auf der wir die Werte der linearen Abbildung vorgeben. In \mathbb{R}^n haben wir bislang dafür immer die Standardbasis gewählt, aber es mag unter Umständen auch sinnvoll sein, eine andere Basis zu wählen. Dass man gerade eine Basis von der Quelle U braucht, kann man sich ungefähr wie folgt veranschaulichen: Wenn man die Werte der Abbildung nicht auf einem Erzeugendensystem von U festgelegt hat, sondern auf einer Ansammlung von Vektoren, deren lineare Hülle strikt in U enthalten ist, so gibt es Vektoren außerhalb der linearen Hülle, und für die wüssten wir nicht, wie die lineare Abbildung aussehen muss. Man muss dementsprechend die lineare Abbildung auf einem Erzeugendensystem von U festgelegt haben, um sie für jeden Vektor in U bestimmen zu können. Ist allerdings das Erzeugendensystem nicht linear unabhängig, so gibt es unterschiedliche Arten, ein und denselben Vektor in Termen von diesem Erzeugendensystem darzustellen, die unter Umständen (allerdings nicht notwendigerweise) zu unterschiedlichen Werten der linearen Abbildung auf ein und demselben Vektor liefern. Dann wäre dies allerdings keine Abbildung, denn jedem Vektor in U muss *genau* ein Element von V zugeordnet werden. Man beachte auch, dass hingegen für die Werte $\varphi(u_1), \varphi(u_2), \dots, \varphi(u_l)$ der linearen Abbildung überhaupt keine Bedingungen vorliegen. Diese Werte können linear abhängig oder linear unabhängig sein, sie können ein Erzeugendensystem von V bilden oder eben nicht; das spielt für die eindeutige Festlegung der Abbildung φ keine Rolle.

Nun wollen wir einige weitere Beispiele für lineare Abbildungen angeben.

Beispiel. • Wir betrachten die $k \times n$ -Matrix N , in der alle Einträge Null sind. Diese definiert eine lineare Abbildung $T_N: \mathbb{R}^n \rightarrow \mathbb{R}^k$, die jedem Vektor $v \in \mathbb{R}^n$ den Nullvektor in \mathbb{R}^k zuordnet. Diese lineare Abbildung wird die *Nullabbildung* genannt. Sie ist zwar nicht besonders spannend, aber manchmal ganz praktisch. Wir sprechen übrigens für alle k und n von einer Nullabbildung; die Verwechslungsgefahr ist hierbei gering.

- In einem der eben behandelten Beispiele haben wir bereits eine weitere spezielle Art von Matrizen gesehen. Wir bezeichnen die quadratische $n \times n$ -Matrix, die auf der Diagonale lauter Einsen und sonst nur Nullen als Einträge hat, als die $n \times n$ -Einheitsmatrix E_n . Es ist also beispielsweise

$$E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Wir betrachten die zu E_n assoziierte lineare Abbildung $\mathbb{R}^n \rightarrow \mathbb{R}^n$. Multipliziert man einen Vektor in \mathbb{R}^n mit E_n , so stellt man fest, dass wir wieder denselben Vektor als Ergebnis erhalten. Die Abbildung, die jedem Vektor einfach diesen Vektor selbst zuordnet, haben wir schon in Mathematischen Grundlagen 1 im allgemeineren Kontext kennengelernt. Ist $X \neq \emptyset$ eine beliebige Menge, so haben wir die *Identitätsabbildung*

$$\text{id}_X: X \rightarrow X,$$

die durch die Vorschrift $\text{id}_X(x) = x$ für alle $x \in X$ gegeben ist. Die Einheitsmatrix E_n ist also die darstellende Matrix der Identitätsabbildung $\text{id}_{\mathbb{R}^n}: \mathbb{R}^n \rightarrow \mathbb{R}^n$. Auch diese Abbildung ist nicht besonders spannend, aber ebenfalls ganz praktisch.

5 Matrizenmultiplikation

Als nächstes wollen wir uns mit der Frage beschäftigen, wie die Hintereinanderausführung zweier linearen Abbildungen aussieht. Abbildungen im Allgemeinen und lineare Abbildungen im Besonderen können als Operationen auf den vorgegebenen Daten oder Transformationen von diesen gesehen werden. Führt man diese Operationen hintereinander aus, so wird das mathematisch als Komposition von Abbildungen kodiert. Das ist etwa sinnvoll, wenn man mehrere Kameraschwenks hintereinander durchführen will, oder wenn man beispielsweise zunächst ein Objekt drehen und dann in die Ebene projizieren; taucht aber auch in vielen weiteren Kontexten auf.

Wir erinnern uns an die Definition der Komposition aus Mathematischen Grundlagen 1, die zunächst nichts mit der Linearität zu tun hat und für beliebige Abbildungen definiert ist.

Definition 5.1. Seien X, Y, Z beliebige Mengen und $f: X \rightarrow Y$, $g: Y \rightarrow Z$ beliebige Abbildungen. Die **Komposition** $g \circ f$ (in Worten „ g nach f “ oder „ g verknüpft mit f “) der Abbildungen g und f ist eine Abbildung, die durch die Vorschrift

$$\begin{aligned} g \circ f: X &\rightarrow Z, \\ x &\mapsto g(f(x)) \end{aligned}$$

definiert ist.

Wir erinnern uns nochmal daran, dass $g \circ f$ nur definiert ist, wenn das Ziel von f und die Quelle von g übereinstimmen, sonst kann man nämlich das Ergebnis $f(x)$ der Anwendung von f auf ein Element $x \in X$ gar nicht in die Abbildung g einsetzen. Die Komposition $g \circ f$ hat die Quelle von f als Quelle und das Ziel von g als Ziel.

Es gibt viele synonyme Namen für Komposition von Abbildungen. Unter dem Namen „Verkettung“ wird es zum Teil bereits in der Schule behandelt und taucht zum Beispiel in der „Kettenregel“ für Ableitungen auf.

Nun wollen wir spezieller auf die Komposition linearer Abbildungen eingehen. Die erste wichtige Erkenntnis besteht darin, dass die Komposition zweier linearer Abbildungen wieder linear ist. Diesen Satz wollen wir beweisen.

Satz 5.2. Seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^k$ und $W \subseteq \mathbb{R}^l$ Unterräume von \mathbb{R}^n , \mathbb{R}^k bzw. \mathbb{R}^l . Seien ferner $\varphi: U \rightarrow V$ und $\psi: V \rightarrow W$ lineare Abbildungen. Dann ist die Komposition $\psi \circ \varphi: U \rightarrow W$ ebenfalls eine lineare Abbildung.

Beweis. Wir müssen nachprüfen, dass für die Abbildung $\psi \circ \varphi: U \rightarrow W$ die Bedingungen (A1) und (A2) gelten.

zu (A1): Seien also $u_1, u_2 \in U$ zwei beliebige Vektoren der Quelle. Wir müssen prüfen, ob $\psi \circ \varphi(u_1 + u_2)$ und $\psi \circ \varphi(u_1) + \psi \circ \varphi(u_2)$ dasselbe sind. Zunächst benutzen wir die Definition der Komposition und erhalten

$$\psi \circ \varphi(u_1 + u_2) = \psi(\varphi(u_1 + u_2)).$$

Nun erinnern wir uns daran, dass φ nach Voraussetzung eine lineare Abbildung ist und die (A1) für φ nun besagt:

$$\varphi(u_1 + u_2) = \varphi(u_1) + \varphi(u_2).$$

Das ist nun ein Element des Unterraums V , das auf zwei verschiedene Arten und Weisen geschrieben wurde. Setzt man aber dasselbe Element von V in ψ ein, egal wie es geschrieben wurde, muss dasselbe rauskommen. Wir erhalten also

$$\psi(\varphi(u_1 + u_2)) = \psi(\varphi(u_1) + \varphi(u_2)).$$

Jetzt nutzen wir, dass auch ψ linear ist und somit ebenfalls der Bedingung (A1) genügt. Diese wollen wir nun auf die Vektoren $\varphi(u_1), \varphi(u_2) \in V$ anwenden. Das liefert nun

$$\psi(\varphi(u_1) + \varphi(u_2)) = \psi(\varphi(u_1)) + \psi(\varphi(u_2)).$$

Führt man sich erneut die Definition von Komposition vor Augen, so führt die obige Gleichungskette zu dem Schluss, dass

$$\psi \circ \varphi(u_1 + u_2) = \psi \circ \varphi(u_1) + \psi \circ \varphi(u_2)$$

gilt. Somit erfüllt die Komposition zweier linearer Abbildungen die Bedingung (A1). Wir bemerken außerdem, dass wir nur benutzt haben, dass sowohl φ als auch ψ die Bedingung (A1) erfüllen; die Bedingung (A2) war hierbei irrelevant.

zu (A2): Nun wollen wir also zeigen, dass die Abbildung $\psi \circ \varphi$ sich „gut“ bezüglich der Skalarmultiplikation verhält. Der Beweis funktioniert recht ähnlich zum ersten Teil. Seien also $u \in U$ und $\lambda \in \mathbb{R}$ beliebig. Nach Definition der Komposition gilt

$$\psi \circ \varphi(\lambda u) = \psi(\varphi(\lambda u)).$$

Da φ linear ist und somit die Bedingung (A2) erfüllt, folgt

$$\varphi(\lambda u) = \lambda \varphi(u).$$

Wir wenden wieder ψ auf beide Seiten der Gleichheit an und erhalten

$$\psi(\varphi(\lambda u)) = \psi(\lambda \varphi(u)).$$

Da ψ ebenfalls lineare Abbildung ist und somit auch die Bedingung (A2) erfüllt, können wir den Faktor λ aus dem letzten Ausdruck „herausziehen“ (als Vorfaktor des Vektors $\varphi(u) \in V$) und erhalten

$$\psi(\lambda \varphi(u)) = \lambda \cdot \psi(\varphi(u)).$$

Erinnert man sich nun noch an die Definition der Komposition, so erhält man insgesamt

$$\psi \circ \varphi(\lambda u) = \lambda \cdot \psi \circ \varphi(u).$$

Somit ist die Bedingung (A2) für $\psi \circ \varphi$ erfüllt. Wir bemerken nochmal, dass hier nur benutzt wurde, dass φ und ψ der Bedingung (A2) genügen.

Insgesamt haben wir also gezeigt, dass die Komposition zweier linearer Abbildungen wieder eine lineare Abbildung ist. \square

Nun beschäftigen wir uns mit dem Spezialfall, dass wir zwei lineare Abbildungen $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^k$ und $\psi: \mathbb{R}^k \rightarrow \mathbb{R}^l$ haben. Wir wissen, dass jede dieser Abbildungen jeweils durch eine Matrix dargestellt sind. Bezeichnen wir die darstellende Matrix von φ als A und die darstellende Matrix von ψ als B . Es sei angemerkt, dass A eine $k \times n$ - und B eine $l \times k$ -Matrix ist. Wir suchen nun die darstellende Matrix der linearen Abbildung $\psi \circ \varphi: \mathbb{R}^n \rightarrow \mathbb{R}^l$. Diese wollen wir in Termen von Matrizen A und B ausdrücken. Bevor wir dafür ein allgemeines Verfahren entwickeln, betrachten wir ein Beispiel.

Beispiel. Seien

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

und

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

die darstellenden Matrizen für $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ bzw. $\psi: \mathbb{R}^3 \rightarrow \mathbb{R}^4$. Wir suchen die darstellende 2×4 -Matrix der Abbildung $\psi \circ \varphi: \mathbb{R}^2 \rightarrow \mathbb{R}^4$. Nach Definition stehen in den Spalten dieser Matrix gerade die Bilder der Standardbasisvektoren der Quelle, also die Vektoren $\psi \circ \varphi \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)$, $\psi \circ \varphi \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$. Diese wollen wir nun berechnen. Dabei ist es einfach, den ersten Teil des Einsetzens durchzuführen: Die Vektoren $\varphi \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)$, $\varphi \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$ sind nach Definition gerade die Spalten der Matrix A . Wir müssen also in ψ die Spalten der darstellenden Matrix von φ einsetzen, um die Spalten der darstellenden Matrix von $\psi \circ \varphi$ zu finden. Wir wissen, dass wir dafür wiederum die Spalten von A jeweils mit der Matrix B multiplizieren müssen. In diesem Fall erhalten wir:

$$\begin{aligned} \psi \circ \varphi \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) &= \psi \left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right) \\ &= B \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 4 & -3 & \frac{1}{2} \\ -\frac{1}{2} & 5 & \frac{3}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} 1 \cdot 1 + 2 \cdot (-1) + 3 \cdot 0 \\ 0 \cdot 1 + (-1) \cdot (-1) + (-2) \cdot 0 \\ 4 \cdot 1 + (-3) \cdot (-1) + \frac{1}{2} \cdot 0 \\ -\frac{1}{2} \cdot 1 + 5 \cdot (-1) + \frac{3}{2} \cdot 0 \end{pmatrix} \\
&= \begin{pmatrix} -1 \\ 1 \\ 7 \\ -\frac{11}{2} \end{pmatrix}.
\end{aligned}$$

Mit der zweiten Spalte verhält es sich ganz ähnlich:

$$\begin{aligned}
\psi \circ \varphi \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) &= \psi \left(\begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix} \right) \\
&= B \cdot \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix} \\
&= \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 4 & -3 & \frac{1}{2} \\ -\frac{1}{2} & 5 & \frac{3}{2} \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix} \\
&= \begin{pmatrix} 1 \cdot 4 + 2 \cdot 2 + 3 \cdot 3 \\ 0 \cdot 4 + (-1) \cdot 2 + (-2) \cdot 3 \\ 4 \cdot 4 + (-3) \cdot 2 + \frac{1}{2} \cdot 3 \\ -\frac{1}{2} \cdot 4 + 5 \cdot 2 + \frac{3}{2} \cdot 3 \end{pmatrix} \\
&= \begin{pmatrix} 17 \\ -8 \\ \frac{23}{2} \\ \frac{25}{2} \end{pmatrix}.
\end{aligned}$$

Insgesamt sehen wir also, dass die darstellende Matrix von $\psi \circ \varphi$ gerade die Matrix

$$C = \begin{pmatrix} -1 & 17 \\ 1 & -8 \\ 7 & \frac{23}{2} \\ -\frac{11}{2} & \frac{25}{2} \end{pmatrix}.$$

Im großen und ganzen haben wir in dem Beispiel bereits gesehen, wie das allgemeine Verfahren zur Bestimmung der darstellenden Matrix der Komposition zweier Abbildungen aussehen muss: Wir multiplizieren die Matrix der als zweites angewandten Abbildung mit den Spalten der Matrix der als erstes angewandten Abbildung, und schreiben die Ergebnisse in die Spalten unserer neuen Matrix. Das ist im Wesentlichen die Definition der Matrizenmultiplikation, einer neuen Operation für Matrizen, die wir jetzt einführen wollen.

Definition 5.3. Sei A eine $k \times n$ -Matrix und B eine $n \times r$ -Matrix. Das Produkt $A \cdot B$ der Matrizen A und B ist definiert als die Matrix, die als Spalten die Matrix-Vektor-Produkte der Matrix A mit den Spalten der Matrix B hat.

Wir bemerken nochmal explizit, dass die Matrizengrößen „zusammenpassen“ müssen: Die Anzahl der Spalten der ersten Matrix muss mit der Anzahl der Zeilen der zweiten Matrix übereinstimmen. Ist das nicht der Fall, so sagt man, das Produkt $A \cdot B$ sei nicht definiert. Ist das Produkt definiert, so hat das Ergebnis so viele Zeilen wie die erste Matrix und so viele Spalten wie die zweite.

Wir sehen nochmal, dass das Produkt von Matrizen gerade die darstellende Matrix der Komposition der zugehörigen Abbildungen liefert, $A \cdot B$ ist gerade die darstellende Matrix für $T_A \circ T_B$. Die Bedingung an die Matrizengrößen entspricht gerade der Tatsache, dass Quelle von T_A und Ziel von T_B übereinstimmen müssen, damit $T_A \circ T_B$ definiert ist.

In einiger Hinsicht ist das Produkt für Matrizen ähnlich dem Produkt von reellen Zahlen. Allerdings besteht bei Matrizen ein fundamentaler Unterschied: Für Matrizenmultiplikation gilt nicht das Kommutativitätsgesetz. Für die Multiplikation reeller Zahlen wissen wir, dass wir die Faktoren nach Belieben vertauschen können; es gilt etwa $5 \cdot 8 = 8 \cdot 5$, oder allgemeiner $x \cdot y = y \cdot x$ für alle $x, y \in \mathbb{R}$. Das funktioniert bei Matrizen im Allgemeinen nicht. Das ist nicht überraschend, wenn man bedenkt, dass es auch bei Komposition von Abbildungen sehr auf die Reihenfolge, in der diese ausgeführt werden, ankommt, wie wir bereits auch in Mathematischen Grundlagen 1 gesehen haben.

Wir verdeutlichen die Behauptung, dass die Matrizenmultiplikation im Allgemeinen nicht kommutativ ist, an einigen Beispielen.

Beispiel. • Zunächst bemerken wir, dass wir bereits eine Matrizenmultiplikation ausgeführt haben. Wir haben berechnet, dass

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 4 & -3 & \frac{1}{2} \\ -\frac{1}{2} & 5 & \frac{3}{2} \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 \\ -1 & 2 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 17 \\ 1 & -8 \\ 7 & \frac{23}{2} \\ -\frac{11}{2} & \frac{25}{2} \end{pmatrix}$$

gilt. Will man aber das Produkt derselben Matrizen in der umgekehrten Reihenfolge ausrechnen, also

$$\begin{pmatrix} 1 & 4 \\ -1 & 2 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 4 & -3 & \frac{1}{2} \\ -\frac{1}{2} & 5 & \frac{3}{2} \end{pmatrix},$$

so stellt man fest, dass dieses Produkt nicht definiert ist: Die erste Matrix hat 2 Spalten, während die zweite 4 Zeilen hat.

Es kann also passieren, dass das Produkt $B \cdot A$ zweier Matrizen definiert ist, während $A \cdot B$ nicht definiert ist und dementsprechend nicht dasselbe ist.

- Wir betrachten ein weiteres Beispiel. Sei C die 2×3 -Matrix

$$C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

und A wieder die 3×2 -Matrix

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

In diesem Fall sehen wir, dass sowohl $C \cdot A$ als auch $A \cdot C$ definiert sind. Allerdings ist $C \cdot A$ eine 2×2 -Matrix, während $A \cdot C$ eine 3×3 -Matrix ist. Die beiden Produkte haben also keine Chance, gleich zu sein. Vollständigkeitshalber rechnen wir beide Produkte aus:

$$\begin{aligned} C \cdot A &= \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 \\ -1 & 2 \\ 0 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 1 + 2 \cdot (-1) + 3 \cdot 0 & 1 \cdot 4 + 2 \cdot 2 + 3 \cdot 3 \\ 4 \cdot 1 + 5 \cdot (-1) + 6 \cdot 0 & 4 \cdot 4 + 5 \cdot 2 + 6 \cdot 3 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 17 \\ -1 & 44 \end{pmatrix}; \\ A \cdot C &= \begin{pmatrix} 1 & 4 \\ -1 & 2 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 1 + 4 \cdot 4 & 1 \cdot 2 + 4 \cdot 5 & 1 \cdot 3 + 4 \cdot 6 \\ (-1) \cdot 1 + 2 \cdot 4 & (-1) \cdot 2 + 2 \cdot 5 & (-1) \cdot 3 + 2 \cdot 6 \\ 0 \cdot 1 + 3 \cdot 4 & 0 \cdot 2 + 3 \cdot 5 & 0 \cdot 3 + 3 \cdot 6 \end{pmatrix} \\ &= \begin{pmatrix} 17 & 22 & 27 \\ 7 & 8 & 9 \\ 12 & 15 & 18 \end{pmatrix} \end{aligned}$$

- Nun betrachten wir zwei 2×2 -Matrizen. Das Produkt zweier 2×2 -Matrizen ist stets definiert und liefert wieder eine 2×2 -Matrix (analog gilt das für beliebige quadratische, also $n \times n$ -Matrizen). Obwohl beide Produkte nun dieselbe Größe haben, müssen diese trotzdem nicht übereinstimmen, wie wir nun zeigen wollen.

Wir betrachten also zunächst

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0 + 2 \cdot 0 & 1 \cdot 1 + 2 \cdot 0 \\ 3 \cdot 0 + 4 \cdot 0 & 3 \cdot 1 + 4 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 3 \end{pmatrix}.$$

Vertauscht man nun die Faktoren, so erhält man

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 \cdot 1 + 1 \cdot 3 & 0 \cdot 2 + 1 \cdot 4 \\ 0 \cdot 1 + 0 \cdot 3 & 0 \cdot 2 + 0 \cdot 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 0 & 0 \end{pmatrix}.$$

Wie bereits angekündigt, stimmen die Ergebnisse trotz gleicher Größe nicht überein. Das zeigt nochmal, dass die Matrizenmultiplikation selbst für quadratische Matrizen nicht kommutativ ist.

6 Invertieren von Matrizen

Als nächstes wollen wir uns mit der Frage beschäftigen, ob der Effekt einer linearen Abbildung durch anschließende Anwendung einer weiteren (linearen) Abbildung aufgehoben werden kann. Manchmal wird das möglich sein; betrachtet man etwa eine Drehung in der Ebene oder einen „Kameraschwenk“ im Raum, so kann man jeweils einfach die Objekte in der Ebene oder die Kamera „zurückdrehen“, um in den ursprünglichen Zustand zurückzukehren. Hingegen wollen wir jetzt ein Beispiel für eine lineare Abbildung sehen, deren Effekt nicht umkehrbar ist.

Beispiel. Wir betrachten die Projektion des Raums auf die x - y -Ebene, also die Abbildung $p: \mathbb{R}^3 \rightarrow \mathbb{R}^2$, die durch $p \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) = \begin{pmatrix} x \\ y \end{pmatrix}$ definiert ist.

Man kann nachprüfen, dass das eine lineare Abbildung ist. Diese Abbildung lässt sich allerdings nicht rückgängig machen, da wir durch die Projektion Informationen verlieren, die nicht mehr aus dem Bildpunkt ablesen können, nämlich die z -Koordinate, also gewissermaßen die Höhe von dem Punkt. Etwa der Ursprung $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ in der Ebene ist sowohl das Bild des Ursprungs $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ im

Raum als auch des Punkts $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Wir haben keine Möglichkeit, festzustellen, welche z -Koordinate der Punkt vor der Projektion hatte. Diese Mehrdeutigkeit kennt man auch aus den Bildern, die einfach durch die Projektion auf eine Ebene entstehen. Das ist mitunter der Grund, warum ein solches Bild nicht „räumlich“ wirkt.

Manche lineare Abbildungen lassen sich hingegen einfach wieder umkehren, wie wir nochmal am folgenden Beispiel verdeutlichen wollen.

Beispiel. Wir betrachten die Abbildung $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die durch $f \left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} \frac{1}{2}y \\ x \end{pmatrix}$ definiert ist. Diese Abbildung ist, wie man leicht nachprüft, linear.

Außerdem sieht man leicht, dass der Effekt dieser Abbildung durch die Abbildung $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die durch $g\left(\begin{pmatrix} z \\ w \end{pmatrix}\right) = \begin{pmatrix} 2w \\ z \end{pmatrix}$ definiert ist, wieder aufgehoben wird, also dass $(g \circ f)(v) = v$ für alle $v \in \mathbb{R}^2$ gilt. Tatsächlich ist auch umgekehrt $(f \circ g)(v) = v$ für alle $v \in \mathbb{R}^2$.

Wir wollen also dem Unterschied zwischen den beiden letzten Beispielen auf den Grund gehen. Wir erinnern uns, dass wir bereits in Mathematischen Grundlagen 1 ein ähnliches, allgemeineres Problem diskutiert haben, nämlich wann der Effekt einer beliebigen Abbildung rückgängig gemacht werden kann. Wir wiederholen einige wichtige Punkte hiervon.

Definition 6.1. Seien X, Y beliebige Mengen. Eine Abbildung $g: Y \rightarrow X$ heißt **inverse Abbildung** (oder: **Umkehrabbildung**) zu der Abbildung $f: X \rightarrow Y$, falls sowohl $f \circ g = \text{id}_Y$ als auch $g \circ f = \text{id}_X$ gilt, ausgeschrieben: $f(g(y)) = y$ für alle $y \in Y$ und $g(f(x)) = x$ für alle $x \in X$.

Man sagt in diesem Fall auch, f und g seien zueinander invers. Dieser Begriff ist möglicherweise in ähnlicher Form auch schon aus der Schule bekannt, etwa im Kontext der Wurzelfunktion auf den positiven reellen Zahlen.

In Mathematischen Grundlagen 1 haben wir das folgende Kriterium für die Existenz einer Umkehrabbildung kennengelernt:

Satz 6.2. Seien X, Y beliebige Mengen. Eine Abbildung $f: X \rightarrow Y$ hat genau dann eine inverse Abbildung, wenn f bijektiv ist.

Wir erinnern uns, dass eine „genau dann, wenn“-Aussage zwei Implikationen zusammenfasst: Hat f eine inverse Abbildung, so ist f auch bijektiv, und umgekehrt besitzt jede bijektive Abbildung f eine Umkehrabbildung. Der Satz sagt zunächst nichts darüber aus, wie man diese Umkehrabbildung ausrechnen kann, es geht bloß um die Existenz der Umkehrabbildung.

Wir wiederholen nochmal die Begriffe bijektiv, injektiv und surjektiv.

Definition 6.3. Seien X, Y beliebige Mengen.

- Eine Abbildung $f: X \rightarrow Y$ heißt **injektiv**, wenn für alle $x_1, x_2 \in X$ aus $f(x_1) = f(x_2)$ folgt: $x_1 = x_2$.
- Eine Abbildung $f: X \rightarrow Y$ heißt **surjektiv**, wenn jedes Element $y \in Y$ ein **Urbild** $x \in X$ hat, d.h. zu jedem $y \in Y$ gibt es ein $x \in X$ mit $f(x) = y$.
- Eine Abbildung $f: X \rightarrow Y$ heißt **bijektiv**, falls sie sowohl injektiv als auch surjektiv ist.

Wir bemerken nochmal, dass injektiv und surjektiv im Allgemeinen unabhängige Eigenschaften sind. Injektivität wird manchmal mit der Definition

der Abbildung verwechselt: Während bei der Abbildung lediglich zugesichert wird, dass *ein* Element der Quelle genau ein Element im Ziel zugeordnet bekommt, wird bei Injektivität zusätzlich noch verlangt, dass unterschiedlichen Elementen der Quelle auch unterschiedliche Elemente im Ziel zugewiesen werden. Vergleicht man das mit Nummern ziehen, etwa im Bürgerbüro, so entspricht eine Abbildung von der Menge der Wartenden in die Menge z.B. der Zahlen von 1 bis 100 dem Zustand, dass jede Person genau eine Nummer gezogen hat. Würden Nummern mehrfach vergeben, so hieße es, dass diese Abbildung nicht injektiv ist. Dass jede Nummer von 1 bis 100 mindestens einmal vergeben wurde, würde der Surjektivität entsprechen.

Übrigens kann Injektivität so umformuliert werden, dass jedes Element im Ziel *höchstens* ein Urbild besitzt. Kombiniert man Injektivität und Surjektivität, so sieht man, dass bei einer bijektiven Abbildung jedes Element im Ziel *genau* ein Urbild in der Quelle besitzt. Das ist auch die entscheidende Beobachtung für den Beweis des soeben zitierten Satzes.

Wir wollen nochmal am Beispiel der Projektion, für die wir uns in etwa die Unmöglichkeit einer inversen Abbildung verdeutlicht haben, diese Eigenschaften nochmal veranschaulichen.

Beispiel. Wir betrachten wieder die lineare Abbildung $p: \mathbb{R}^3 \rightarrow \mathbb{R}^2$, die durch $p \left(\begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) = \begin{pmatrix} x \\ y \end{pmatrix}$ definiert ist.

Diese Abbildung ist nicht injektiv. Beispielsweise werden die unterschiedlichen Elemente $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$ der Quelle auf dasselbe Element im Ziel, nämlich $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ abgebildet. Damit ist die Abbildung p insbesondere nicht bijektiv, und wir sehen nochmal, dass sie auch keine inverse Abbildung besitzen kann.

Wir wollen trotzdem noch zu Übungszwecken die Surjektivität der Abbildung p nachprüfen. Tatsächlich ist die Abbildung p surjektiv, denn jedes Element $\begin{pmatrix} x \\ y \end{pmatrix}$ des Ziels ein (und tatsächlich sogar unendlich viele, auch wenn es für die Surjektivität nicht relevant ist) Urbild in der Quelle besitzt; beispielsweise gilt:

$$p \left(\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \right) = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Die Abbildung p ist also surjektiv, aber nicht injektiv.

Nun wollen wir die Bedeutung der Injektivität und Surjektivität im speziellen für lineare Abbildungen untersuchen. Dabei stellt es sich heraus, dass

injektive lineare Abbildungen lineare Unabhängigkeit erhalten, während surjektive lineare Abbildungen Erzeugendensysteme auf Erzeugendensysteme abbilden. In den Übungen haben wir bereits gesehen, dass beliebige lineare Abbildungen weder lineare Unabhängigkeit noch Erzeugendensysteme zu erhalten brauchen; die Injektivität bzw. Surjektivität sind also für diesen Satz essentiell. Wir präzisieren die vorherigen Behauptungen im folgenden Satz.

Satz 6.4. *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^k$ Unterräume von \mathbb{R}^n bzw. \mathbb{R}^k . Sei $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^k$ eine lineare Abbildung.*

- 1) *Sei φ zusätzlich injektiv und w_1, \dots, w_l linear unabhängige Vektoren in U . Dann sind die Vektoren $\varphi(w_1), \dots, \varphi(w_l)$ linear unabhängig.*
- 2) *Sei φ zusätzlich surjektiv. Sei ferner u_1, \dots, u_m ein Erzeugendensystem von U . Dann bilden die Vektoren $\varphi(u_1), \dots, \varphi(u_m)$ ein Erzeugendensystem von V .*

Es sei nochmal angemerkt, dass beide Teile des Satzes nicht wirklich miteinander zu tun haben: Im ersten Teil braucht die Abbildung nicht surjektiv zu sein, im zweiten kann sie injektiv oder eben nicht injektiv sein. Nun stellen wir einen Beweis für diesen Satz vor.

Beweis. zu 1): Wir müssen zeigen, dass die Vektoren $\varphi(w_1), \dots, \varphi(w_l)$ in V linear unabhängig sind. Dafür müssen wir nach Definition zeigen, dass jede Linearkombination dieser Vektoren, die 0 ergibt, bereits nur Nullen als Koeffizienten hat. Seien also $\mu_1, \dots, \mu_l \in \mathbb{R}$ reelle Zahlen, sodass

$$\mu_1\varphi(w_1) + \mu_2\varphi(w_2) + \dots + \mu_l\varphi(w_l) = 0$$

gilt. Nun nutzen wir die Linearität von φ , um die linke Seite dieser Gleichung umzuformen. Wir benutzen die Eigenschaften (A1) und (A2) allerdings „rückwärts“, also um die Vorfaktoren bzw. die Summe in die Abbildung φ „reinzuziehen“. Bis jetzt haben wir diese Bedingungen nur andersrum verwendet, also um Summe bzw. Vorfaktoren aus der Abbildung „herauszuziehen“. Nun aber durch die umgekehrte Anwendungen erhalten wir

$$\begin{aligned} & \mu_1\varphi(w_1) + \mu_2\varphi(w_2) + \dots + \mu_l\varphi(w_l) \\ = & \varphi(\mu_1w_1) + \varphi(\mu_2w_2) + \dots + \varphi(\mu_lw_l) \\ = & \varphi(\mu_1w_1 + \mu_2w_2 + \dots + \mu_lw_l). \end{aligned}$$

Wir wissen also, dass der Vektor $\mu_1w_1 + \mu_2w_2 + \dots + \mu_lw_l \in U$ von φ auf 0 abgebildet wird. Wir kennen allerdings schon einen Vektor aus U , der von φ auf 0 abgebildet werden muss: Wir haben bereits diskutiert,

dass eine lineare Abbildung immer den Nullvektor auf den Nullvektor abbildet. Also gilt

$$\varphi(\mu_1 w_1 + \mu_2 w_2 + \dots + \mu_l w_l) = \varphi(0).$$

Nun ist φ eine injektive Abbildung, ordnet also unterschiedlichen Elementen der Quelle unterschiedliche Werte zu. Da den Vektoren $\mu_1 w_1 + \mu_2 w_2 + \dots + \mu_l w_l$ und 0 derselbe Wert zugeordnet wird, können diese also nicht unterschiedlich sein, sondern sind gleich, also

$$\mu_1 w_1 + \mu_2 w_2 + \dots + \mu_l w_l = 0.$$

Nun nutzen wir zuletzt die Voraussetzung, dass die Vektoren w_1, \dots, w_l linear unabhängig sind. Wir wissen also nach Definition, dass alle Koeffizienten dieser Linearkombination der Vektoren w_1, \dots, w_l , die Null ergibt, auch 0 sein müssen, also

$$\mu_1 = \mu_2 = \dots = \mu_l = 0.$$

Zusammengefasst haben wir also gezeigt, dass aus der Annahme

$$\mu_1 \varphi(w_1) + \mu_2 \varphi(w_2) + \dots + \mu_l \varphi(w_l) = 0$$

folgt, dass $\mu_1 = \mu_2 = \dots = \mu_l = 0$ gilt. Das zeigt gerade, dass die Vektoren $\varphi(w_1), \dots, \varphi(w_l)$ linear unabhängig sind, und schließt damit den Beweis des ersten Teils ab.

zu 2): Sei nun φ eine surjektive lineare Abbildung (nicht notwendigerweise injektiv) und seien $u_1, \dots, u_m \in U$ ein Erzeugendensystem von U . Wir müssen zeigen, dass die Vektoren $\varphi(u_1), \dots, \varphi(u_m) \in V$ ein Erzeugendensystem von V bilden. Das heißt gerade, dass die lineare Hülle dieser Vektoren ganz V ist, also $\text{Span}(\varphi(u_1), \dots, \varphi(u_m)) = V$. Dies sind zwei Mengen, und die Gleichheit zweier Mengen kann gezeigt werden, indem man von jeder der beiden Mengen zeigt, dass sie eine Teilmenge der jeweils anderen Menge ist. Nun haben wir gleich zu Anfang unserer Untersuchung der linearen Hüllen beobachtet, dass die lineare Hülle einer beliebigen Ansammlung von Vektoren in einem Unterraum ebenfalls in diesem Unterraum enthalten ist. In diesem Fall folgt also aus $\varphi(u_1), \dots, \varphi(u_m) \in V$, dass

$$\text{Span}(\varphi(u_1), \dots, \varphi(u_m)) \subseteq V$$

gilt. Wir müssen also lediglich

$$V \subseteq \text{Span}(\varphi(u_1), \dots, \varphi(u_m))$$

beweisen. Dafür müssen wir zeigen, dass jedes Element von V auch in $\text{Span}(\varphi(u_1), \dots, \varphi(u_m))$ enthalten ist, also dass jedes Element von

V sich als Linearkombination der Vektoren $\varphi(u_1), \dots, \varphi(u_m)$ darstellen lässt.

Sei also $v \in V$ beliebig. Wir versuchen, diesen Vektor als Linearkombination von $\varphi(u_1), \dots, \varphi(u_m)$ auszudrücken. Da wir bereits ein Erzeugendensystem von U kennen, wollen wir das ausnutzen. Dafür beobachten wir, dass $\varphi: U \rightarrow V$ eine surjektive Abbildung ist, also $v \in V$ ein Urbild $u \in U$ besitzt, d.h. es gibt einen Vektor $u \in U$ mit $\varphi(u) = v$.

Da nun u_1, \dots, u_m nach Voraussetzung ein Erzeugendensystem von U bilden, lässt sich jedes Element von U , insbesondere auch u , als Linearkombination der Vektoren u_1, \dots, u_m darstellen. Es gibt also reelle Zahlen $c_1, c_2, \dots, c_m \in \mathbb{R}$, sodass

$$u = c_1 u_1 + c_2 u_2 + \dots + c_m u_m$$

gilt. Nun wenden wir auf diesen Ausdruck φ an. Einerseits wissen wir, dass $\varphi(u) = v$ ist, da wir u gerade so ausgewählt haben. Andererseits können wir für u die soeben gefundene Linearkombination in Termen von u_1, \dots, u_m einsetzen. Darauf können wir dann die Bedingungen (A1) und (A2) anwenden, und wir erhalten

$$\begin{aligned} v &= \varphi(u) = \varphi(c_1 u_1 + c_2 u_2 + \dots + c_m u_m) \\ &= c_1 \varphi(u_1) + c_2 \varphi(u_2) + \dots + c_m \varphi(u_m). \end{aligned}$$

Wir haben nun v als Linearkombination der Vektoren $\varphi(u_1), \dots, \varphi(u_m)$ mit Koeffizienten c_1, c_2, \dots, c_m dargestellt. Also haben wir gezeigt, dass jedes Element von V sich als Linearkombination der Vektoren $\varphi(u_1), \dots, \varphi(u_m)$ darstellen lässt und somit insgesamt

$$V = \text{Span}(\varphi(u_1), \dots, \varphi(u_m))$$

gezeigt. Folglich sind $\varphi(u_1), \dots, \varphi(u_m)$ ein Erzeugendensystem von V , was zu beweisen war. □

Nun wollen wir die Erkenntnisse des vorherigen Satzes nutzen, um das folgende Korollar zu beweisen.

Korollar 6.5. *Seien $U \subseteq \mathbb{R}^n$ und $V \subseteq \mathbb{R}^k$ Unterräume von \mathbb{R}^n bzw. \mathbb{R}^k und sei $U \neq \{0\}$. Sei ferner $\varphi: U \rightarrow V$ eine bijektive lineare Abbildung. Dann gilt: $\dim U = \dim V$.*

Beweis. Wir haben bereits besprochen, dass jeder Unterraum $\neq \{0\}$ von \mathbb{R}^n , insbesondere also U , eine Basis besitzt. Sei u_1, \dots, u_l eine Basis von U . Nach Definition ist die Dimension von U gerade die Anzahl der Elemente in einer beliebigen Basis von U ; es gilt also $l = \dim U$. Nun ist eine Basis von U nach

Definition ein linear unabhängiges Erzeugendensystem von U . Da φ eine surjektive lineare Abbildung ist, bildet sie das Erzeugendensystem u_1, \dots, u_l von U auf ein Erzeugendensystem $\varphi(u_1), \dots, \varphi(u_l)$ von V . Da φ außerdem eine injektive lineare Abbildung ist, bildet sie die linear unabhängigen Vektoren u_1, \dots, u_l auf linear unabhängige Vektoren $\varphi(u_1), \dots, \varphi(u_l)$ in V . Die Vektoren $\varphi(u_1), \dots, \varphi(u_l)$ in V sind also, da φ bijektive lineare Abbildung ist, eine Basis von V , die gerade l Elemente besitzt. Also ist $\dim V = l = \dim U$, was zu beweisen war. \square

Es sei angemerkt, dass wir für surjektive lineare Abbildung $\varphi: U \rightarrow V$ analog die Abschätzung $\dim U \geq \dim V$ und für injektive lineare Abbildung $\varphi: U \rightarrow V$ die Abschätzung $\dim U \leq \dim V$ erhalten können.

Wir bemerken außerdem, dass sowohl die Bijektivität als auch die Linearität im Allgemeinen notwendige Voraussetzungen für dieses Korollar sind. Bijektive Abbildungen kann es auch zwischen Unterräumen unterschiedlicher Dimension geben; diese sind dann allerdings nie linear. An dieser Stelle sei nochmal daran erinnert, dass die „Größe“ unendlicher Mengen ein recht unintuitiver Begriff ist. So haben wir etwa in Mathematischen Grundlagen 1 gesehen, dass es eine Bijektion zwischen der Menge der natürlichen und der Menge der ganzen Zahlen gibt, obwohl man intuitiv die Menge der ganzen Zahlen als „größer“ ansehen würde. Genauso gibt es auch zwischen \mathbb{R} und \mathbb{R}^2 eine Bijektion, die wieder unserer Intuition widerspricht. Allerdings ist eine solche Bijektion recht schwer zu beschreiben, weswegen wir hier darauf verzichten müssen.

Wir spezialisieren das Korollar nochmal für den Fall, dass wir eine lineare Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^k$ vorliegen haben. Diese kann nach dem obigen Korollar höchstens dann bijektiv sein, wenn $n = k$ gilt. (Das heißt selbstverständlich nicht, dass jede lineare Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ bijektiv ist.) Es ist eine wichtige Beobachtung, die wir allerdings nicht beweisen wollen, dass die inverse Abbildung zu einer linearen Abbildung, sofern sie existiert, selbst eine lineare Abbildung sein muss.

Wir wollen die Existenz einer Umkehrabbildung zu einer linearen Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^k$ nochmal in der Sprache der Matrizen formulieren. Sei A die darstellende Matrix von φ . Wie bereits erwähnt, ist die Umkehrabbildung $\psi: \mathbb{R}^k \rightarrow \mathbb{R}^n$ von φ , sofern sie existiert, auch linear, also durch eine Matrix B dargestellt. Nun ist die Abbildung $\varphi \circ \psi$, die von AB dargestellt ist, gleich der Identitätsabbildung von \mathbb{R}^k . Also ist die darstellende Matrix von $\varphi \circ \psi$ gerade die $k \times k$ -Einheitsmatrix, d.h. $AB = E_k$. Genauso sehen wir aus der zweiten Bedingung in der Definition einer Umkehrabbildung von φ , dass auch $BA = E_n$ gelten muss.

Wie wir soeben bereits gesagt haben, kann zu der vorgegeben $k \times n$ -Matrix A eine solche Matrix B höchstens dann existieren, wenn $n = k$ gilt, also höchstens dann, wenn die Matrix quadratisch ist. Dieser Art von Matrizen geben wir einen Namen:

Definition 6.6. Sei A eine $n \times n$ -Matrix. Eine $n \times n$ -Matrix B heißt **invers zu der Matrix A** , falls gilt:

$$AB = BA = E_n.$$

Ist diese Bedingung erfüllt, so wird für B auch A^{-1} geschrieben. Eine Matrix A , zu der es eine inverse Matrix B gibt, heißt **invertierbar**.

Wir haben bereits besprochen, dass eine Matrix genau dann invertierbar ist, wenn die von ihr dargestellte lineare Abbildung eine Inverse besitzt, und diese inverse Abbildung ist in einem solchen Fall auch von der inversen Matrix dargestellt wird.

Die Notation steht im Einklang mit der Notation $x^{-1} = \frac{1}{x}$ für reelle Zahlen $x \neq 0$: Für den Kehrwert x^{-1} von $x \neq 0$ gilt nämlich: $x \cdot x^{-1} = x^{-1} \cdot x = 1$. Ähnlich wie man bei den reellen Zahlen die 0 ausschließen muss, muss man auch beim Invertieren von Matrizen gewisse Matrizen ausschließen; das werden jedoch mehr als nur eine sein.

Wir wollen uns als nächstes mit der Frage beschäftigen, wie man bei einer quadratischen Matrix A prüft, ob die von ihr dargestellte Abbildung bijektiv ist und was gegebenenfalls die Umkehrabbildung ist.

Betrachtet man $AB = E_n$ als eine Gleichung für die vorgegebene $n \times n$ -Matrix A , so kann man diese in mehrere inhomogene lineare Gleichungssysteme aufspalten. Die erste Spalte der Matrix B genügt beispielsweise dem linearen Gleichungssystem, das durch die erweiterte Koeffizientenmatrix $(A|e_1)$ vorgegeben ist, wobei e_1 den ersten Standardbasisvektor in \mathbb{R}^n bezeichnet. Ähnlich sieht das mit den weiteren Spalten von B aus. Löst man die jeweiligen Gleichungssysteme, so erhält man nach und nach die Spalten von B . Wir wollen dafür ein etwas praktischeres Verfahren vorstellen.

Die entscheidende Beobachtung hierbei ist, dass wir beim Lösen eines linearen Gleichungssystems die sinnvollen elementaren Zeilentransformationen alleine anhand der Koeffizientenmatrix bestimmen und dabei nicht die rechte Seite betrachten, obwohl wir die elementaren Zeilentransformationen auch auf die rechte Seite anwenden. Es heißt aber, dass wir mehrere „rechte Seiten“ auf einmal abhandeln können. In unserem Fall wollen wir alle Standardbasisvektoren auf der rechten Seite haben.

Wir betrachten also zu einer $n \times n$ -Matrix A , die wir invertieren wollen, eine $n \times 2n$ -Matrix, die in der linken Hälfte genau die Matrix A hat und auf der rechten Seite die Matrix, die als Spalten gerade die Standardbasisvektoren hat. Das ist allerdings gerade die Einheitsmatrix. Wir benutzen in der Notation wieder einen vertikalen Strich, um die linke Seite von der rechten zu trennen. Diese $n \times 2n$ -Matrix kann also kurz als $(A|E_n)$ notiert werden.

Nun kommen wir zu der zweiten entscheidenden Idee, die wir in diesem Verfahren benutzen wollen: Dem *Gauß-Jordan-Eliminationsverfahren*. Das ist eine Erweiterung des Gauß'sches Eliminationsverfahrens, die das Rückeinsetzen unnötig macht. Dieses Verfahren ist aus theoretischen Überlegungen

sehr hilfreich; es sei aber gleich vorweg angemerkt, dass für die meisten praktischen Anwendungen das Rückeinsetzen sich als sinnvoller, d.h. meist schneller und präziser, erweist. Trotzdem wollen wir dieses theoretische Verfahren, das auch für die Rechnungen „per Hand“ hilfreich sein kann, kennenlernen.

Bislang haben wir, wenn wir ein lineares Gleichungssystem lösen wollten, mit den elementare Zeilentransformationen aufgehört, sobald wir die linke Seite der erweiterten Koeffizientenmatrix in die Zeilenstufenform gebracht haben, und haben von da an mit dem Rückeinsetzen angefangen. Wenn wir allerdings unsere erweiterte Koeffizientenmatrix schon in der Form

$$\left(\begin{array}{cccc|c} 1 & 0 & \dots & 0 & y_1 \\ 0 & 1 & \dots & 0 & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & y_n \end{array} \right),$$

also in der Form $(E_n|y)$ für einen Vektor $y \in \mathbb{R}^n$, geschrieben hätten (was natürlich nur für lineare Gleichungssysteme mit n Gleichungen in n Variablen möglich ist), so ist das Rückeinsetzen besonders einfach: Die einzige Lösung dieses linearen Gleichungssystems ist $y \in \mathbb{R}^n$.

Das *Gauß-Jordan-Verfahren* zum Lösen linearer Gleichungssysteme mit n Gleichungen in n Variablen sieht also vor, die Matrix auf der linken Seite der erweiterten Koeffizientenmatrix so lange umzuformen, bis sie nicht nur in Zeilenstufenform ist, sondern sogar die Einheitsmatrix ist. Es sei angemerkt, dass es auch für nicht-quadratische Koeffizientenmatrizen Varianten des Gauß-Jordan-Verfahrens existieren.

Wir zeigen zunächst in einem sehr einfachen Beispiel, wie man durch Gauß-Jordan-Verfahren zum Lösen der linearen Gleichungssysteme auf ein Verfahren zum Invertieren von Matrizen kommt.

Beispiel. Wir wollen untersuchen, ob die Matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

invertierbar ist.

Dafür müssen wir, wie wir uns im Vorfeld überlegt haben, das lineare Gleichungssystem, das zu der erweiterten Koeffizientenmatrix

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & y_1 \\ 0 & 1 & 1 & y_2 \\ 0 & 0 & 1 & y_3 \end{array} \right)$$

gehört, lösen, und zwar für die drei Fälle

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ oder } \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ oder } \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Wie wir bereits erwähnt haben, wollen wir alle drei lineare Gleichungssysteme auf einmal lösen. Dafür betrachten wir die Matrix

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right).$$

Wir haben die Spalten der rechten Seite, die zu unterschiedlichen Gleichungssystemen gehören, zur Unterscheidung farblich markiert.

Die linke Seite dieser Matrix wollen wir nach Gauß-Jordan-Verfahren für lineare Gleichungssysteme nun durch elementare Zeilentransformationen zur Einheitsmatrix machen. Dabei dürfen wir nicht vergessen, dass die Zeilentransformationen auf beide Seiten gleichermaßen angewandt werden. Wir erhalten also:

$$\begin{array}{ccc} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) & \xrightarrow{I-II} & \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \\ & & \xrightarrow{II-III} & \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right). \end{array}$$

Die Matrix auf der linken Seite ließ sich also durch elementare Zeilentransformationen auf die Einheitsmatrix bringen.

Also ist die eindeutige Lösung des linearen Gleichungssystems, das zu der erweiterten Koeffizientenmatrix

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

gehört, durch die erste Spalte der rechten Seite gegeben und ist gerade der Vektor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Die eindeutige Lösung des linearen Gleichungssystems, das zu der erweiterten Koeffizientenmatrix

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

gehört, finden wir in der zweiten Spalte der rechten Seite und ist somit der Vektor $\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

Schließlich finden wir in der dritten Spalte der rechten Seite die eindeutige Lösung des linearen Gleichungssystems, das zur erweiterten Koeffizientenmatrix

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

gehört, und das ist gerade der Vektor $\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$.

Wir haben uns vorher überlegt, dass diese Lösungen von den obigen Gleichungssystemen gerade die Spalten der Matrix B bilden müssen, die invers zu der uns vorgegebenen Matrix A ist, sofern diese überhaupt existiert. Wir wissen, dass die so erhaltene Matrix B der Gleichung $AB = E_n$ genügt; allerdings ist es nicht offensichtlich, wenn auch wahr, dass daraus in diesem speziellen Fall zweier *quadratischer* Matrizen bereits $BA = E_n$ folgt. Wir wollen auf diesen Beweis verzichten.

Die Matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

mit der wir angefangen haben, ist also invertierbar, und hat als inverse Matrix die Matrix

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

die wir soeben auf der rechten Seite der „großen“ Matrix erhalten haben.

Nun wollen wir das Verfahren etwas allgemeiner zusammenfassen.

Vorher bemerken wir noch, dass die linke Seite nicht immer auf die Form der Einheitsmatrix gebracht werden kann: Wäre es stets möglich, eine $n \times n$ -Matrix durch elementare Zeilentransformationen auf die Einheitsmatrix zu bringen, so hätte jedes lineare Gleichungssystem mit n Gleichungen in n Variablen genau eine Lösung, was natürlich nicht der Fall ist. Insbesondere ist es so, dass wir, sobald wir eine Nullzeile erhalten, bei einer *quadratischen* Matrix wissen, dass wir weniger Stufen als Variablen und somit unendlich viele Lösungen des Gleichungssystems haben. Hat man also durch elementare Zeilentransformationen einer $n \times n$ -Matrix A eine Nullzeile bekommen, so wird man diese Matrix niemals durch elementare Zeilentransformationen zur Matrix E_n machen können.

Wir fassen das bislang gesagte nun in Form eines Verfahrens zusammen. Wir werden, wie auch bei früheren Gelegenheiten, keinen Beweis dafür geben, dass das Verfahren korrekt ist oder stets terminiert.

Gauß-Jordan-Verfahren zum Invertieren von Matrizen

Sei also eine $n \times n$ -Matrix A vorgegeben, von der wir wissen wollen, ob diese invertierbar ist, und gegebenenfalls die inverse Matrix zu A bestimmen wollen.

1. Man betrachte die $n \times 2n$ -Matrix $(A|E_n)$ und versuche, die linke Seite durch elementare Zeilenumformungen, die man wie gewohnt gleichermaßen auf der linken und rechten Seite durchführt, auf die Form der $n \times n$ -Matrix zu bringen.
2. Dabei können zwei Fälle auftreten. Entsteht dabei eine Nullzeile, so ist das Verfahren beendet und die Matrix ist nicht invertierbar. Andernfalls wird es möglich sein, durch elementare Zeilentransformationen die Form $(E_n|B)$ zu erreichen. In diesem Fall ist die Matrix A invertierbar und die Matrix B auf der rechten Seite die inverse Matrix zu A .

Es ist nicht ohne weitere Argumentation klar, warum nur die beiden in dem Verfahren beschriebenen Fälle auftreten können. Außerdem sollte nicht gänzlich klar sein, warum die Matrix B invers zu A ist. Wir erinnern uns, dass die k -te Spalte von B gerade die Lösungen der Gleichungen $Ax = e_k$ für $1 \leq k \leq n$ sind, also die Gleichung $AB = E_n$ dadurch erfüllt sein muss. Dass hingegen die Gleichung $BA = E_n$ erfüllt ist, bedarf weiterer Argumentation.

Es sei ferner angemerkt, dass es unter Umständen, zum Beispiel um das Verfahren konkreter oder vollständiger zu machen, sinnvoll sein kann, die Reihenfolge der elementaren Zeilentransformationen festzulegen, also etwa, jeweils spaltenweise vorzugehen und versuchen, jede Spalte zu dem jeweiligen Standardbasisvektor zu machen. Für den Menschen ist ein solches „striktes“ Verfahren allerdings nicht unbedingt schneller, ähnlich wie das bereits beim Gauß'schen Eliminationsverfahren der Fall gewesen war.

Wir machen an dieser Stelle ein einfaches Beispiel zum Invertieren von 3×3 -Matrizen.

Beispiel. Wir betrachten ein nicht allzu kompliziertes Beispiel. In diesem wollen wir die Matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ 3 & 5 & 1 \end{pmatrix}$$

auf Invertierbarkeit untersuchen. Dabei gehen wir nach dem obigen Verfahren vor: Wir schreiben die vorgegebene Matrix und daneben die 3×3 -Einheitsmatrix auf und führen so lange elementare Zeilentransformationen durch, bis links entweder eine Nullzeile oder eine Einheitsmatrix entsteht.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 2 & 5 & 0 & 0 & 1 & 0 \\ 3 & 5 & 1 & 0 & 0 & 1 \end{array} \right) \xrightarrow{\text{II}-2\text{I}, \text{III}-3\text{I}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 5 & 0 & -2 & 1 & 0 \\ 0 & 5 & 1 & -3 & 0 & 1 \end{array} \right)$$

$$\begin{array}{l} \xrightarrow{\text{III-II}} \\ \xrightarrow{\text{II/5}} \end{array} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 5 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right)$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{2}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right)$$

Wir haben die Einheitsmatrix auf der linken Seite durch elementare Zeilentransformationen erreichen können. Die Matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ 3 & 5 & 1 \end{pmatrix}$$

ist also invertierbar und hat als inverse Matrix die Matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ -\frac{2}{5} & \frac{1}{5} & 0 \\ -1 & -1 & 1 \end{pmatrix}.$$

Nun wollen wir, zunächst nur anhand des Beispiels und dann allgemeiner, eine Verbindung zu Koordinaten eines Vektors bezüglich einer Basis von \mathbb{R}^n , in diesem Fall von \mathbb{R}^3 , herstellen.

Wir bemerken, dass die Spalten der Matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ 3 & 5 & 1 \end{pmatrix}$$

eine Basis von \mathbb{R}^3 bilden, und zwar eine, die wir bereits im letzten Beispiel des Kapitels 3 betrachtet haben. Die Koordinaten eines beliebigen Vektors in \mathbb{R}^3 bezüglich dieser Basis sind, wie wir in dem früheren Beispiel ermittelt haben, gegeben durch

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \frac{y-2x}{5} \cdot \begin{pmatrix} 0 \\ 5 \\ 5 \end{pmatrix} + (z-x-y) \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Nun behaupten wir, dass diese Koordinaten in Verbindung mit den Einträgen der inversen Matrix von

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ 3 & 5 & 1 \end{pmatrix}$$

stehen. Dafür multiplizieren wir den Vektor $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$, dessen Koordinaten be-

züglich der Basis $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ wir bestimmen wollen, mit der inversen

Matrix, die wir soeben berechnet haben:

$$\begin{aligned} \begin{pmatrix} 1 & 0 & 0 \\ -\frac{2}{5} & \frac{1}{5} & 0 \\ -1 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 1 \cdot x + 0 \cdot y + 0 \cdot z \\ -\frac{2}{5} \cdot x + \frac{1}{5} \cdot y + 0 \cdot z \\ -1 \cdot x + (-1) \cdot y + 1 \cdot z \end{pmatrix} \\ &= \begin{pmatrix} x \\ \frac{-2x+y}{5} \\ -x - y + z \end{pmatrix}. \end{aligned}$$

Bei genauer Betrachtung stellt man fest, dass das genau die Koordinaten des Vektors $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ bezüglich der Basis $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ sind.

Nachfolgend wollen wir uns, zumindest im Ansatz, überlegen, dass dies kein Zufall ist, sondern ein allgemeines Verfahren zur Berechnung der Koordinaten eines Vektors in \mathbb{R}^n bezüglich einer beliebigen Basis von \mathbb{R}^n bilden.

Eine wichtige Beobachtung, die wir im vorherigen Beispiel gemacht haben, war, dass die Spalten der vorgegebenen Matrix eine Basis von \mathbb{R}^3 bildeten. Dass auch das kein Zufall ist, wollen wir uns als nächstes überlegen.

Bemerkung. Sei A eine $n \times n$ -Matrix. Diese Matrix ist genau dann invertierbar, wenn ihre Spalten eine Basis von \mathbb{R}^n bilden.

Wir wollen diese Aussage zwar nicht beweisen, jedoch einen Beweis skizzieren. Zunächst bemerken wir nochmal, dass diese Aussage nur für quadratische Matrizen Sinn ergibt, da andere Matrizen nicht invertierbar sein können. Auf der anderen Seite muss die Anzahl der Vektoren in jeder Basis von \mathbb{R}^n genau n sein, weswegen die zweite Hälfte der Aussage nur für quadratische Matrizen gelten kann.

Wir erinnern uns, dass n Vektoren in \mathbb{R}^n genau dann eine Basis bilden, wenn sie linear unabhängig sind. Die Spalten einer $n \times n$ -Matrix A bilden also genau dann eine Basis von \mathbb{R}^n , wenn diese Spalten linear unabhängig sind.

Als nächstes erinnern wir uns an das Verfahren, um die lineare Abhängigkeit von Vektoren in \mathbb{R}^n festzustellen. Dafür schrieben wir die Vektoren in eine Matrix und brachten diese auf Zeilenstufenform durch elementare Zeilentransformationen. Die Vektoren sind genau dann linear unabhängig, wenn wir dabei genauso viele Stufen wie Vektoren erhalten. Bei einer *quadratischen* Matrix (und nur bei einer solchen!) ist das genau dann der Fall, wenn wir in der Zeilenstufenform keine Nullzeile erhalten.

Schreibt man die Spalten von A als Spalten in eine Matrix, so erhält man gerade die Matrix A . Die Spalten von A sind also genau dann linear unabhängig, wenn man, sobald diese Matrix durch elementare Transformationen in Zeilenstufenform gebracht hat, lauter Einsen auf der Diagonale hat und somit genauso viele Stufen wie Spalten (und insbesondere keine Nullzeile).

An dieser Stelle vergleichen wir das mit dem Gauß-Jordan-Verfahren zum Invertieren von A . Hat man eine $n \times n$ -Matrix A durch elementare Zeilentransformationen auf Zeilenstufenform mit lauter Einsen auf der Diagonale gebracht, so ist es nicht schwer einzusehen, dass man auch alle Einträge oberhalb der Diagonale eliminieren kann, und die Matrix deshalb invertierbar ist. Lässt sich umgekehrt die Matrix durch elementare Zeilentransformationen auf die Form der Einheitsmatrix bringen, so ist es insbesondere Zeilenstufenform mit lauter Einsen auf der Diagonale.

Zusammenfassend würde also bei genauer Ausführung folgen, dass die $n \times n$ -Matrix genau dann invertierbar ist, wenn ihre Spalten eine Basis von \mathbb{R}^n bilden.

Wir notieren die Zwischenschritte nochmal schematisch für eine $\mathbf{n} \times \mathbf{n}$ -Matrix A :

$$\begin{aligned}
 & \text{Spalten von } A \text{ bilden eine Basis von } \mathbb{R}^n \\
 \Leftrightarrow & \text{Spalten von } A \text{ sind linear unabhängig} \\
 \Leftrightarrow & A \text{ hat } n \text{ Stufen, wenn in ZSF gebracht} \\
 \Leftrightarrow & A \text{ lässt sich durch el. Zeilentransformationen auf } E_n \text{ bringen} \\
 \Leftrightarrow & A \text{ ist invertierbar.}
 \end{aligned}$$

Dass die Beobachtung des obigen Beispiels kein Zufall ist, wollen wir im folgenden Satz festhalten, auf dessen Beweis wir verzichten wollen.

Satz 6.7. *Sei $u_1, \dots, u_n \in \mathbb{R}^n$ eine Basis von \mathbb{R}^n , und sei A die $n \times n$ -Matrix, die gerade die Vektoren u_1, \dots, u_n als Spalten hat. Dann ist die Matrix A invertierbar und die Koordinaten jedes Vektors $v \in \mathbb{R}^n$ bezüglich der Basis u_1, \dots, u_n sind gerade die Einträge des Vektors $A^{-1} \cdot v$.*

Wir wollen jedoch die Beweisidee kurz erläutern. Wollen wir unter den Voraussetzungen des Satzes die Koordinaten von $v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ bezüglich der Basis u_1, \dots, u_n bestimmen, also reelle Zahlen $\lambda_1, \dots, \lambda_n$, sodass

$$\lambda_1 u_1 + \lambda_2 u_2 + \dots + \lambda_n u_n = v$$

gilt. Das ist nun, schreibt man diese Gleichung aus, ein lineares Gleichungssystem mit der erweiterten Koeffizientenmatrix $(A|v)$. Man kann diese Gleichung mit Hilfe der Matrix-Vektor-Multiplikation anders schreiben, und zwar gerade in die Form

$$A \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Wäre uns eine Gleichung $a\lambda = x$ in reellen Zahlen vorgegeben, so würden wir, um λ in Termen von x und a zu bestimmen, beide Seiten der Gleichung durch a teilen, also gerade mit a^{-1} multiplizieren, was natürlich nur möglich ist, wenn a nicht Null ist.

Hier kann man ähnlich verfahren: Ist A invertierbare Matrix, so können wir beide Seiten mit A^{-1} multiplizieren und erhalten

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = A^{-1} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Das ist genau die Behauptung des Satzes.

7 Eigenwerte und Eigenvektoren

Als nächstes fragen wir uns, wann eine lineare Abbildung eine besonders einfache Form hat. Diese Frage ist natürlich viel zu vage und hat viele unterschiedliche Antworten. Wir wollen zunächst Beispiele von besonders einfachen linearen Abbildung präsentieren, und danach die Frage genauer formulieren.

Beispiel. Für jeden Unterraum $U \subseteq \mathbb{R}^n$ und jede reelle Zahl $c \in \mathbb{R}$ ist die Abbildung

$$\begin{aligned} \varphi: U &\rightarrow U, \\ w &\mapsto cw, \end{aligned}$$

eine lineare Abbildung. Tatsächlich lassen sich die Bedingungen (A1) und (A2) für Linearität leicht nachprüfen.

Sind nämlich $w_1, w_2 \in U$ zwei Vektoren, so gilt

$$\varphi(w_1 + w_2) = c \cdot (w_1 + w_2) = c \cdot w_1 + c \cdot w_2 = \varphi(w_1) + \varphi(w_2),$$

wobei wir nur die Definition von φ und das Distributivgesetz ausgenutzt haben. Somit ist (A1) für φ erfüllt.

Hat man nun eine weitere reelle Zahl $\lambda \in \mathbb{R}$ und einen Vektor $w \in U$ vorgegeben, dann gilt:

$$\varphi(\lambda w) = c \cdot (\lambda w) = (c \cdot \lambda) \cdot w = (\lambda \cdot c) \cdot w = \lambda \cdot (c \cdot w) = \lambda \varphi(w).$$

Hierbei haben wir die Definition von φ , die Assoziativität der Skalarmultiplikation und die Kommutativität der Multiplikation für die reellen Zahlen ausgenutzt. Somit gilt auch die Bedingung (A2) für φ , und φ ist tatsächlich eine lineare Abbildung.

Diese Abbildung ist sehr einfach, insbesondere ist es einfach, den anschaulichen Effekt dieser Abbildung zu beschreiben: Das Bild $\varphi(w)$ jedes Vektors $w \in U$ zeigt in dieselbe (oder genau entgegengesetzte, falls $c < 0$, oder in gar keine, wenn $c = 0$) Richtung wie w , und unterscheidet sich von w nur durch einen Streckungs- bzw. Stauchungsfaktor.

Es sei noch angemerkt, dass im Fall $U = \mathbb{R}^n$ auch die darstellende Matrix von φ besonders einfache Form hat. Da die Bilder der Standardbasisvektoren ebenfalls durch Multiplikation mit dem festen Faktor c gegeben sind, haben wir in jeder Spalte der darstellenden Matrix von φ gerade das c -fache des entsprechenden Standardbasisvektors, sodass wir insgesamt eine Matrix, die nur c 's auf der Diagonale und sonst nur Nullen als Einträge hat, also etwa

$$\begin{pmatrix} c & 0 & \dots & 0 \\ 0 & c & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c \end{pmatrix},$$

Nun betrachten wir eine etwas kompliziertere Variante dieses Beispiels.

Beispiel. Manchmal möchte man Abbildungen untersuchen, die in unterschiedliche Richtungen mit unterschiedlichen Streckungsfaktoren strecken. Das kennt man etwa von Bildbearbeitungsprogrammen, wo man horizontal und vertikal mit unterschiedlichen Streckungsfaktoren strecken kann und dadurch meist ein verzerrtes Bild erhält.

Ein Beispiel einer solchen Abbildung kann man präziser beschreiben als die lineare Abbildung mit der darstellenden Matrix

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Alle Vektoren „in x -Richtung“, also von der Form $\begin{pmatrix} x \\ 0 \end{pmatrix}$, werden von dieser Abbildung um den Faktor 2 gestreckt:

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} 2x \\ 0 \end{pmatrix}.$$

Alle Vektoren „in y -Richtung“, also von der Form $\begin{pmatrix} 0 \\ y \end{pmatrix}$, werden von dieser Abbildung um den Faktor 3 gestreckt:

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 3y \end{pmatrix}.$$

Bilder von anderen Vektoren unter dieser Abbildung werden im Allgemeinen allerdings nicht in dieselbe Richtung wie der ursprüngliche Vektor

zeigen. Beispielsweise wird der Vektor $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ nicht auf Vielfaches dieses Vektors abgebildet:

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Diese Abbildung ist also anders und durchaus komplizierter als eine bloße Streckung in alle Richtungen um denselben Faktor, aber in gewisser Weise immer noch sehr einfach. Wir wollen nun ähnliche, allerdings etwas allgemeinere Abbildungen beschreiben.

Wir interessieren uns nun dafür, ob eine Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ Vektoren, die in gewisse Richtungen zeigen, nur streckt und ihre Richtung (im Wesentlichen) nicht verändert. Diese Richtungen müssen natürlich im Allgemeinen nicht, wie in dem obigen Beispiel, gerade in Richtung der Standardbasisvektoren zeigen, was die Suche nach solchen Richtungen verkompliziert. Wir geben nun die formalen Definitionen, die das Konzept der „Streckungen in gewisse Richtungen“ präzisieren.

Definition 7.1. Sei $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung. Ein Vektor $v \in \mathbb{R}^n \setminus \{0\}$ heißt **Eigenvektor** der Abbildung φ zum **Eigenwert** $\lambda \in \mathbb{R}$, falls

$$\varphi(v) = \lambda v$$

gilt.

Eine reelle Zahl $\lambda \in \mathbb{R}$ heißt **Eigenwert** der Abbildung φ , falls es mindestens einen Vektor $v \in \mathbb{R}^n$ mit $v \neq 0$ gibt, der Eigenvektor von φ zum Eigenwert λ ist.

Die Abbildung φ heißt **diagonalisierbar**, falls es eine Basis von \mathbb{R}^n gibt, die aus Eigenvektoren von φ besteht.

Wir wollen nun einige erste Kommentare zu dieser Definition geben.

Bemerkung. • Der Nullvektor wird nicht als Eigenvektor einer linearen Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ betrachtet. Dieser würde für jede beliebige reelle Zahl μ Eigenvektor sein, wenn man ihn nicht explizit ausschließen würde, da

$$\varphi(0) = 0 = \mu \cdot 0$$

immer gilt. Dann wäre auch jede reelle Zahl ein Eigenwert von φ . Wir sind aber daran interessiert, die „wirklichen“ Streckungsfaktoren zu finden. Deswegen schließen wir 0 als Eigenvektor aus.

Hingegen ist die 0 als *Eigenwert* durchaus zulässig. In diesem fragt man sich also, ob es andere Vektoren als den Nullvektor gibt, die von der Abbildung φ auf Null abgebildet werden. Das ist eine wichtige Information über die Abbildung φ .

- Eine lineare Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist, wie wir uns bereits klargemacht haben, durch ihre darstellende $n \times n$ -Matrix A vollständig festgelegt. Da man beim Bestimmen der Eigenwerte und Eigenvektoren häufig mit der Matrix arbeitet, spricht man auch dafür, dass es Eigenvektoren und Eigenwerte der Matrix A sind, womit man dasselbe wie Eigenvektoren bzw. Eigenwerte von φ meint. Mit anderen Worten heißt ein Vektor $v \in \mathbb{R}^n$, der nicht der Nullvektor ist, ein Eigenvektor von der Matrix A zum Eigenwert λ , falls für diesen Vektor $Av = \lambda v$ gilt.
- Sei $v \in \mathbb{R}^n$ ein Eigenvektor der linearen Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ zum Eigenwert λ . Dann ist cv ebenfalls ein Eigenvektor von φ zum Eigenwert λ für jede reelle Zahl $c \in \mathbb{R} \setminus \{0\}$. Den Fall $c = 0$ müssen wir ausschließen, da $0 \cdot v = 0$ wie bereits erläutert nicht als Eigenvektor betrachtet wird. Für die anderen Vielfachen cv (also mit $c \neq 0$) ist das Ergebnis nicht der Nullvektor, da $v \neq 0$ ist, und es gilt:

$$\varphi(cv) = c\varphi(v) = c(\lambda v) = (c\lambda)v = (\lambda c)v = \lambda \cdot (cv).$$

Hierbei haben wir im ersten Schritt ausgenutzt, dass die Abbildung φ linear ist und somit der Bedingung (A2) genügt. Im zweiten wurde die Definition des Eigenvektors eingesetzt, und danach haben wir die Assoziativität der Skalarmultiplikation und die Kommutativität der Multiplikation der reellen Zahlen ausgenutzt.

Insgesamt haben wir also gesehen, dass das Einsetzen des Vielfachen cv des Eigenvektors v in φ als Ergebnis das λ -fache von cv liefert, also ist cv ebenfalls ein Eigenvektor von φ zum Eigenwert λ .

Informell gesagt lässt sich diese Aussage wie folgt zusammenfassen: Wird ein Vektor v von der Abbildung φ um einen Streckungsfaktor λ gestreckt, so wird auch jedes Vielfache dieses Vektors von φ um denselben Faktor gestreckt.

Zunächst betrachten wir ein sehr einfaches Beispiel.

Beispiel. Wir betrachten die lineare Abbildung T_B von \mathbb{R}^2 nach \mathbb{R}^2 , die durch die Matrix

$$B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

dargestellt wird. Dafür ist der Vektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ein Eigenvektor, und zwar zum Eigenwert 2, denn es gilt:

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Nach der letzten Bemerkung ist auch jedes Vielfache $\begin{pmatrix} x \\ 0 \end{pmatrix}$ mit $x \neq 0$ des Vektors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ebenfalls ein Eigenvektor von B zum Eigenwert 2.

Ähnlich sieht man, dass $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und somit jedes $\begin{pmatrix} 0 \\ y \end{pmatrix}$ mit $y \neq 0$ ein Eigenvektor von B zum Eigenwert 3 ist:

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix} = 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Insbesondere bemerken wir, dass die Standardbasis von \mathbb{R}^2 aus Eigenvektoren von B besteht. Also ist B diagonalisierbar. Wir bemerken, dass B sogar eine *Diagonalmatrix* ist, also eine quadratische Matrix, die nur auf der Diagonale von Null verschiedene Einträge hat. Der Begriff „diagonalisierbar“ hat damit zu tun, dass alle diagonalisierbaren Matrizen im gewissen Sinne ähnlich zu einer Diagonalmatrix sind, wenn auch im allgemeinen komplizierter.

Nun stellt sich die Frage, wie man die Eigenvektoren und Eigenwerte einer vorgegebenen Matrix bestimmen kann. Wir beginnen damit, das in einem Beispiel zu erläutern. Der Einfachheit halber beschränken wir uns bei den Beispielen auf 2×2 -Matrizen. Wir werden aber kurz ein mögliches allgemeines Vorgehen angeben.

Beispiel. Wir betrachten die 2×2 -Matrix

$$B = \begin{pmatrix} 3 & 5 \\ 4 & 2 \end{pmatrix}$$

und wollen alle Eigenwerte und Eigenvektoren von dieser Matrix finden. Wir suchen also alle Vektoren $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ und alle reellen Zahlen $\lambda \in \mathbb{R}$, sodass

$$\begin{pmatrix} 3 & 5 \\ 4 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

gilt. Dabei wissen wir bereits, dass $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ dieser Gleichung für alle $\lambda \in \mathbb{R}$ genügt, und betrachten diese Lösung als uninteressant.

Wir schreiben diese Gleichung nun etwas anders auf:

$$\begin{aligned} & \begin{pmatrix} 3 & 5 \\ 4 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix} \\ \Leftrightarrow & \begin{cases} 3x + 5y = \lambda x \\ 4x + 2y = \lambda y \end{cases} \end{aligned}$$

$$\Leftrightarrow \begin{cases} (3 - \lambda)x + 5y = 0 \\ 4x + (2 - \lambda)y = 0 \end{cases}$$

Nun ist die folgende Überlegung hilfreich. Zwar sind sowohl λ als auch x, y Unbekannte, deren Werte wir bestimmen wollen. Jedoch stellt es sich als sinnvoll heraus, ihnen unterschiedliche Rollen zukommen zu lassen. Wir behandeln λ als Parameter, und x, y als Variablen. Dann ist das obige Gleichungssystem linear und homogen in den Variablen x, y und hat die Koeffizientenmatrix

$$\begin{pmatrix} 3 - \lambda & 5 \\ 4 & 2 - \lambda \end{pmatrix}.$$

Würden wir λ ebenfalls als Variable behandeln, so wäre das Gleichungssystem nicht linear, und wir hätten zunächst keine gute Methode, um es zu lösen. Wir beobachten außerdem, dass die Koeffizientenmatrix dieses Gleichungssystem aus der ursprünglichen Matrix B dadurch entsteht, dass wir von allen Diagonaleinträgen ein λ subtrahieren.

Nun fragen wir uns, was die Lösungen dieses linearen homogenen Gleichungssystems sind. Wir wissen bereits, dass $x = y = 0$ eine Lösung des Gleichungssystems ist, die wir allerdings als „uninteressant“ betrachten. Wir suchen also alle Werte des Parameters λ , für die das Gleichungssystem mehr als diese eine (und somit unendlich viele) Lösungen besitzt. Wir formen wie immer die Matrix durch elementare Zeilentransformationen um:

$$\begin{aligned} \begin{pmatrix} 3 - \lambda & 5 \\ 4 & 2 - \lambda \end{pmatrix} & \xrightarrow{I \leftrightarrow II} \begin{pmatrix} 4 & 2 - \lambda \\ 3 - \lambda & 5 \end{pmatrix} \\ & \xrightarrow{I/4} \begin{pmatrix} 1 & \frac{1}{2} - \frac{\lambda}{4} \\ 3 - \lambda & 5 \end{pmatrix} \\ & \xrightarrow{II - (3 - \lambda)I} \begin{pmatrix} 1 & \frac{1}{2} - \frac{\lambda}{4} \\ 0 & 5 - (3 - \lambda)\left(\frac{1}{2} - \frac{\lambda}{4}\right) \end{pmatrix} \end{aligned}$$

Diese Matrix ist nun in Zeilenstufenform. Sie hat entweder eine oder zwei Stufen, abhängig davon, ob der Eintrag in der zweiten Spalte, zweiten Zeile, also der Ausdruck

$$5 - (3 - \lambda)\left(\frac{1}{2} - \frac{\lambda}{4}\right)$$

Null ist oder nicht. Ist dieser Ausdruck nicht Null, so hat das lineare Gleichungssystem nur die Lösung $x = y = 0$ und somit ist jedes λ , für das dieser Ausdruck nicht Null ist, *kein* Eigenwert von B . Ist für ein λ der Ausdruck

$$5 - (3 - \lambda)\left(\frac{1}{2} - \frac{\lambda}{4}\right)$$

hingegen 0, so hat die Koeffizientenmatrix des betrachteten linearen Gleichungssystems nur eine Stufe und zwei Variablen, was bedeutet, dass das lineare Gleichungssystem unendlich viele Lösungen hat, sodass solche $\lambda \in \mathbb{R}$ Eigenwerte von B sind.

Wir müssen also die Gleichung

$$5 - (3 - \lambda) \left(\frac{1}{2} - \frac{\lambda}{4} \right) = 0$$

lösen, und die Lösungen dieser Gleichung werden gerade die Eigenwerte von B sein.

Wir lösen also die Klammern auf und fassen dann die Summanden zusammen, um die Gleichung zu vereinfachen:

$$\begin{aligned} & 5 - (3 - \lambda) \left(\frac{1}{2} - \frac{\lambda}{4} \right) = 0 \\ \Leftrightarrow & 5 - \left(\frac{3}{2} - \frac{3\lambda}{4} - \frac{\lambda}{2} + \frac{\lambda^2}{4} \right) = 0 \\ \Leftrightarrow & 5 - \left(\frac{3}{2} - \frac{5\lambda}{4} + \frac{\lambda^2}{4} \right) = 0 \\ \Leftrightarrow & 5 - \frac{3}{2} + \frac{5\lambda}{4} - \frac{\lambda^2}{4} = 0 \mid \cdot (-4) \\ \Leftrightarrow & -20 + 6 - 5\lambda + \lambda^2 = 0 \\ \Leftrightarrow & \lambda^2 - 5\lambda - 14 = 0. \end{aligned}$$

Nun haben wir die Gleichung durch Termumformungen und Äquivalenzumformungen so weit vereinfacht, dass dies eine quadratische Gleichung ist, und zwar in der Standardform, auf die wir die $p - q$ -Formel anwenden können. Diese liefert nun:

$$\begin{aligned} & \lambda^2 - 5\lambda - 14 = 0 \\ \Leftrightarrow & \lambda = \frac{5}{2} + \sqrt{\left(\frac{5}{2}\right)^2 + 14} \text{ oder } \lambda = \frac{5}{2} - \sqrt{\left(\frac{5}{2}\right)^2 + 14} \\ \Leftrightarrow & \lambda = \frac{5}{2} + \sqrt{\frac{25}{4} + 14} \text{ oder } \lambda = \frac{5}{2} - \sqrt{\frac{25}{4} + 14} \\ \Leftrightarrow & \lambda = \frac{5}{2} + \sqrt{\frac{25}{4} + \frac{56}{4}} \text{ oder } \lambda = \frac{5}{2} - \sqrt{\frac{25}{4} + \frac{56}{4}} \\ \Leftrightarrow & \lambda = \frac{5}{2} + \sqrt{\frac{81}{4}} \text{ oder } \lambda = \frac{5}{2} - \sqrt{\frac{81}{4}} \\ \Leftrightarrow & \lambda = \frac{5}{2} + \frac{9}{2} \text{ oder } \lambda = \frac{5}{2} - \frac{9}{2} \\ \Leftrightarrow & \lambda = 7 \text{ oder } \lambda = -2. \end{aligned}$$

Die einzigen Eigenwerte der Matrix B sind also die reellen Zahlen $\lambda = 7$ und $\lambda = -2$. Nun wollen wir die zugehörigen Eigenvektoren bestimmen. Wir wissen, dass es (bis auf den Nullvektor) genau die Lösungen des linearen Gleichungssystems mit der Koeffizientenmatrix

$$\begin{pmatrix} 1 & \frac{1}{2} - \frac{\lambda}{4} \\ 0 & 5 - (3 - \lambda)\left(\frac{1}{2} - \frac{\lambda}{4}\right) \end{pmatrix}$$

ist. Um die Eigenvektoren zu bestimmen, setzen wir also die beiden reellen Zahlen, die wir als einzige Eigenwerte bestimmt haben, jeweils ein und lösen das zugehörige lineare Gleichungssystem. Wir beachten dabei, dass wir den Eintrag $(2, 2)$ nicht auszurechnen brauchen, da wir die Werte von λ gerade so bestimmt haben, dass dieser Eintrag Null ist. (Es kann jedoch sinnvoll sein, diese Rechnung durchzuführen, um eventuell bei der Rechnung entstandene Fehler aufzuspüren.)

Für $\lambda = 7$ wird die Koeffizientenmatrix zu

$$\begin{pmatrix} 1 & \frac{1}{2} - \frac{7}{4} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -\frac{5}{4} \\ 0 & 0 \end{pmatrix}.$$

Die Lösungsmenge des zugehörigen linearen Gleichungssystems ist gegeben durch

$$\left\{ \mu \begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix} \mid \mu \in \mathbb{R} \right\}.$$

Also ist jeder Vektor von der Form $\begin{pmatrix} 5 \\ 4 \\ y \end{pmatrix}$ mit $y \neq 0$ ein Eigenvektor der Matrix B zum Eigenwert 7, und das sind auch alle Eigenvektoren von B zum Eigenwert 7.

Als nächstes machen wir mit dem Eigenwert $\lambda = -2$ weiter. Die Koeffizientenmatrix wird in diesem Fall zu

$$\begin{pmatrix} 1 & \frac{1}{2} - \frac{-2}{4} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Die Lösungsmenge des zugehörigen linearen Gleichungssystems ist gegeben durch

$$\left\{ \gamma \begin{pmatrix} 1 \\ -1 \end{pmatrix} \mid \gamma \in \mathbb{R} \right\}.$$

Also ist jeder Vektor von der Form $\begin{pmatrix} \gamma \\ -\gamma \end{pmatrix}$ mit $\gamma \neq 0$ ein Eigenvektor der Matrix B zum Eigenwert -2 , und das sind auch alle Eigenvektoren von B zum Eigenwert -2 .

Zuletzt fragen wir uns, ob die Matrix B diagonalisierbar ist. Das ist tatsächlich der Fall: Beispielsweise bilden die Vektoren $\begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ eine Basis von \mathbb{R}^2 (es kann zum Beispiel leicht mit unserem Kriterium nachgeprüft werden, dass diese Vektoren linear unabhängig sind.) Ferner ist jeder dieser Vektoren jeweils ein Eigenvektor der Matrix B , jeweils zum Eigenwert 7 bzw. -2 . Somit hat \mathbb{R}^2 eine Basis aus Eigenvektoren von B , und B ist diagonalisierbar.

Wir wollen in diesem Beispiel etwas genauer erläutern, in wie fern eine Basis aus Eigenvektoren für eine Matrix ähnliche Eigenschaften wie bei einer Diagonalmatrix bedeutet.

Da die Vektoren $\begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ eine Basis von \mathbb{R}^2 bilden, lässt sich jeder Vektor $v \in \mathbb{R}^2$ eindeutig schreiben als

$$v = c_1 \begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

mit reellen Zahlen $c_1, c_2 \in \mathbb{R}$, die für jeden Vektor $v \in \mathbb{R}^2$ jeweils eindeutig bestimmt sind. Wendet man nun die von B dargestellte lineare Abbildung T_B auf v an, so erhält man unter Ausnutzung der Bedingungen (A1) und (A2) für die lineare Abbildung T_B :

$$\begin{aligned} T_B(v) &= T_B \left(c_1 \begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right) \\ &= T_B \left(c_1 \begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix} \right) + T_B \left(c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right) \\ &= c_1 T_B \left(\begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix} \right) + c_2 T_B \left(\begin{pmatrix} 1 \\ -1 \end{pmatrix} \right). \end{aligned}$$

Nun nutzen wir aus, dass beide Vektoren der Basis $\begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ auch Eigenvektoren von B bzw. T_B sind, und die Werte von T_B auf diesen Vektoren einfach Vielfache von diesen Vektoren sind. Eingesetzt ergibt das:

$$\begin{aligned} T_B(v) &= c_1 T_B \left(\begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix} \right) + c_2 T_B \left(\begin{pmatrix} 1 \\ -1 \end{pmatrix} \right) \\ &= c_1 \cdot 7 \begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix} + c_2 \cdot (-2) \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= 7c_1 \begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix} + (-2c_2) \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \end{aligned}$$

Die eindeutigen Koordinaten des Vektors $T_B(v)$ bezüglich der Basis $\begin{pmatrix} 5 \\ 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ können also leicht aus den Koordinaten c_1, c_2 von v bezüglich derselben Basis

berechnet werden. Die Koordinaten von $T_B(v)$ sind dann gerade $7c_1, -2c_2$. Diese werden also durch Multiplikation mit den Eigenwerten aus den Koordinaten von v erhalten. Ähnlich ist es auch bei Diagonalmatrizen, allerdings werden da die Koordinaten bezüglich der Standardbasis, also gerade die Einträge des Vektors in \mathbb{R}^n benutzt: Multiplikation mit einer Diagonalmatrix bewirkt gerade, dass die jeweiligen Einträge des Vektors mit den Zahlen auf der Diagonale der Diagonalmatrix skaliert werden. Das motiviert, zumindest zum Teil, den Namen „diagonalisierbar“.

Wir wollen kurz darauf eingehen, wie man bei einer beliebigen vorgegebenen Matrix A die Frage nach Eigenvektoren und Eigenwerten von A untersuchen könnte. Sei A eine $n \times n$ -Matrix, die wie gewohnt in der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}.$$

Wir erinnern uns nochmal daran, dass Eigenvektoren und Eigenwerte nur für quadratische Matrizen Sinn ergeben: Geht die von der Matrix dargestellte lineare Abbildung in einen Raum anderer Dimension, so ergibt es wenig Sinn zu fragen, ob der Bildvektor ein Vielfaches des ursprünglichen Vektors ist, da die beiden Elemente in unterschiedlichen Vektorräumen sind und in dieser Form nicht verglichen werden können.

Will man also die Eigenvektoren der Matrix A finden, also Vektoren $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in$

\mathbb{R}^n , die nicht der Nullvektor sind und der Gleichung

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

für ein $\lambda \in \mathbb{R}$ genügen, so wollen wir das wieder zu einem linearen Gleichungssystem in Standardform mit Parameter λ umwandeln. Dabei gehen wir wie im obigen Beispiel vor:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$\begin{aligned}
&\Leftrightarrow \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nn}x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix} \\
&\Leftrightarrow \begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = \lambda x_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = \lambda x_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nn}x_n = \lambda x_n \end{cases} \\
&\Leftrightarrow \begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = 0 \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + a_{23}x_3 + \dots + a_{2n}x_n = 0 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + (a_{nn} - \lambda)x_n = 0 \end{cases}
\end{aligned}$$

Hierbei haben wir im ersten Schritt das Produkt der Matrix A mit dem Vektor $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ ausgerechnet, im zweiten die Gleichheit von Vektoren als kom-

ponentenweise Gleichheit umgeschrieben und im letzten Schritt das lineare Gleichungssystem in Standardform gebracht. Hierbei bemerken wir, dass auf der rechten Seite in der i -ten Zeile nur die Variable x_i vorkommt, also im Vergleich zur Matrix A die Koeffizientenmatrix dieses linearen Gleichungssystem nur in den Diagonaleinträgen anders ist. Die Diagonaleinträge der Matrix A werden um ein $-\lambda$ abgeändert. Die Koeffizientenmatrix des linearen Gleichungssystems zur Bestimmung der Eigenvektoren zum potentiellen Eigenwert λ ist also genau die Matrix A , die auf der Diagonale durch $-\lambda$ abgeändert wurde, also in diesem Fall:

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} - \lambda \end{pmatrix}.$$

Ein exaktes Verfahren zur Auffindung der Eigenvektoren der Matrix A besteht also darin, diese modifizierte Matrix als Koeffizientenmatrix eines homogenen linearen Gleichungssystems in Abhängigkeit von λ zu betrachten und alle Werte von λ zu finden, für die das zugehörige lineare Gleichungssystem mehr als eine Lösung besitzt. Das sind genau die Eigenwerte von A . Bestimmung der Eigenvektoren besteht dann darin, die Lösungsmengen für die entsprechenden Werte von λ tatsächlich anzugeben. Dabei dürfen wir nicht vergessen, dass der Nullvektor nicht als Eigenvektor der Matrix A betrachtet wird.

Es sei angemerkt, dass dieses Verfahren unter Umständen - wenn auch nicht notwendigerweise - auf eine Lösung einer Gleichung bezüglich λ hinausläuft, die höhere Potenzen von λ , also $\lambda, \lambda^2, \dots, \lambda^n$ enthält (wobei n genau die Anzahl der Zeilen bzw. der Spalten von A ist). In dem obigen Beispiel einer 2×2 -Matrix erhielten wir so eine quadratische Gleichung, die leicht durch die $p - q$ -Formel zu lösen ist. Für $n = 3$ und $n = 4$ hat man da noch präzise, wenn auch komplizierte Formeln; jedoch muss man für höhere Werte von n im Allgemeinen zum Lösen solcher Gleichungen auf numerische Methoden zurückgreifen und erhält keine präzise Lösung. Es sei ferner noch erwähnt, dass in der Praxis, wo häufig Eigenvektoren viel größerer Matrizen bestimmt werden müssen, ganz andere Näherungsverfahren existieren, auf die wir möglicherweise später zu sprechen kommen.

Nun wollen wir ein Beispiel einer nicht-diagonalisierbaren Matrix geben. Tatsächlich wird sich sogar zeigen, dass diese Matrix überhaupt keine Eigenwerte hat, also informell gesagt gar keine Richtungen, die im wesentlichen erhalten bleiben.

Beispiel. Wir betrachten die Matrix

$$C = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Wie vorher besprochen, müssen wir, um Eigenwerte dieser Matrix zu finden, diejenigen Werte $\lambda \in \mathbb{R}$ bestimmen, für die das zu der Koeffizientenmatrix

$$\begin{pmatrix} 1 - \lambda & -1 \\ 1 & 1 - \lambda \end{pmatrix}.$$

gehörende homogene lineare Gleichungssystem mehr als eine Lösung hat. Wir formen die Matrix wie gewohnt durch elementare Zeilentransformationen um:

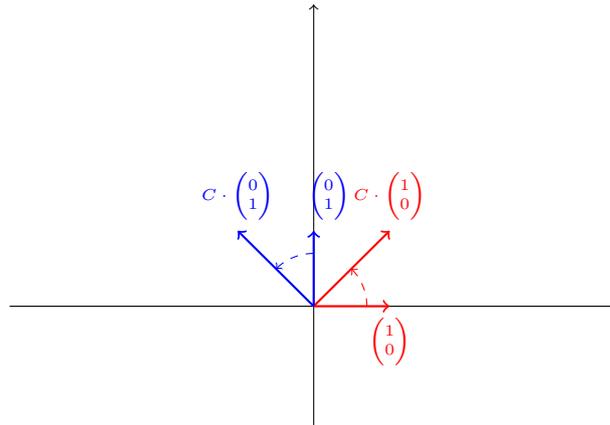
$$\begin{pmatrix} 1 - \lambda & -1 \\ 1 & 1 - \lambda \end{pmatrix} \xrightarrow{I \leftrightarrow II} \begin{pmatrix} 1 & 1 - \lambda \\ 1 - \lambda & -1 \end{pmatrix} \xrightarrow{II - (1 - \lambda)I} \begin{pmatrix} 1 & 1 - \lambda \\ 0 & -1 - (1 - \lambda)^2 \end{pmatrix}.$$

Diese Matrix hat also genau dann nur eine Stufe und das zugehörige lineare Gleichungssystem genau dann mehr als eine Lösung, wenn $-1 - (1 - \lambda)^2 = 0$ gilt. Diese Gleichung ist jedoch äquivalent zu der Gleichung

$$(1 - \lambda)^2 = -1,$$

die für keine reelle Zahl λ erfüllt sein kann, da das Quadrat der reellen Zahl $1 - \lambda$ niemals negativ sein kann. Folglich hat diese Matrix keine Eigenwerte und dementsprechend auch keine Eigenvektoren. Wie angekündigt, kann man informell sagen, dass keine Richtung erhalten bleibt.

Macht man sich klar, wie die Abbildung geometrisch in der Ebene aussieht, so ist es nicht weiter überraschend. Man kann diese Abbildung als Drehstreckung in der Ebene betrachten, also eine Operation, bei der jeder Vektor in der Ebene um 45° gedreht und um den Faktor $\sqrt{2}$ gestreckt wird. Wir zeichnen etwa die Bilder der Standardbasisvektoren ein:



Das exakte Ergebnis entspricht also unserer Anschauung: Bei einer Drehung um 45° zeigt kein Vektor in dieselbe Richtung wie vorher, auch dann nicht, wenn zusätzlich gleichzeitig eine Streckung erfolgt.

Wir haben also zwei extreme Beispiele gesehen: Eine diagonalisierbare Matrix und eine, die gar keine Eigenwerte hat. Nun wollen wir noch ein Beispiel einer Matrix anführen, die zwar Eigenwerte und Eigenvektoren besitzt, allerdings trotzdem nicht diagonalisierbar ist. Informell gesagt, gibt es nicht genug Eigenvektoren, um eine Basis des \mathbb{R}^n , in diesem Fall \mathbb{R}^2 , zu bilden.

Beispiel. Wir betrachten nun die Matrix

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

Wir verfahren wie eben, um die Eigenwerte dieser Matrix zu finden, und betrachten das homogene lineare Gleichungssystem mit Parameter λ und der Koeffizientenmatrix

$$\begin{pmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{pmatrix}.$$

Hier müssen wir zwischen $\lambda = 1$ und $\lambda \neq 1$ unterscheiden. Zunächst behandeln wir den zweiten Fall, denn dieser ist etwas einfacher. Ist $\lambda \neq 1$, so ist $1 - \lambda \neq 0$ und wir können beide Zeilen der Matrix durch diese Zahl teilen:

$$\begin{pmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{pmatrix} \xrightarrow{I/(1-\lambda), II/(1-\lambda)} \begin{pmatrix} 1 & \frac{1}{1-\lambda} \\ 0 & 1 \end{pmatrix}.$$

Diese Matrix hat zwei Spalten und zwei Stufen, also hat das zugehörige homogene lineare Gleichungssystem die einzige Lösung $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Folglich ist keine der reellen Zahlen $\lambda \neq 1$ ein Eigenwert der Matrix A .

Nun müssen wir noch untersuchen, ob $\lambda = 1$ ein Eigenwert der Matrix A ist. Dafür setzen wir den Wert $\lambda = 1$ in das obige homogene lineare Gleichungssystem mit Parameter λ und erhalten die Koeffizientenmatrix

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

Das zugehörige lineare Gleichungssystem hat mehr als eine Lösung. Die Lösungsmenge ist gegeben durch

$$\left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\}.$$

Also ist jeder Vektor der Form $\begin{pmatrix} x \\ 0 \end{pmatrix}$ mit $x \neq 0$ ein Eigenvektor der Matrix A zum Eigenwert 1. Allerdings sind das auch alle Eigenvektoren der Matrix A , und somit lässt sich daraus keine Basis von \mathbb{R}^2 aus Eigenvektoren von A finden, da je zwei Vektoren der Form $\begin{pmatrix} x \\ 0 \end{pmatrix}$ linear abhängig sind.

Es kann also vorkommen, dass eine $n \times n$ -Matrix gar keine Eigenwerte besitzt, oder dass eine $n \times n$ -Matrix zwar Eigenwerte besitzt, aber nicht „genug“ Eigenvektoren, um eine Basis von \mathbb{R}^n zu bilden, oder eine $n \times n$ -Matrix kann n Eigenvektoren haben, die eine Basis von \mathbb{R}^n bilden, und somit diagonalisierbar sein.

8 Skalarprodukt, Abstände und Winkel

In diesem letzten Abschnitt der linearen Algebra geht es um einen Begriff von Abständen und von Winkeln in \mathbb{R}^n . In der Ebene und im Raum haben wir eine klare intuitive Vorstellung von diesen Begriffen, und unser allgemeiner Begriff sollte mit der Anschauung nach Möglichkeit übereinstimmen. Ferner soll er auf beliebige Dimensionen anwendbar sein. Sowohl für die Abstände als auch für die Winkel ist das Skalarprodukt auf \mathbb{R}^n entscheidend. Häufig wird in der Schule das Skalarprodukt auf \mathbb{R}^2 und/oder \mathbb{R}^3 bereits behandelt. Wir fangen damit an, das Skalarprodukt zu definieren.

Definition 8.1. Seien $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ zwei beliebige Vektoren in \mathbb{R}^n . Das

Skalarprodukt von $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ und $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ ist definiert als

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = x_1y_1 + x_2y_2 + \dots + x_ny_n.$$

Es ist wichtig zu bemerken, dass beim Skalarprodukt aus zwei Vektoren als Produkt nicht wieder ein Vektor, sondern eine Zahl, also ein Skalar, gemacht wird. Das geschieht, indem man die ersten, zweiten, usw. bis zu den n -ten Einträgen jeweils paarweise multipliziert und dann die Summe von den so erhaltenen Produkten bildet. Das Skalarprodukt sollte nicht mit Skalarmultiplikation verwechselt werden: Die Skalarmultiplikation ist, informell gesagt, „Streckung“ des Vektors mit einem Vorfaktor, Skalarprodukt bildet aus zwei Vektoren eine Zahl, die, wie wir später sehen werden, etwas über den Winkel zwischen diesen beiden Vektoren aussagt. Es sei außerdem angemerkt, dass in der Literatur auch allgemeinere „Skalarprodukte“ behandelt werden und unser Skalarprodukt dann „Standardskalarprodukt“ genannt wird, weil es das einfachste und verbreitetste Skalarprodukt ist. Wir betrachten jedoch nur dieses und sagen deswegen nur „Skalarprodukt“.

Wir rechnen einige Beispiele für das Skalarprodukt aus.

Beispiel. • Das Skalarprodukt der Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix} \in \mathbb{R}^5$ und $\begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ -2 \end{pmatrix}$ ist

gegeben durch

$$\begin{aligned} \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ -2 \end{pmatrix} \right\rangle &= 1 \cdot (-1) + 2 \cdot 0 + 3 \cdot (-1) + 4 \cdot 0 + 5 \cdot (-2) \\ &= -14. \end{aligned}$$

Im Allgemeinen kann Skalarprodukt zweier Vektoren positiv, negativ oder Null sein.

- Das Skalarprodukt der Vektoren $\begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ -2 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 2 \\ 0 \\ 4 \\ 0 \end{pmatrix}$ ist gegeben durch

$$\left\langle \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 4 \\ 0 \end{pmatrix} \right\rangle = (-1) \cdot 0 + 0 \cdot 2 + (-1) \cdot 0 + 0 \cdot 4 + (-2) \cdot 0 = 0.$$

In diesem Fall ist das Skalarprodukt Null, weil in jeder Komponente entweder in dem ersten oder in dem zweiten Vektor eine Null steht. Das ist jedoch nicht die einzige Möglichkeit dafür, dass das Skalarprodukt

Null wird. So ist etwa das Skalarprodukt von $\begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ -2 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \\ -1 \\ 17 \\ 0 \end{pmatrix}$ auch

Null:

$$\begin{aligned} \left\langle \begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 17 \\ 0 \end{pmatrix} \right\rangle &= (-1) \cdot 1 + 0 \cdot 0 + (-1) \cdot (-1) + 0 \cdot 17 + (-2) \cdot 0 \\ &= 0. \end{aligned}$$

Wie schon erwähnt, ist das Skalarprodukt mit Winkeln in \mathbb{R}^n verbunden. In \mathbb{R}^2 und \mathbb{R}^3 haben wir eine klare Vorstellung davon, was ein Winkel zwischen zwei Vektoren ist. Versucht man, diese Vorstellung zu präzisieren, so stellt sich das als nicht ganz einfach heraus. Die Frage, was ein sinnvoller und präziser Begriff von Winkel ist, stellte sich bereits Euklid, der Autor einer der bekanntesten Abhandlungen über Geometrie aus der Antike. Auch heute gibt es unterschiedliche Möglichkeiten, den Begriff des Winkels genau durch Axiome festzulegen. Wir wollen uns nicht in diese Eigenschaften des Winkels vertiefen, wollen aber anmerken, dass man auf \mathbb{R}^n durch das Skalarprodukt einen sinnvollen Begriff für Winkel bekommt. Am einfachsten ist es für rechte Winkel zu definieren.

Definition 8.2. Zwei Vektoren $u, v \in \mathbb{R}^n$ heißen **orthogonal** (oder senkrecht zueinander), falls ihr Skalarprodukt Null ist, in Formeln: $\langle u, v \rangle = 0$.

Wir haben bereits einige Beispiele von orthogonalen Vektoren gesehen:

$\begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ -2 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 2 \\ 0 \\ 4 \\ 0 \end{pmatrix}$ sind orthogonale Vektoren in \mathbb{R}^5 , aber auch $\begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ -2 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \\ -1 \\ 17 \\ 0 \end{pmatrix}$. Es sei angemerkt, dass daraus nicht folgt, dass $\begin{pmatrix} 0 \\ 2 \\ 0 \\ 4 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \\ -1 \\ 17 \\ 0 \end{pmatrix}$ orthogonal sind. Wir rechnen nach, dass diese Vektoren nicht senkrecht zueinander sind, da ihr Skalarprodukt nicht 0 ist:

$$\left\langle \begin{pmatrix} 0 \\ 2 \\ 0 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 17 \\ 0 \end{pmatrix} \right\rangle = 0 \cdot 1 + 2 \cdot 0 + 0 \cdot (-1) + 4 \cdot 17 + 0 \cdot 0 = 68.$$

Ferner wollen wir anmerken, dass je zwei unterschiedliche Vektoren der Standardbasis orthogonal sind. Etwa für $e_2, e_4 \in \mathbb{R}^4$:

$$\left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle = 0 \cdot 0 + 1 \cdot 0 + 0 \cdot 0 + 0 \cdot 1 = 0.$$

Das verallgemeinert die Tatsache, dass im \mathbb{R}^2 und \mathbb{R}^3 die Standardbasisvektoren paarweise senkrecht zueinander sind.

Bevor wir weiter darauf eingehen, in welchem Sinne das Skalarprodukt ein Maß für die Winkel ist, wollen wir zunächst sehen, dass man aus dem Skalarprodukt ein Maß für die Länge von Vektoren erhält.

Definition 8.3. Sei $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ ein Vektor in \mathbb{R}^n . Die **Norm** (oder Länge) von diesem Vektor ist eine reelle Zahl, die gegeben ist durch

$$\left\| \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right\| = \sqrt{\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right\rangle}.$$

Damit diese Definition Sinn ergibt, darf das Skalarprodukt eines Vektors mit sich selbst nie negativ sein. Das sieht man allerdings sofort, wenn man

das Skalarprodukt in der Definition ausrechnet:

$$\begin{aligned} \left\| \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right\| &= \sqrt{\left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right\rangle} \\ &= \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}. \end{aligned}$$

Die Quadrate reeller Zahlen sind nie negativ, und somit ist auch die Summe $x_1^2 + x_2^2 + \dots + x_n^2$ entweder positiv oder Null. Damit ist die Norm eines Vektors in \mathbb{R}^n stets definiert. Sie ist eine nicht-negative reelle Zahl, also positiv oder 0. In \mathbb{R}^2 bzw. \mathbb{R}^3 ist dieser Längenbegriff häufig schon aus der Schule bekannt.

Mit Hilfe der Norm können wir jetzt einen Zusammenhang zwischen Winkeln und Skalarprodukt feststellen. Dabei ist die folgende Beobachtung wichtig: Sicher ist eine wünschenswerte Eigenschaft des Winkels, dass der Winkel zwischen zwei Vektoren u und v derselbe ist wie zwischen u und $2v$ (oder $3v$ oder allgemein für positive Vielfache von v). Allerdings ist das Skalarprodukt von u und v und das von u und $2v$ im Allgemeinen nicht dasselbe, etwa

$$\begin{aligned} \left\langle \begin{pmatrix} 0 \\ 2 \\ 0 \\ 4 \\ 0 \end{pmatrix}, 2 \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \\ 17 \\ 0 \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} 0 \\ 2 \\ 0 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -2 \\ 34 \\ 0 \end{pmatrix} \right\rangle \\ &= 0 \cdot 2 + 2 \cdot 0 + 0 \cdot (-2) + 4 \cdot 34 + 0 \cdot 0 \\ &= 136, \end{aligned}$$

also doppelt so groß, wie das Skalarprodukt von $\begin{pmatrix} 0 \\ 2 \\ 0 \\ 4 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 1 \\ 0 \\ -1 \\ 17 \\ 0 \end{pmatrix}$. (Dass die

Skalarmultiplikation eines der Vektoren mit 2 das Skalarprodukt doppelt so groß werden lässt, ist übrigens kein Zufall, wie wir später nochmal aufgreifen werden.)

Also muss das Skalarprodukt modifiziert werden, um ein vernünftiges Maß für Winkel zu liefern. Diese Modifikation besteht darin, für zwei Vektoren $u, v \in \mathbb{R}^n$, von denen keiner der Nullvektor ist, die Größe

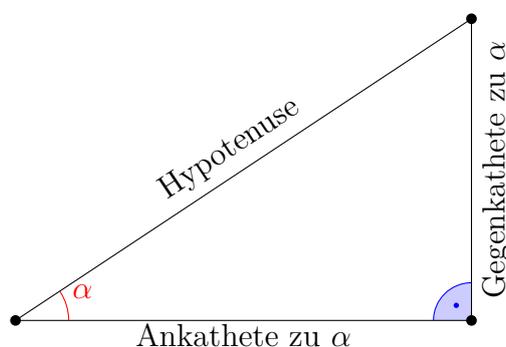
$$\frac{\langle u, v \rangle}{\|u\| \cdot \|v\|}$$

zu betrachten. Es stellt sich als sinnvoll heraus, den Winkel φ zwischen den Vektoren $u, v \in \mathbb{R}^n$, von denen keiner der Nullvektor ist, durch die Identität

$$\cos(\varphi) = \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|}$$

zu definieren.

Wir erinnern uns, dass der Kosinus von einem spitzen Winkel α in einem rechtwinkligen Dreieck als das Verhältnis der Längen der Ankathete und der Hypotenuse definiert ist. Wir erinnern uns daran, dass in einem rechtwinkligen Dreieck die beiden Seiten, die an den rechten Winkel anliegen, *Katheten* genannt werden, und die dem rechten Winkel gegenüberliegende Seite *Hypotenuse* genannt wird. Ist einer der beiden spitzen Winkel im rechtwinkligen Dreieck der Winkel α , so nennt man die Kathete, die an dem Winkel α anliegt, die *Ankathete* von α , und die dem Winkel α gegenüberliegende Kathete die *Gegenkathete* von α . Wir veranschaulichen das nochmal im Bild:



In einem solchen rechtwinkligen Dreieck ist Kosinus von α gegeben durch

$$\cos(\alpha) = \frac{\text{Länge der Ankathete zu } \alpha}{\text{Länge der Hypotenuse}}.$$

Wir werden später nochmal darauf eingehen, wie Kosinus von stumpfen Winkeln definiert ist; zur Erinnerung: Für den rechten Winkel haben wir $\cos(90^\circ) = 0$. Das passt schon mit der Definition der Orthogonalität zusammen.

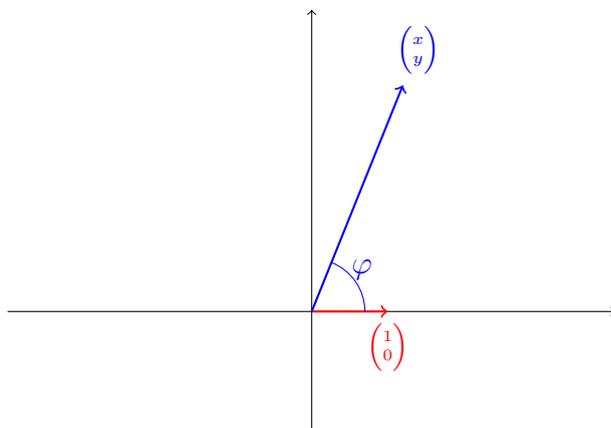
Als nächstes brauchen wir noch, um die obige Definition des Winkels mit unserer Anschauung in einigen Beispielen zu vergleichen, den Satz des Pythagoras, den wir an dieser Stelle nochmal angeben.

Satz 8.4 (Pythagoras). *Seien a, b die Kathetenlängen und c Hypotenusenlänge in einem rechtwinkligen Dreieck. Dann gilt für diese Zahlen:*

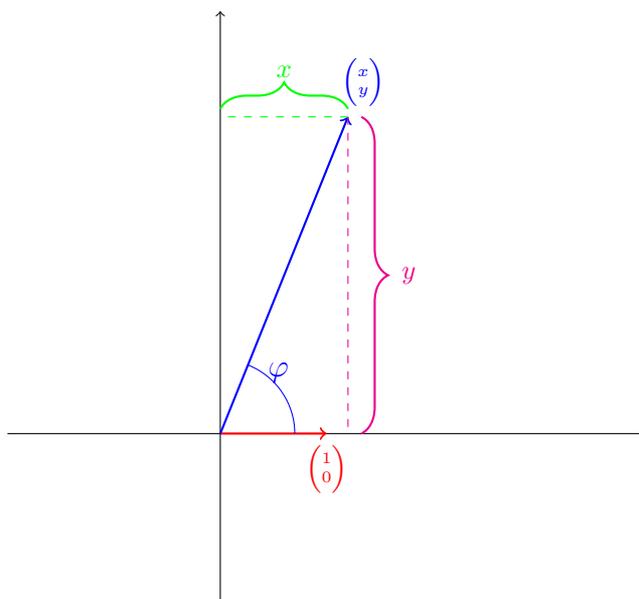
$$a^2 + b^2 = c^2.$$

Nun können wir diese Schulkenntnisse im folgenden Beispiel verwenden.

Beispiel. Wir betrachten den Winkel zwischen dem Vektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und einem anderen Vektor $\begin{pmatrix} x \\ y \end{pmatrix}$ in \mathbb{R}^2 . Damit die Bilder einfacher werden, nehmen wir an, dass der Vektor $\begin{pmatrix} x \\ y \end{pmatrix}$ im ersten Quadranten liegt. Wir erhalten das folgende Bild als Veranschaulichung:



Nun erinnern wir uns, dass wir die Koordinaten x und y des Vektors $\begin{pmatrix} x \\ y \end{pmatrix}$ jeweils als die Längen der Projektionen auf die x - bzw. y -Achse sehen können, wie im Bild angedeutet:



Nun haben wir ein rechtwinkliges Dreieck mit dem Winkel φ darin, der von dem Vektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und dem Vektor $\begin{pmatrix} x \\ y \end{pmatrix}$ eingeschlossen wird. Die Ankathete von φ ist dabei die Länge der Projektion von $\begin{pmatrix} x \\ y \end{pmatrix}$ auf die x -Achse, also gerade x . Die Länge der Hypotenuse lässt sich mit dem Satz von Pythagoras ermitteln. Da in diesem rechtwinkligen Dreieck die beiden Katheten die Längen x und y haben, ist die Hypotenusenlänge gerade $\sqrt{x^2 + y^2}$, was nach unserer Definition auch passenderweise die Länge des Vektors $\begin{pmatrix} x \\ y \end{pmatrix}$ ist. Folglich ist der Kosinus von φ nach der üblichen Schuldefinition in diesem

Fall

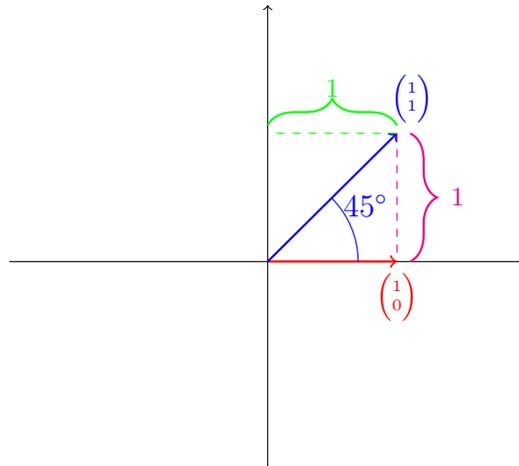
$$\cos(\varphi) = \frac{x}{\sqrt{x^2 + y^2}}.$$

Wir vergleichen das mit dem Ausdruck, den wir mit dem Skalarprodukt erhalten:

$$\begin{aligned} \frac{\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| \cdot \left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|} &= \frac{1 \cdot x + 0 \cdot y}{\sqrt{1^2 + 0^2} \cdot \sqrt{x^2 + y^2}} \\ &= \frac{x}{\sqrt{x^2 + y^2}}. \end{aligned}$$

In diesem Fall, wo wir einen der beiden Vektoren konkret festgelegt haben und den anderen nur im ersten Quadranten betrachtet haben, haben wir also gesehen, dass unser neuer Begriff mit dem alten übereinstimmt.

Wir wollen das für ein konkretes Beispiel im ersten Quadranten nochmal expliziter sehen. Sei $x = y = 1$. Dann ist der Vektor $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ gerade die Diagonale des Quadrates mit Seitenlänge 1, deren zwei Seiten auf den Koordinaten liegen. Diese Diagonale bildet mit der Seite des Quadrats ein 45° -Winkel.



Die Länge der Diagonale ist wieder durch den Satz von Pythagoras in dem rechtwinklig-gleichschenkligen Dreieck durch $\sqrt{1^2 + 1^2} = \sqrt{2}$ gegeben. Also ist

$$\cos(45^\circ) = \frac{1}{\sqrt{2}}.$$

Auf der anderen Seite ist

$$\frac{\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| \cdot \left\| \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\|} = \frac{1 \cdot 1 + 0 \cdot 1}{\sqrt{1^2 + 0^2} \cdot \sqrt{1^2 + 1^2}}$$

$$= \frac{1}{\sqrt{2}},$$

also auch eine Übereinstimmung. (Es sei nochmal benont, dass dies kein allgemeiner Beweis ist, sondern nur eine Veranschaulichung am Beispiel.)

Das Skalarprodukt zweier Vektoren ist also für den Winkel, den sie einschließen, relevant. Während die Größe des Skalarproduktes jedoch nicht gleich sein muss beim selben Winkel, können wir auch ohne den „Korrekturfaktor“ der Normen bereits eine Information über den Winkel erhalten: Ist das Skalarprodukt positiv, so schließen die beiden Vektoren einen spitzen Winkel ein, beim Skalarprodukt Null ist der eingeschlossene Winkel - wie wir bereits gesagt haben - gerade 90° , und beim negativen Skalarprodukt ist der eingeschlossene Winkel stumpf.

Nun wollen wir damit fortfahren, den allgemeinen Begriff der Orthogonalität genauer zu untersuchen. Häufig stellt es sich als hilfreich heraus, mehrere Vektoren zu betrachten, die orthogonal zueinander sind. Dafür ist die folgende Definition wichtig.

Definition 8.5. Eine Ansammlung von Vektoren $v_1, \dots, v_k \in \mathbb{R}^n$ heißt **orthogonales System**, falls alle paarweise Skalarprodukte unterschiedlicher Vektoren aus dieser Ansammlung Null ergeben, also in Formeln: Für alle $1 \leq i, j \leq k$ mit $i \neq j$ gilt:

$$\langle v_i, v_j \rangle = 0.$$

Ein orthogonales System $v_1, \dots, v_n \in \mathbb{R}^n$, das zusätzlich eine Basis von \mathbb{R}^n ist und für alle $1 \leq i \leq n$ die Bedingung $\|v_i\| = 1$ erfüllt, nennt man **Orthonormalbasis** von \mathbb{R}^n .

Wir bemerken, dass eine Orthonormalbasis von \mathbb{R}^n insbesondere eine Basis von \mathbb{R}^n ist, also in jedem Fall genau n Vektoren haben muss. Die Bedingung $\|v_i\| = 1$ für die Länge nennt man auch „Normierung“, worauf sich der „normal“-Teil des Wortes „Orthonormalbasis“ bezieht.

Wir wollen nun einige Eigenschaften der Norm, des Skalarprodukts und der Orthogonalität im folgenden Satz zusammenfassen.

Satz 8.6.

- 1) Sei v ein Vektor in \mathbb{R}^n . Dann ist die Norm von v genau dann Null, wenn v der Nullvektor ist. In Formeln:

$$\|v\| = 0 \Leftrightarrow v = 0.$$

- 2) Sei $w \in \mathbb{R}^n$ fest. Dann ist die Abbildung $\psi: \mathbb{R}^n \rightarrow \mathbb{R}$, die durch $\psi(v) = \langle v, w \rangle$ definiert ist, linear.

3) Ein orthogonales System in \mathbb{R}^n , in dem kein Vektor der Nullvektor ist, ist stets linear unabhängig.

4) Für alle Vektoren $v, w \in \mathbb{R}^n$ gilt: $\langle v, w \rangle = \langle w, v \rangle$.

Bevor wir einige der Behauptungen dieses Satzes beweisen, machen wir noch einige kleine Bemerkungen: Die letzte Eigenschaft wird auch *Symmetrie* des Skalarproduktes genannt. Bei der vorletzten Eigenschaft ist es klar, dass eine Ansammlung von Vektoren, die den Nullvektor enthält, nicht linear unabhängig sein kann. Es lässt sich aber leicht ein orthogonales System angeben, das den Nullvektor enthält. Die Zusatzbedingung im dritten Teil des Satzes ist also notwendig.

Nun geben wir einen Beweis für einige Teile dieses Satzes an.

Beweis. zu 1): Sei $v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ ein beliebiger Vektor. Dann ist die

Norm von diesem Vektor, wie wir bereits gesehen haben,

$$\left\| \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

Also ist $\|v\| = 0$ genau dann erfüllt, wenn $\sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = 0$ ist. Das ist genau dann der Fall, wenn $x_1^2 + x_2^2 + \dots + x_n^2 = 0$ ist. Die Quadrate von reellen Zahlen sind immer nicht-negativ, und somit ist die Summe von einigen Quadraten von reellen Zahlen genau dann Null, wenn jeder der Summanden Null ist. Also ist die Aussage $x_1^2 + x_2^2 + \dots + x_n^2 = 0$ für reelle Zahlen $x_1, \dots, x_n \in \mathbb{R}$ äquivalent zu

$$x_1 = x_2 = \dots = x_n = 0.$$

Dass alle Komponenten vom Vektor v Null sind, ist natürlich äquivalent dazu, dass v der Nullvektor ist. Somit haben wir eine Kette von äquivalenten Aussagen hergestellt, die die Behauptung des ersten Teils beweist.

zu 2): Sei $w = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ ein beliebiger, aber fest gewählter Vektor in \mathbb{R}^n .

Wir „raten“ nun die darstellende Matrix B der Abbildung $\psi: \mathbb{R}^n \rightarrow \mathbb{R}$, die durch $\psi(v) = \langle v, w \rangle$ gegeben ist. Wenn wir dann zeigen, dass $T_B =$

ψ gilt, so wissen wir insbesondere, dass ψ eine lineare Abbildung ist. (Wir erinnern uns daran, dass es nicht reicht, nur die Einheitsvektoren in ψ einzusetzen, um auf die darstellende Matrix zu kommen. Auf diese Weise bekommt man zwar zu jeder Abbildung eine Matrix, aber diese Matrix wird im Allgemeinen die Abbildung nicht beschreiben. Deshalb müssen wir auch noch prüfen, dass die Abbildung ψ denselben Effekt hat wie die Multiplikation mit der Matrix B .)

Als erstes schreiben wir die Abbildung ψ für jedes $v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$

etwas expliziter hin:

$$\begin{aligned} \psi(v) &= \psi \left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right) = \langle v, w \rangle \\ &= \left\langle \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right\rangle \\ &= x_1 y_1 + x_2 y_2 + \dots + x_n y_n. \end{aligned}$$

Dabei müssen wir bedenken, dass y_i die Komponenten des *festen* Vektors w sind und somit als Konstanten oder Parameter betrachtet werden müssen. Hingegen sind die x_i die Komponenten des Vektors v , dessen Wert unter der Abbildung ψ ermittelt wird, diese ändern sich also, sobald man v ändert.

Nun zu der Matrix B , die potentiell die darstellende Matrix von ψ ist. Zunächst machen wir uns klar, welche Größe diese Matrix haben muss. Da ψ eine Abbildung von \mathbb{R}^n nach $\mathbb{R} = \mathbb{R}^1$ ist, wird B eine $1 \times n$ -Matrix: Diese Matrix hat also eine einzige Zeile, in der n Zahlen (also n Spalten der „Höhe 1“) stehen. An der i -ten Stelle in dieser Zeile steht jeweils die reelle Zahl, die dem Standardbasisvektor e_i von \mathbb{R}^n zugeordnet wird. Bildet man das Skalarprodukt $\langle e_i, w \rangle$, so sieht man schnell ein, dass dieses nur einen einzigen Summanden hat, der nicht 0 ist, und das ist $1 \cdot y_i$, sodass

$$\langle e_i, w \rangle = y_i$$

gilt. Also ist die Matrix, die ψ darstellen würde, falls denn ψ eine lineare Abbildung ist, die Matrix mit einer Spalte, in der gerade die Einträge von w stehen, also die Matrix

$$B = (y_1 \quad y_2 \quad \dots \quad y_n).$$

Nun multiplizieren wir diese Matrix mit dem Vektor $v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. Wie gewohnt müssen wir die Einträge jeder Zeile der Matrix (davon gibt es allerdings nur eine) komponentenweise mit den Einträgen des Vektors multipliziert werden und diese Produkte dann addiert werden. Wir erhalten:

$$\begin{pmatrix} y_1 & y_2 & \dots & y_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = y_1x_1 + y_2x_2 + \dots + y_nx_n.$$

Da die Multiplikation von reellen Zahlen kommutativ ist, folgt daraus

nun $B \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \psi \left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right)$ für alle $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$. Die Abbildung ψ

stimmt also mit der Abbildung T_B überein und ist somit linear.

zu 3): Seien Vektoren $v_1, \dots, v_l \in \mathbb{R}^n$ ein orthogonales System derart, dass keiner dieser Vektoren der Nullvektor ist. Wir wollen die Definition der linearen Unabhängigkeit verwenden, um zu beweisen, dass diese Vektoren linear unabhängig sind. Seien also $\mu_1, \dots, \mu_l \in \mathbb{R}$ reelle Zahlen, sodass die Linearkombination der Vektoren v_1, \dots, v_l mit Koeffizienten μ_1, \dots, μ_l Null ergibt, d.h. sodass

$$\mu_1v_1 + \mu_2v_2 + \dots + \mu_lv_l = 0.$$

Wir wollen zeigen, dass dann bereits alle μ_i gleich Null sein müssen.

Nun bemerken wir als erstes, dass das Skalarprodukt von einem beliebigen Vektor mit dem Nullvektor Null ergibt. Das nutzen wir nun aus und bilden die Skalarprodukte $\langle 0, v_i \rangle = 0$ für jedes $1 \leq i \leq l$. Da $\mu_1v_1 + \mu_2v_2 + \dots + \mu_lv_l = 0$ gilt, können wir stattdessen auch schreiben

$$\langle \mu_1v_1 + \mu_2v_2 + \dots + \mu_lv_l, v_i \rangle = 0.$$

Nun wissen wir nach dem vorigen Teil dieses Satzes, dass für ein festes v_i das Skalarprodukt mit v_i linear ist. Folglich können wir die Eigenschaften (A1) und (A2) verwenden, um die Summe im ersten Teil des Skalarprodukts „auseinanderzuziehen“ und erhalten:

$$\begin{aligned} 0 &= \langle \mu_1v_1 + \mu_2v_2 + \dots + \mu_lv_l, v_i \rangle \\ &= \langle \mu_1v_1, v_i \rangle + \langle \mu_2v_2, v_i \rangle + \dots + \langle \mu_lv_l, v_i \rangle \end{aligned}$$

$$= \mu_1 \langle v_1, v_i \rangle + \mu_2 \langle v_2, v_i \rangle + \dots + \mu_l \langle v_l, v_i \rangle.$$

Als nächstes nutzen wir die Tatsache aus, dass die Vektoren v_1, v_2, \dots, v_l ein orthogonales System bilden. Das heißt nach Definition gerade, dass Skalarprodukte von je zwei unterschiedlichen aus diesen Vektoren Null ergeben. Das bedeutet wiederum, dass die meisten Skalarprodukte in der folgenden Summe bereits Null sind, nämlich alle, wo der erste und der zweite Vektor im Skalarprodukt unterschiedlich sind. Es bleibt ein einziger Summand über, und zwar $\mu_i \langle v_i, v_i \rangle$, sodass wir nun erhalten:

$$\mu_i \langle v_i, v_i \rangle = 0.$$

Nun erinnern wir uns, dass nach Definition der Norm $\langle v_i, v_i \rangle = \|v_i\|^2$ gilt. Da v_i nicht der Nullvektor ist, ist nach dem ersten Teil des Theorems auch $\|v_i\|$ nicht Null und somit auch $\langle v_i, v_i \rangle$. Da diese Zahl nicht Null ist, aber multipliziert mit μ_i Null ergibt, muss also μ_i Null sein. Da diese Überlegung für jedes i zwischen 1 und l gleichermaßen galt, folgt also

$$\mu_1 = \mu_2 = \dots = \mu_l = 0.$$

Das ist genau das, was wir beweisen wollten. Also ist jedes orthogonales System, das nicht den Nullvektor enthält, linear unabhängig.

zu 4): Hierzu wollen wir keinen Beweis führen. Es sei jedoch erwähnt, dass die Kommutativität der Multiplikation von reellen Zahlen hier die entscheidende Zutat ist.

□

Korollar 8.7. *Seien v_1, \dots, v_n Vektoren in \mathbb{R}^n , die der Bedingung*

$$\langle v_i, v_j \rangle = 0$$

für alle $1 \leq i < j \leq n$ genügen. Für die Skalarprodukte dieser Vektoren jeweils mit sich selbst gelte außerdem: $\langle v_i, v_i \rangle = 1$ für alle $1 \leq i \leq n$. Dann bilden diese Vektoren eine Orthonormalbasis von \mathbb{R}^n .

Beweis. Wir erinnern uns, dass $v_1, \dots, v_n \in \mathbb{R}^n$ eine Orthonormalbasis von \mathbb{R}^n bilden, falls sie ein orthogonales System sind, das zusätzlich eine Basis von \mathbb{R}^n ist und für alle $1 \leq i \leq n$ die Bedingung $\|v_i\| = 1$ erfüllt,

Zunächst wollen wir uns überlegen, dass die vorgegebenen Vektoren tatsächlich ein orthogonales System in \mathbb{R}^n bilden. Dafür müssen alle Skalarprodukte von Paaren unterschiedlicher Vektoren $\langle v_i, v_j \rangle$ Null sein. Wir wissen nach Voraussetzung, dass ein solches Skalarprodukt Null ist, wenn der erste Index kleiner als der zweite ist. Nun nutzen wir die Symmetrie des Skalarproduktes aus, d.h. den Teil 4) des vorherigen Satzes. Dieser besagt insbesondere,

dass $\langle v_i, v_j \rangle = \langle v_j, v_i \rangle$ für alle $1 \leq i, j \leq n$ gilt. Ist also $i > j$, so ist nach Voraussetzung das Skalarprodukt $\langle v_j, v_i \rangle$ Null und folglich auch

$$\langle v_i, v_j \rangle = \langle v_j, v_i \rangle = 0.$$

Insgesamt folgt $\langle v_i, v_j \rangle = 0$ für alle $1 \leq i, j \leq n$, also bilden v_1, \dots, v_n ein orthogonales System.

Als nächstes wollen wir uns überlegen, dass alle Vektoren Norm 1 haben. Das folgt unmittelbar aus der Voraussetzung $\langle v_i, v_i \rangle = 1$ für alle $1 \leq i \leq n$, da die Norm eines Vektors w nach Definition gerade die Quadratwurzel aus dem Skalarprodukt $\langle w, w \rangle$ von w mit sich selbst ist. Angewandt auf v_i , erhalten wir für alle $1 \leq i \leq n$

$$\|v_i\| = \sqrt{\langle v_i, v_i \rangle} = \sqrt{1} = 1.$$

Zuletzt müssen wir nur noch nachweisen, dass die Vektoren v_1, \dots, v_n eine Basis von \mathbb{R}^n bilden. Aus $\|v_i\| = 1$ für alle $1 \leq i \leq n$ folgt nach Teil 1) des obigen Satzes insbesondere, dass keiner der Vektoren v_1, \dots, v_n der Nullvektor ist. Wir haben also ein orthogonales System aus Vektoren, die alle nicht Null sind, und können Teil 3) des obigen Satzes anwenden, um zu sehen, dass dieses System dann linear unabhängig sein muss.

Nun erinnern wir uns daran, dass ein linear unabhängiges System aus n Vektoren in \mathbb{R}^n automatisch eine Basis von \mathbb{R}^n ist. Folglich sind $v_1, \dots, v_n \in \mathbb{R}^n$ eine Orthonormalbasis von \mathbb{R}^n . \square

Als nächstes wollen wir uns der Frage widmen, wann Abbildungen von \mathbb{R}^n in sich selbst Längen und Winkel erhalten, wie etwa Spiegelungen und Drehungen der Ebene es tun sollten. Insbesondere würden wir gerne an der darstellenden Matrix der Abbildung prüfen können, wie die Abbildungen zum Beispiel die Längen der Vektoren verändert. Dafür brauchen wir zunächst einige Vorbereitungen.

Zunächst führen wir den Begriff „transponieren“ ein, der im Wesentlichen bedeutet, aus den Spalten einer Matrix Zeilen einer neuen Matrix zu machen. Diese Operation stellt sich häufig als hilfreich heraus.

Definition 8.8. Sei A die $k \times n$ -Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \dots & a_{kn} \end{pmatrix}.$$

Die **transponierte Matrix** A^T zu der Matrix A ist eine $n \times k$ -Matrix, die genau die Spalten von A als Zeilen hat, in Formeln:

$$A^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{k1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{kn} \end{pmatrix}.$$

Wir verdeutlichen die Definition nochmal an einem Beispiel.

Beispiel. Wir betrachten die 2×3 -Matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

Die dazu transponierte Matrix ist dann die 3×2 -Matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Nun betrachten wir $n \times n$ -Matrizen, deren Spalten eine Orthonormalbasis von \mathbb{R}^n bilden. Es stellt sich heraus, dass diese Matrizen besondere Eigenschaften haben. Wir haben uns bereits überlegt, dass jede $n \times n$ -Matrix, deren Spalten eine Basis von \mathbb{R}^n bilden, invertierbar ist. Sind die Spalten zusätzlich eine Orthonormalbasis, so lässt sich die inverse Matrix dieser Matrix besonders einfach berechnen. Das ist der Gegenstand des folgenden Satzes.

Satz 8.9. *Sei A eine $n \times n$ -Matrix, deren Spalten eine Orthonormalbasis von \mathbb{R}^n bilden. Dann ist die inverse Matrix zu A gleich der transponierten Matrix zu A , in Formeln $A^{-1} = A^T$.*

Beweisidee. Wir wollen keinen vollständigen Beweis von diesem Satz geben, jedoch eine der Grundideen des Beweises skizzieren. Wir müssten eigentlich, um zu überprüfen, dass die transponierte Matrix A^T tatsächlich invers zu A ist, nachrechnen, dass sowohl $A^T \cdot A = E_n$ als auch $A \cdot A^T = E_n$ gilt. Wir wollen nur die wesentliche Idee von dem Beweis der ersten Identität skizzieren. Dazu schreiben wir A in der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}.$$

Dann ist die Matrix A^T nach Definition gegeben durch

$$A^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{n1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{nn} \end{pmatrix}.$$

Wir betrachten das Produkt dieser Matrizen und insbesondere den Eintrag in der ersten Zeile, ersten Spalte des Produkts etwas genauer: In

$$A^T \cdot A = \begin{pmatrix} a_{11} & a_{21} & a_{31} & \dots & a_{n1} \\ a_{12} & a_{22} & a_{32} & \dots & a_{n2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & a_{3n} & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$$

müssen wir die Einträge der ersten Zeile von A^T mit den jeweils entsprechenden Einträge der ersten Spalte von A multiplizieren und diese Produkte addieren. Die Zeilen von A^T sind jedoch nach Definition gerade die Spalten von A , sodass wir im Eintrag an der Stelle $(1, 1)$ im Produkt erhalten:

$$a_{11}^2 + a_{21}^2 + a_{31}^2 + \dots + a_{n1}^2,$$

was gerade das Skalarprodukt

$$\left\langle \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} \right\rangle$$

der ersten Spalte von A mit sich selbst ist. Da die Spalten von A nach Voraussetzung eine Orthonormalbasis von \mathbb{R}^n bilden, muss dieses Skalarprodukt, das gleichzeitig das Quadrat der Norm der ersten Spalte von A ist, also 1 sein, und ist damit gleich dem ersten Eintrag der Matrix E_n .

Allgemeiner stellt man durch ähnliche Überlegungen fest, dass der Eintrag an der Stelle (i, j) der Produktmatrix $A^T \cdot A$ gerade das Skalarprodukt der i -ten Spalte von A mit der j -ten Spalte von A ist. Dabei ist die Argumentation sehr ähnlich zu der Argumentation im Beweis des Teils 2) vom Satz 8.6. Nach Definition der Orthonormalbasis sind all diese Einträge also Null, außer im Fall $i = j$, also außer auf der Diagonale der Produktmatrix. Dort stehen jeweils die Quadrate der Normen der einzelnen Spalten, die alle Eins sind. Somit würden wir auf diese Weise genau die Einheitsmatrix erhalten. Das soll uns als Beweisidee für diesen Satz genügen. \square

Zu dem obigen Beispiel wollen wir ein Beispiel angeben.

Beispiel. Wir betrachten die Matrix

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

Diese Matrix hat eine Orthonormalbasis des \mathbb{R}^3 als Spalten. Wir wollen das mit dem Kriterium aus dem Korollar 8.7 nachrechnen. Dafür müssen wir also die Skalarprodukte der ersten Spalte mit der ersten, zweiten und dritten Spalte, der zweiten mit der zweiten und dritten und schließlich der dritten nur mit sich selbst ausrechnen. Diese sind nachfolgend aufgelistet:

$$\left\langle \begin{pmatrix} \frac{4}{5} \\ \frac{3}{5} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{4}{5} \\ \frac{3}{5} \\ 0 \end{pmatrix} \right\rangle = \frac{4}{5} \cdot \frac{4}{5} + \frac{3}{5} \cdot \frac{3}{5} + 0 \cdot 0 = \frac{16}{25} + \frac{9}{25} = 1,$$

$$\begin{aligned} \left\langle \begin{pmatrix} 4 \\ 3 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle &= \frac{4}{5} \cdot 0 + \frac{3}{5} \cdot 0 + 0 \cdot 1 = 0, \\ \left\langle \begin{pmatrix} 4 \\ 3 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{3}{5} \\ \frac{4}{5} \\ 0 \end{pmatrix} \right\rangle &= \frac{4}{5} \cdot \left(-\frac{3}{5}\right) + \frac{3}{5} \cdot \frac{4}{5} + 0 \cdot 0 = -\frac{12}{25} + \frac{12}{25} = 0, \\ \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle &= 0 \cdot 0 + 0 \cdot 0 + 1 \cdot 1 = 1, \\ \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -\frac{3}{5} \\ \frac{4}{5} \\ 0 \end{pmatrix} \right\rangle &= 0 \cdot \left(-\frac{3}{5}\right) + 0 \cdot \frac{4}{5} + 1 \cdot 0 = 0, \\ \left\langle \begin{pmatrix} -\frac{3}{5} \\ \frac{4}{5} \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{3}{5} \\ \frac{4}{5} \\ 0 \end{pmatrix} \right\rangle &= -\frac{3}{5} \cdot \left(-\frac{3}{5}\right) + \frac{4}{5} \cdot \frac{4}{5} + 0 \cdot 0 = \frac{9}{25} + \frac{16}{25} = 1. \end{aligned}$$

Also sind die Bedingungen aus dem Korollar 8.7 für die Spalten der Matrix A erfüllt, und somit bilden die Spalten von A eine Orthonormalbasis des \mathbb{R}^3 .

Also lässt sich der obige Satz darauf anwenden, und wir können die inverse Matrix zu A einfach durch Transponieren berechnen. Es folgt also:

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

Wir wollen in diesem konkreten Fall nachrechnen, dass $A \cdot A^T = E_3$ gilt. Das ist zwar dank dem obigen Satz mathematisch nicht notwendig, verleiht uns allerdings ein gewisses Gefühl für die Sache.

$$\begin{aligned} &\begin{pmatrix} \frac{4}{5} & 0 & -\frac{3}{5} \\ \frac{3}{5} & 0 & \frac{4}{5} \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{4}{5} & \frac{3}{5} & 0 \\ 0 & 0 & 1 \\ -\frac{3}{5} & \frac{4}{5} & 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{4}{5} \cdot \frac{4}{5} + 0 \cdot 0 + \left(-\frac{3}{5}\right) \cdot \left(-\frac{3}{5}\right) & \frac{4}{5} \cdot \frac{3}{5} + 0 \cdot 0 + \left(-\frac{3}{5}\right) \cdot \frac{4}{5} & \frac{4}{5} \cdot 0 + 0 \cdot 1 + \left(-\frac{3}{5}\right) \cdot 0 \\ \frac{3}{5} \cdot \frac{4}{5} + 0 \cdot 0 + \frac{4}{5} \cdot \left(-\frac{3}{5}\right) & \frac{3}{5} \cdot \frac{3}{5} + 0 \cdot 0 + \frac{4}{5} \cdot \frac{4}{5} & \frac{3}{5} \cdot 0 + 0 \cdot 1 + \frac{4}{5} \cdot 0 \\ 0 \cdot \frac{4}{5} + 1 \cdot 0 + 0 \cdot \left(-\frac{3}{5}\right) & 0 \cdot \frac{3}{5} + 1 \cdot 0 + 0 \cdot \frac{4}{5} & 0 \cdot 0 + 1 \cdot 1 + 0 \cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

an dieser Stelle haben wir also durch Rechnung im Spezialfall die allgemeine Theorie bestätigt.

Nun wollen wir also lineare Abbildungen im Zusammenhang mit Längen und Winkeln untersuchen. Im Allgemeinen wird eine lineare Abbildung weder

die Längen von Vektoren erhalten noch die Winkel dazwischen. Etwa die Streckung, die wir bereits kennengelernt haben, besteht ja gerade daraus, dass Längen von allen Vektoren verändert werden. Im Allgemeinen werden durch lineare Abbildungen auch die Winkel zwischen den Vektoren „verzerrt“. Wir wollen nun näher Abbildungen untersuchen, die dies nicht tun.

Definition 8.10. Eine lineare Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt **orthogonal**, falls sie Längen und Winkel erhält ist, d.h. für alle Vektoren $v, w \in \mathbb{R}^n$ muss φ erfüllen:

$$\langle \varphi(v), \varphi(w) \rangle = \langle v, w \rangle$$

und $\|\varphi(v)\| = \|v\|$.

Eine $n \times n$ -Matrix heißt **orthogonal**, falls die Spalten dieser Matrix eine Orthonormalbasis von \mathbb{R}^n bilden.

Bei einer orthogonalen Abbildung hat also insbesondere der Bildvektor eines Vektors $v \in \mathbb{R}^n$ dieselbe Länge wie v .

Nun wollen wir erläutern, dass die Begriffe der orthogonalen Matrix und orthogonalen Abbildung sinnvoll gewählt sind, denn sie beschreiben denselben Sachverhalt. Das ist der Inhalt des folgenden Satzes.

Satz 8.11. *Eine lineare Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist genau dann orthogonal, wenn ihre darstellende Matrix orthogonal ist.*

Beweisidee. Die Aussage des Satzes ist eine „genau dann, wenn“ Aussage. Es müssen also gewissermaßen zwei Aussagen gezeigt werden. Zunächst muss gezeigt werden, dass die darstellende Matrix einer orthogonalen Abbildung eine orthogonale Matrix ist. Dann muss auch gezeigt werden, dass eine lineare Abbildung, deren darstellende Matrix orthogonal ist, auch eine orthogonale Abbildung ist. Wir wollen uns an dieser Stelle allerdings nur mit dem ersten Teil der Aussage beschäftigen.

Wir betrachten also eine orthogonale Abbildung $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$, und wollen beweisen, dass ihre darstellende Matrix orthogonal ist. Also müssen wir nach Definition einer orthogonalen Matrix gerade zeigen, dass die Spalten der darstellenden Matrix von φ eine Orthonormalbasis von \mathbb{R}^n bilden. Nun sind die Spalten der darstellenden Matrix einer linearen Abbildung gerade die Bilder der Einheitsvektoren unter dieser Abbildung. Folglich müssen wir zeigen, dass die Vektoren $\varphi(e_1), \varphi(e_2), \dots, \varphi(e_n)$ eine Orthonormalbasis von \mathbb{R}^n bilden.

Dafür beobachten wir zunächst, dass die Standardbasisvektoren in \mathbb{R}^n eine Orthonormalbasis von \mathbb{R}^n bilden, wie sich unschwer nachrechnen lässt. Nun wollen wir das Korollar 8.7 nutzen, um nachzuprüfen dass die Bilder $\varphi(e_1), \varphi(e_2), \dots, \varphi(e_n)$ ebenfalls eine Orthonormalbasis von \mathbb{R}^n sind. Dafür müssen wir zunächst die Skalarprodukte der Form

$$\langle \varphi(e_i), \varphi(e_j) \rangle$$

für $1 \leq i < j \leq n$ bilden. Dabei nutzen wir aus, dass die Abbildung φ orthogonal ist, also das Skalarprodukt zweier Vektoren dasselbe ist wie das Skalarprodukt ihrer Bilder unter φ . In Formeln heißt es in diesem Fall:

$$\langle \varphi(e_i), \varphi(e_j) \rangle = \langle e_i, e_j \rangle.$$

Nun ist das Skalarprodukt zweier unterschiedlicher Standardbasisvektoren, wie wir soeben erwähnt haben, stets gleich Null, sodass wir

$$\langle \varphi(e_i), \varphi(e_j) \rangle = 0$$

für $1 \leq i < j \leq n$ erhalten.

Nun müssen wir außerdem nachprüfen, dass die Skalarprodukte $\langle \varphi(e_i), \varphi(e_i) \rangle$ den Wert 1 für alle $1 \leq i \leq n$ haben. Wir gehen genauso wie eben vor:

$$\langle \varphi(e_i), \varphi(e_i) \rangle = \langle e_i, e_i \rangle$$

für alle $1 \leq i \leq n$. Nun ist das Skalarprodukt jedes Standardbasisvektors in \mathbb{R}^n mit sich selbst gerade 1, sodass wir

$$\langle \varphi(e_i), \varphi(e_i) \rangle = 1$$

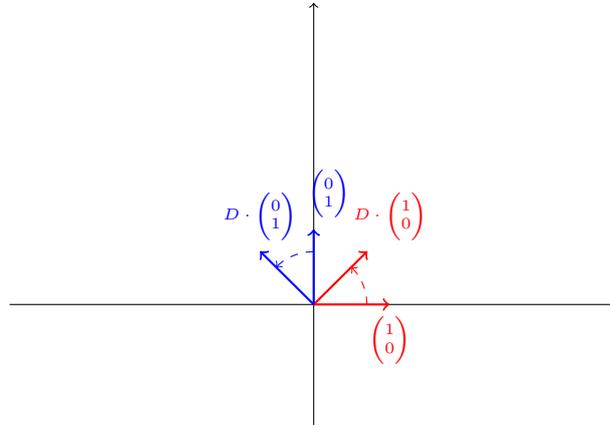
für alle $1 \leq i \leq n$ erhalten.

Damit sind die Voraussetzungen des Korollars 8.7 erfüllt, und die Vektoren $\varphi(e_1), \varphi(e_2), \dots, \varphi(e_n)$ bilden eine Orthonormalbasis von \mathbb{R}^n . Somit ist die darstellende Matrix von φ orthogonal, was zu beweisen war. \square

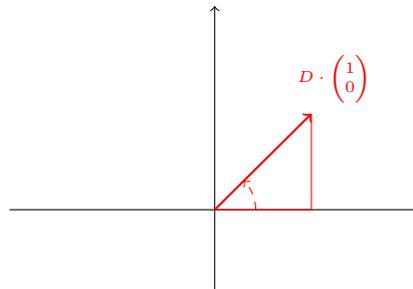
Nun wollen wir ein Beispiel einer orthogonalen Abbildung betrachten.

Beispiel. Wir betrachten eine Drehung der Ebene gegen den Uhrzeigersinn um den Winkel 45° , und wollen eine Begründung dafür skizzieren, warum das eine orthogonale Abbildung ist. Tatsächlich ließe sich das mit jeder Drehung machen, allerdings ist die darstellende Matrix für die Drehung um 45° besonders einfach zu bestimmen.

Wir werden nicht nachweisen, dass diese Drehung eine lineare Abbildung ist. Dafür müsste man sich überlegen, wie die Koordinaten eines beliebigen Vektors $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ nach der Drehung um 45° aussehen. Wir wollen das nur in den Spezialfällen der Standardbasisvektoren von \mathbb{R}^2 tatsächlich durchführen. Wir bezeichnen die darstellende Matrix unserer Drehung mit D .



Da das Bild eines Vektors der Länge 1 nach der Drehung nach wie vor Länge 1 haben muss, hat der Vektor $D \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ die Länge 1. Um dessen Koordinaten zu bestimmen, betrachten wir das rechtwinklige Dreieck, dessen Katheten gerade die Koordinaten des Vektors sind, und der als Hypotenuse den Vektor selbst hat:

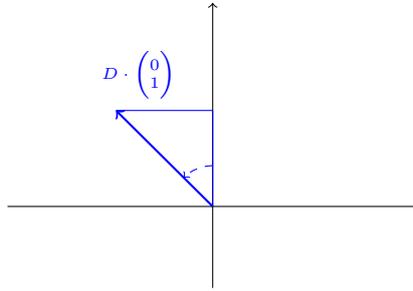


In diesem rechtwinkligen Dreieck sind also die beiden spitzen Winkel gerade 45° , und die Hypotenuse hat die Länge 1. Da beide spitzen Winkel gleich sind, ist unser Dreieck rechtwinklig-gleichschenkelig, also sind die beiden Katheten gleich lang. Die Länge x der Ankathete des 45° -Winkels lässt sich nun mit Hilfe des Kosinus von 45° bestimmt, da in diesem rechtwinkligen Dreieck gilt:

$$\cos(45^\circ) = \frac{x}{1}.$$

Nun haben wir $\cos(45^\circ)$ in einem früheren Beispiel in diesem Abschnitt bereits berechnet und erhielten $\cos(45^\circ) = \frac{1}{\sqrt{2}}$. Also hat das Bild des Vektors $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ unter der Drehung um 45° die Koordinaten $\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$.

Für die Koordinaten des Vektors $D \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ zeichnet man ein ganz ähnliches Dreieck ein, dessen Seitenlängen wieder gleich sind:



Allerdings muss man hier beachten, dass die erste Koordinate negativ ist. Insgesamt erhält man also

$$D \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Die darstellende Matrix D der Drehung um 45° gegen den Uhrzeigersinn in der Ebene ist also (unter der Annahme, dass diese Abbildung linear ist):

$$D = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Wir rechnen nach, dass diese Matrix eine Orthonormalbasis von \mathbb{R}^2 als Spalten hat. dafür muss das Skalarprodukt der beiden Vektoren in den Spalten 0 sein und die Länge dieser Vektoren jeweils 1 sein. Wir rechnen das jeweils nach:

$$\begin{aligned} \left\langle \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right\rangle &= -\frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} = 0 \\ \left\| \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right\| &= \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2} = \sqrt{\frac{1}{2} + \frac{1}{2}} = 1, \\ \left\| \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \right\| &= \sqrt{\left(-\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2} = \sqrt{\frac{1}{2} + \frac{1}{2}} = 1. \end{aligned}$$

Also sagt uns der obige Satz, dass die Abbildung mit der darstellenden Matrix D orthogonal ist, d.h. Längen und Winkel erhält. Die Drehung um 45° erhält Längen und Winkel, also ist unsere Rechnung konsistent mit unserer Vorstellung. (Man merke, dass wir die Längenerhaltung bereits bei der Bestimmung der Matrix D eigentlich benutzt haben. Es ist also kein Beweis, dass die Drehung Längen und Winkel erhält; wir haben uns lediglich veranschaulicht, dass unsere Begriffe zusammenpassen.)

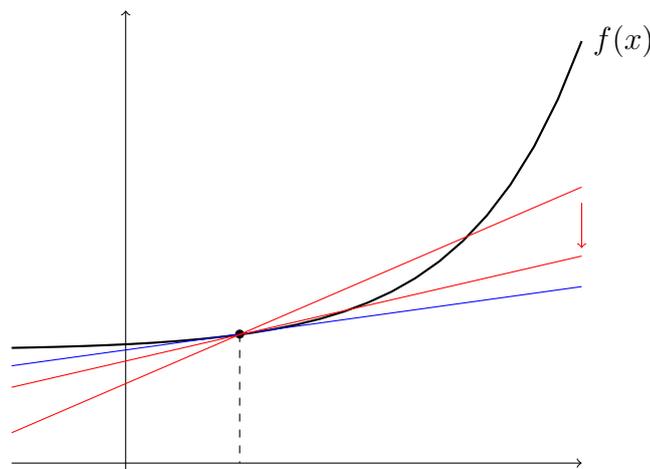
Teil II

Analysis: Differential- und Integralrechnung

Nun kommen wir zu einem kurzen Überblick über einige Themen in der Analysis, wobei unser Schwerpunkt auf der Differential- und Integralrechnung liegen wird. Einige Themen der Differential- und Integralrechnung werden bereits in der Schule behandelt. Wir werden uns deshalb darauf konzentrieren, uns mit Hintergründen der in der Schule bereits erlernten Begriffe wie Ableitung bzw. Rechenregeln dafür auseinanderzusetzen.

Wir werden uns in der Analysis vor allem mit dem Studium von Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ beschäftigen; manchmal wird es auch erforderlich sein, die Definitionsmenge kleiner zu wählen, etwa als ein offenes Intervall I in \mathbb{R} . Im Allgemeinen betrachtet man in der Analysis auch Funktionen in mehreren Variablen; wir beschränken uns jedoch auf den einfachsten Fall von nur einer Variablen.

In der Schule lernt man bereits den Begriff der Ableitung einer solchen Funktion $f: I \rightarrow \mathbb{R}$ in einem Punkt $x_0 \in I$. Die Idee dabei ist, die Steigung der Tangente an den Graphen der Funktion f im Punkt $(x_0, f(x_0))$ zu bestimmen, indem man diese Tangente durch Sekanten annähert. Dabei stellt sich heraus, dass die Steigungen der Sekanten bei den meisten bislang bekannter Funktionen „immer näher“ an die Steigung der Tangente kommen, je näher die Sekante an die Tangente ist. (Das ist im nachfolgenden Bild veranschaulicht. Die roten Geraden sind Sekanten am Funktionsgraphen, die blaue Gerade ist die Tangente in dem markierten Punkt.)



Übrigens kann man die Steigung der Sekante an f als die durchschnittliche Änderungsrate von f zwischen den Schnittpunkten betrachten, während die Tangentensteigung dann die momentane Änderungsrate von f beschreibt.

Das ist etwa für physikalische Funktionen von Bedeutung. Auch historisch geht die Definition der Ableitung auf die Untersuchung derartiger physikalischer Fragestellungen zurück.

Wir wollen diesen Begriff von „immer weiter annähern“ nun präzisieren. Bevor wir das jedoch tun können, sollten wir uns nochmal vor Augen führen, was reelle Zahlen überhaupt sind. Dabei bietet sich die Gelegenheit, eine „Erweiterung“ von reellen Zahlen, nämlich die komplexen Zahlen, kennenzulernen. Das wird der Inhalt des folgenden Kapitels werden.

9 Reelle und komplexe Zahlen

Bevor wir mit den reellen Zahlen anfangen, wollen wir noch etwas weiter ausholen. Die ersten Zahlen, die wir kennenlernen (und auch historisch die ältesten) sind die **natürlichen Zahlen** \mathbb{N} . Man kann sagen, dass die natürlichen Zahlen diejenigen Zahlen sind, die wir zum Zählen verwenden. Etwas präziser wäre es vielleicht zu sagen, dass die natürlichen Zahlen zum Zählen verwendet werden könnten, denn beispielsweise die Zahl 10^{100} wird nicht wirklich zum Zählen benutzt. Das ist natürlich keine präzise mathematische Definition. Allerdings sollte man bedenken, dass wir nicht für jeden Begriff eine präzise mathematische Definition geben können, da wir zunächst einige Grundobjekte brauchen, auf denen wir aufbauen können. (Ähnlich dazu können wir einen Begriff in einer Sprache erst erklären, wenn wir bereits eine bestimmte Menge von Worten dieser Sprache bereits kennen.) Man kann tatsächlich auch natürliche Zahlen durch andere Grundobjekte axiomatisch, also durch Angabe von bestimmten Eigenschaften, beschreiben. Dabei spielt das Induktionsaxiom, das wir in Mathematischen Grundlagen 1 kennengelernt haben, eine zentrale Rolle.

Für unsere Zwecke sollen jedoch die natürlichen Zahlen auch ein Grundobjekt sein, das keiner gänzlich präzisen Definition bedarf. Wir wollen nochmal hervorheben, dass wir mit natürlichen Zahlen rechnen können. Natürliche Zahlen können uneingeschränkt addiert und multipliziert werden, um neue natürliche Zahlen zu erhalten. (Wir merken an, dass hingegen die Differenz oder der Quotient zweier natürlicher Zahlen nicht immer eine natürliche Zahl ist.) Dabei haben diese Verknüpfungen einige Eigenschaften, die wir hier auflisten wollen. Diese besagen im Wesentlichen, dass wir wie gewohnt mit Summen und Produkten rechnen können:

Assoziativität der Addition: Für je drei natürliche Zahlen a, b, c gilt:

$$(a + b) + c = a + (b + c)$$

Kommutativität der Addition: Für je zwei beliebige natürliche Zahlen

a, b gilt:

$$a + b = b + a.$$

Assoziativität der Multiplikation: Für je drei natürliche Zahlen a, b, c gilt:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c).$$

Kommutativität der Addition: Für je zwei beliebige natürliche Zahlen a, b gilt:

$$a \cdot b = b \cdot a.$$

Distributivgesetz: Für je drei natürliche Zahlen a, b, c gilt:

$$a \cdot (b + c) = a \cdot b + a \cdot c.$$

Die Assoziativität besagt jeweils, dass wir in einer Summe (bzw. in einem Produkt) beliebig umklammern können. Die Kommutativität sagt, dass wir die Summanden in einer Summe bzw. die Faktoren in einem Produkt beliebig umordnen können. Das Distributivgesetz erklärt, wie man Klammern auflöst.

Es gibt bei der Multiplikation außerdem noch ein ausgezeichnetes Element, das dadurch auffällt, dass die Multiplikation mit diesem „nichts verändert“. Man sagt auch, 1 sei das neutrale Element der Multiplikation auf den natürlichen Zahlen. Das notieren wir nochmal:

Neutrales Element der Multiplikation: Für jede natürliche Zahl a gilt:

$$1 \cdot a = a \cdot 1 = a.$$

Das neutrale Element der Addition ist keine der oben beschriebenen natürlichen Zahlen, wurde jedoch auch bereits sehr früh eingeführt: die Null. In \mathbb{N}_0 , den natürlichen Zahlen, zu denen wir die Null hinzugefügt haben, gelten alle Rechenregeln, die wir bislang aufgelistet haben. Zusätzlich gilt nun auch:

Neutrales Element der Addition: Für jede Zahl $a \in \mathbb{N}_0$ gilt:

$$0 + a = a + 0 = a.$$

Nun hat man schon sehr früh (sowohl geschichtlich als auch in der Schule) die natürlichen Zahlen mit Null auf zwei Arten erweitert, indem man negative Zahlen und Brüche eingeführt hat.

Die negativen Zahlen sollen etwa Schulden beschreiben, und werden auch notwendig, wenn man die Gleichungen der Form $x + 5 = 2$ lösen können

will. Während die Gleichung $x + 2 = 5$ eine eindeutige Lösung in \mathbb{N}_0 besitzt, hat die Gleichung $x + 5 = 2$ keine Lösung in \mathbb{N}_0 . Also erweitert man den Zahlenbereich und definiert die negativen ganzen Zahlen. Alle *ganzen Zahlen* zusammen, die wir mit \mathbb{Z} bezeichnen, haben wieder eine Addition und Multiplikation, die alle bis hier aufgezählten Eigenschaften besitzen, und zusätzlich noch die Möglichkeit, die Addition einer Zahl wieder durch Addition einer anderen Zahl rückgängig zu machen. Ähnlich wie wir die Abbildung, die den Effekt einer Abbildung „rückgängig macht“, inverse Abbildung zu der ersten nennen, und bei Matrizen die entsprechende Matrix die inverse Matrix nennen, spricht man hier von der Existenz der Inversen Elemente bezüglich der Addition.

Inverse Elemente bezüglich der Addition: Zu jeder ganzen Zahl $a \in \mathbb{Z}$ gibt es eine ganze Zahl $-a \in \mathbb{Z}$, sodass gilt:

$$a + (-a) = (-a) + a = 0.$$

Beispiel. • Zu $5 \in \mathbb{Z}$ ist $-5 \in \mathbb{Z}$ das inverse Element bezüglich der Addition, denn es gilt $5 + (-5) = (-5) + 5 = 0$.

- Zu $-7 \in \mathbb{Z}$ ist $7 \in \mathbb{Z}$ das inverse Element bezüglich der Addition, denn es gilt $(-7) + 7 = 7 + (-7) = 0$.

Jetzt hat man ein ähnliches Problem bezüglich der Multiplikation. Man will beispielsweise Teilen auch in den Fällen beschreiben können, wenn die Aufteilung in ganzen Zahlen „nicht aufgeht“. Beispielsweise will man die Gleichung der Form $2x = 5$ lösen können, was in den ganzen Zahlen nicht möglich ist. Genauso führt man auch hier, gewissermaßen künstlich, inverse Elemente ein. Allerdings tritt hier ein Problem auf, das bereits aus der Schule bekannt ist: Durch Null „darf“ man nicht teilen. Tatsächlich stellt es sich heraus, dass man, wenn man durch Null teilen dürfte, die übrigen bis jetzt aufgezählten Rechenregeln nicht konsistent aufrechterhalten könnte. Deshalb schließt man die Null aus, wenn man Inverse bezüglich der Multiplikation hinzufügt.

Dabei erhält man die *rationalen Zahlen*, die wir üblicherweise mit \mathbb{Q} bezeichnen. Das sind alle Brüche der Form $\frac{p}{q}$, wobei p eine ganze Zahl (also eventuell auch 0 oder negativ) und q eine natürliche Zahl $\neq 0$ ist. Für rationale Zahlen hat man wieder Addition und Multiplikation, die alle bislang aufgelisteten Eigenschaften besitzt. Wir erinnern uns daran, dass man bei der Durchführung der Addition von Brüchen aufpassen muss: Man muss die beiden Summanden auf den gemeinsamen Nenner bringen, der auch der Nenner der Summe sein wird, und addiert nur die Zähler, um den Zähler der Summe zu erhalten. Gleichzeitig ist die Multiplikation von Brüchen einfacher: Man multipliziert einfach sowohl Zähler als auch Nenner der beiden Faktoren, um den Zähler bzw. den Nenner des Produktes zu erhalten. Neben den obigen Eigenschaften hat nun die Multiplikation in den rationalen Zahlen zusätzlich noch die folgende.

Inverse Elemente bezüglich der Multiplikation: Zu jeder rationalen Zahl $a \in \mathbb{Q} \setminus \{0\}$ gibt es eine ganze Zahl $a^{-1} \in \mathbb{Q}$, sodass gilt:

$$a \cdot a^{-1} = a^{-1} \cdot a = 1.$$

Einige Beispiele zur Erinnerung:

Beispiel. • $\left(\frac{3}{4}\right)^{-1} = \frac{4}{3}$,

• $\left(\frac{1}{2}\right)^{-1} = 2$,

• $(-7)^{-1} = -\frac{1}{7}$.

Nun ist es für viele Zwecke üblicher, anstatt mit der Darstellung der rationalen Zahlen als Bruch mit der Dezimaldarstellung dieser zu arbeiten. Wir erinnern uns, dass wir bereits in der Schule ein Kriterium kennengelernt haben, wie man rationale Zahlen an der Dezimaldarstellung erkennt: Das sind diejenige Dezimalzahlen, die entweder nur endlich viele Nachkommastellen haben oder eine ab irgendeiner Stelle periodische Dezimaldarstellung haben. Das erläutern wir nochmal in etwas mehr Detail an einigen Beispielen.

Beispiel. • Die Zahl 27,345 lässt sich als Bruch schreiben, indem wir uns erinnern, dass 0,345 genau 345 Tausendstel sind, sodass gilt:

$$27,345 = 27 + \frac{345}{1000} = \frac{27345}{1000}.$$

(Dieser Bruch kann außerdem noch gekürzt werden.)

Allgemeiner kann man jede Zahl mit endlich vielen, beispielsweise k , Nachkommastellen, als Bruch mit Nenner 10^k schreiben, also etwa bei einer Nachkommastelle reicht der Nenner 10, bei zwei 100, und so weiter.

- Die Zahl $0,\bar{3}$ (sprich: „0 Komma Periode 3“, es ist $0,\bar{3} \approx 0,3333\dots$), also die Zahl, deren Dezimaldarstellung aus einer Null vor dem Komma und an jeder Nachkommastelle eine drei als Ziffer hat, ist rational. Hier wiederholen sich die Nachkommastellen von Anfang an, und zwar hat die Periode die Länge 1. Im Allgemeinen kann es sein, dass sich ein „Block“ aus mehreren Ziffern wiederholt, und im Allgemeinen kann die Wiederholung erst ab einer späteren Nachkommastelle beginnen. $0,\bar{3}$ ist als Bruch gerade $\frac{1}{3}$. Das kann man sich veranschaulichen, indem man sich die schriftliche Division zur Bestimmung der ersten Nachkommastellen ausführt:

$$\begin{array}{r}
 1,000 \quad : \quad 3 = 0,333 \dots \\
 \underline{-9} \\
 10 \\
 \underline{-9} \\
 10 \\
 \underline{-9} \\
 \dots
 \end{array}$$

Man beachte, dass $0,3333\dots$ keine gültige präzise Schreibweise ist, während $0,\overline{3}$ in unserer Konvention alle Nachkommastellen der Zahl festlegt.

- Die Zahl $0,\overline{09}$ (es ist $0,\overline{09} \approx 0,090909\dots$) ist in der Bruchdarstellung gerade $\frac{1}{11}$. Hier hat die Periode, also der Block, der sich immer wiederholt, die Länge 2.
- Bei der Zahl $0,5\overline{3}$ (mit $0,5\overline{3} \approx 0,53333\dots$) fängt die Periode erst ab der zweiten Nachkommastelle an. In der Bruchdarstellung lässt sich $0,5\overline{3}$ als $\frac{8}{15}$ schreiben.

Nun lassen sich die Grundrechenarten in \mathbb{Q} immer ausführen, bis auf die Division durch 0. Allerdings gibt es nach wie vor Gleichungen, die keine Lösungen in den rationalen Zahlen besitzen. Auch hier gibt es ein Beispiel, was oftmals an dieser Stelle in der Schule behandelt wird, und zwar die Gleichung $x^2 = 2$. Diese Gleichung hat in den reellen Zahlen, die wir nun definieren wollen und die eine erneute Erweiterung unseres Zahlenbereichs darstellen, die Lösungen $\sqrt{2}$ und $-\sqrt{2}$, und bereits im Schulunterricht wird meist bewiesen, dass dies keine rationalen Zahlen sind. Wir brauchen also mehr Zahlen als bloß die rationalen Zahlen.

Die reellen Zahlen kann man beispielsweise dadurch definieren, dass wir (fast) alle Dezimalzahlen zulassen. Das wollen wir nochmal genauer formulieren. Es sei angemerkt, dass es vor allem Gewohnheit ist, die uns gerade Dezimaldarstellung verwenden lässt; prinzipiell lässt sich die Darstellung bezüglich beliebiger Basis, etwa die Binärdarstellung, gleichermaßen an dieser Stelle verwenden.

Definition 9.1. Die **reellen Zahlen** \mathbb{R} sind alle Zahlen mit Dezimaldarstellung

$$\pm b_r b_{r-1} \dots b_1 b_0, a_1 a_2 \dots,$$

wobei a_i, b_i Ziffern, d.h. Elemente von $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ sind, und wir die folgenden Einschränkungen haben:

- $b_r \neq 0$, falls $r > 0$,
- Es gibt keine Stelle N , sodass alle a_i für $i > N$ die Ziffer 9 sind,
- die Zahl, die nur aus Nullen besteht, hat kein Vorzeichen.

Die erste Einschränkung kennen wir bereits von der Dezimaldarstellung der ganzen Zahlen: Wir wollen eine eindeutige Darstellung erreichen, und mit 0 als erster Ziffer wäre etwa 12 und 012 dieselbe Zahl. Ähnlich verhält es sich mit der zweiten Bedingung: Man kann sich überlegen, dass die Zahl $0,\bar{9}$, wenn diese als Dezimaldarstellung zulässig wäre, genau den Wert 1 hätte. Das entspricht im Wesentlichen unserer Intuition, denn die Zahlen der Form $0,9999\dots$ kommen immer näher an die 1 dran: Bei 0,999 ist die Abweichung nur 0,001, bei 0,9999 bereits nur 0,0001, und so ist es plausibel, dass $0,\bar{9}$ denselben Wert wie 1 hätte. Diese Uneindeutigkeit wollen wir ausschließen, daher die zweite Bedingung.

Aus der Informatik ist vermutlich eine andere Art der Zahlendarstellung bekannt, nämlich die Gleitkommadarstellung. Diese ist besonders dafür geeignet, sowohl sehr große Zahlen als auch sehr kleine positive Zahlen mit etwa gleicher Genauigkeit im Computer darzustellen. Insbesondere ist die Darstellung der Zahlen in dem üblichen Format im Computer stets endlich. Für die theoretischen Zwecke ist die Unendlichkeit der Dezimaldarstellung von zentraler Bedeutung.

Nun hat auch $\sqrt{2}$ sich als Dezimalzahl darstellen und ist eine reelle Lösung der Gleichung $x^2 = 2$. Der Aufwand wäre natürlich übertrieben, wenn das die einzige Gleichung wäre, die sich zwar in den reellen, aber nicht in den rationalen Zahlen lösen ließe. Tatsächlich lassen sich beispielsweise in den reellen Zahlen die Gleichungen $x^2 = a$ für alle reellen Zahlen $a > 0$ lösen, und auch viel kompliziertere Zahlen produzieren, etwa

$$\sqrt[7]{\sqrt{2 + \sqrt{2}} + \sqrt[4]{7}}.$$

Es gibt übrigens auch reelle Zahlen, die nicht als Lösungen von „einfachen“ Gleichungen entstehen, insbesondere auch nicht als eine Verschachtelung von Wurzeln wie oben entstehen. Die Kreiszahl π ist ein Beispiel solcher Zahl: Es gibt keine Gleichung der Form

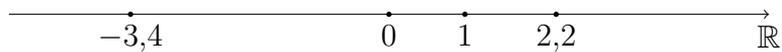
$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0 = 0$$

mit rationalen Koeffizienten $a_n, a_{n-1}, \dots, a_1, a_0 \in \mathbb{Q}$ (und $a_n \neq 0$), deren Lösung die Zahl π ist. Die reellen Zahlen sind bereits sehr vielfältig. Trotzdem wollen wir uns in Kürze einer Erweiterung der reellen Zahlen zuwenden.

Mit den reellen Zahlen können wir wieder die Grundrechenarten wie gewohnt ausführen, und dabei haben die Addition und Multiplikation die früher in diesem Kapitel aufgelisteten Eigenschaften.

Wir wollen außerdem noch anmerken, dass die reelle Zahlen eine weitere wichtige Eigenschaft haben, die uns selbstverständlich erscheint, jedoch von großer Wichtigkeit ist. Auf den reellen Zahlen haben wir die \leq -Ordnung, mit anderen Worten können wir reelle Zahlen sinnvoll vergleichen, und außerdem hat diese Ordnung einige sinnvolle Eigenschaften, die wir stets verwenden, wenn wir mit Ungleichungen arbeiten. Insbesondere lassen sich je zwei reelle Zahlen immer vergleichen: Für zwei reelle Zahlen a, b muss entweder $a < b$ oder $a > b$ oder $a = b$ gelten. Außerdem lässt sich insbesondere mit positiven Zahlen, das heißt mit solchen, die in dieser Ordnung größer als 0 sind, sinnvoll rechnen: Sowohl die Summe als auch das Produkt zweier positiver Zahlen ist stets wieder positiv.

Zuletzt erinnern wir uns daran, dass die reellen Zahlen häufig mit Hilfe der Zahlengerade veranschaulicht werden, etwa wie folgt:



Insbesondere kommt auch die erwähnte Ordnung zum Tragen: Je weiter rechts eine Zahl sich auf der Zahlengerade befindet, desto größer ist sie.

Nun kommen wir zu einer weiteren Art von Gleichungen, die in den reellen Zahlen nach wie vor keine Lösungen besitzt. In den reellen Zahlen ist das Quadrat einer beliebigen Zahl eine nicht-negative Zahl, sodass die Gleichung $x^2 = -1$ keine Lösungen in den reellen Zahlen besitzt. In den komplexen Zahlen fügt man nun künstlich eine neue Zahl hinzu, die man mit i bezeichnet und die die Eigenschaft $i^2 = -1$ hat. Es sei angemerkt, dass i jetzt eine feste Bezeichnung für eine solche Zahl ist und damit nicht mehr als Variable verwendet werden sollte (jedenfalls in den Kontexten, in denen man mit komplexen Zahlen arbeitet).

Man will nun mit dieser neuen Zahl und mit reellen Zahlen genauso rechnen können wie man es früher mit den reellen Zahlen getan hat. Dafür müssen wir noch weitere Zahlen einfach definieren: Etwa für Multiplikation von i mit reellen Zahlen wird man neue Zahlen brauchen, und diese definiert man etwa als $2 \cdot i$, $-3 \cdot i$, $\frac{1}{2} \cdot i$, allgemein als $b \cdot i$ für $b \in \mathbb{R}$. Ferner werden neue Zahlen für Summen dieser Vielfachen von i mit reellen Zahlen benötigt. Es stellt sich heraus, dass diese neuen Zahlen nun für alle Grundrechenarten ausreichen, wie wir in Kürze genauer ausführen werden. Wir beginnen mit einer Definition.

Definition 9.2. Die **komplexen Zahlen** sind alle Zahlen der Form $a + bi$ mit $a, b \in \mathbb{R}$, wobei i eine feste Zahl mit der Eigenschaft $i^2 = -1$ ist. Wir schreiben

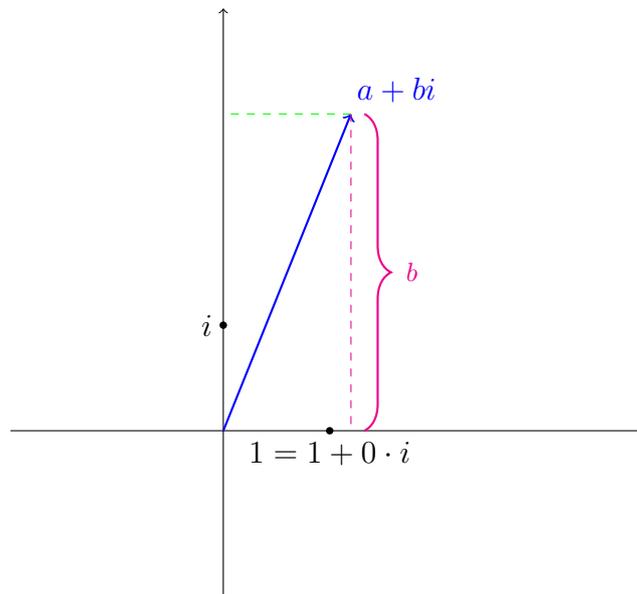
$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}.$$

Beispiel. • $\sqrt{2} + 2 \cdot i$ ist eine komplexe Zahl,

- $7i = 0 + 7 \cdot i$ ist eine komplexe Zahl,

- $\sqrt{2} = \sqrt{2} + 0 \cdot i$ ist eine komplexe Zahl.

Nun zur Veranschaulichung dieser Zahlen: Auf der reellen Gerade ist, informell gesagt, „kein Platz“ für die Zahl i . Es stellt sich jedoch als sinnvoll heraus, die komplexen Zahlen als Punkte in der Ebene oder Vektoren in \mathbb{R}^2 darzustellen, wobei die reellen Zahlen darin auf der x -Achse als die übliche Zahlengerade wiederzufinden sind, und i als der Einheitsvektor in y -Richtung dargestellt wird. Passenderweise ist eine komplexe Zahl $a+bi$ durch die beiden reellen Zahlen $a, b \in \mathbb{R}$ spezifiziert, und auch ein Punkt in der Ebene ist gerade durch seine zwei Koordinaten festgelegt. Die komplexe Zahlenebene sieht also in etwa wie folgt aus:



Ist eine komplexe Zahl $a + bi$ vorgegeben, so nennt man a ihren **Realteil** und b ihren **Imaginärteil**. (Es hat sich eingebürgert, „Realteil“ zu sagen, obwohl man „reelle Zahlen“ sagt.) Für einige Zwecke, etwa für Gleichheit und für Addition komplexer Zahlen, ist es hilfreich, sich diese als Vektoren in \mathbb{R}^2 vorzustellen; für die Multiplikation komplexer Zahlen ist diese Sichtweise jedoch weniger praktisch.

Nun wollen wir uns überlegen, wie man mit den komplexen Zahlen rechnen sollte. Dabei wollen wir unsere Rechenregeln, die wir bereits aufgelistet haben, nach Möglichkeit behalten; das wird die Addition und die Multiplikation bereits festlegen.

Zunächst definieren wir die Addition komplexer Zahlen. Hat man zwei komplexe Zahlen $a_1 + b_1i$ und $a_2 + b_2i$ vorgegeben, und will diese addieren, so erhält man, indem man die gewünschte Assoziativität und Kommutativität der Multiplikation ausnutzt:

$$(a_1 + b_1i) + (a_2 + b_2i) = a_1 + b_1i + a_2 + b_2i = (a_1 + a_2) + (b_1 + b_2)i,$$

und wir konnten die Summe in der Form reelle Zahl + reelle Zahl $\cdot i$ angeben. Wir addieren die komplexen Zahlen also, genau wie Vektoren, komponentenweise: Der Realteil der Summe ist Summe der Realteile der einzelnen Summanden, und der Imaginärteil der Summe ist wiederum die Summe der Imaginärteile der einzelnen Summanden.

Beispiel. Wir berechnen nach der obigen Vorschrift die Summe der komplexen Zahlen $2 + 3i$ und $-1 + (-2)i$.

$$(2 + 3i) + (-1 + (-2)i) = (2 + (-1)) + (3 + (-2))i = 1 + i.$$

Nun kommen wir zu der Multiplikation von komplexen Zahlen. Diese ist etwas komplizierter, lässt sich jedoch leicht herleiten, wenn man bedenkt, dass wir wieder das Distributivgesetz als Rechenregel haben wollen, also dass Klammern auflösen wie früher möglich sein sollte. Wir beginnen mit einem Zahlenbeispiel und verallgemeinern dieses dann.

Beispiel. Wir benutzen das gewünschte Distributivgesetz, um das Produkt der komplexen Zahlen zu bestimmen:

$$(2 + 3i)(-1 - 2i) = 2 \cdot (-1) - 2 \cdot 2i + 3i \cdot (-1) + 3i \cdot (-2i) = -2 - 6i^2 - 7i.$$

Bis hierhin wäre die Rechnung genauso, wenn i einfach eine Variable wäre. Das ändert sich nun, denn wir nutzen jetzt die Definition von i aus und verwenden die Identität $i^2 = -1$. Wir erhalten

$$-2 - 6i^2 - 7i = -2 - 6 \cdot (-1) - 7i = 4 - 7i.$$

Das ist wieder eine komplexe Zahl.

Im allgemeinen Fall können wir genauso verfahren und Distributivgesetz sowie $i^2 = -1$ ausnutzen, um die folgende Definition für das Produkt zweier komplexen Zahlen $a + bi$ und $c + di$ mit beliebigen $a, b, c, d \in \mathbb{R}$ zu erhalten:

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

Man bemerke, dass beim Produkt - ganz anders als bei der Summe - sowohl die Realteile als auch die Imaginärteile der Faktoren den Realteil des Produkts beeinflussen; genauso verhält es sich auch mit dem Imaginärteil des Produktes.

Nun stellt es sich heraus, dass diese Addition und Multiplikation auf den komplexen Zahlen wieder die oben aufgelisteten Eigenschaften hat. Man sollte anmerken, dass jede reelle Zahl in der Form $a + 0 \cdot i$ auch eine komplexe Zahl ist. Insbesondere sind $0 = 0 + 0 \cdot i$ und $1 = 1 + 0 \cdot i$ komplexe Zahlen, die auch für die Addition bzw. Multiplikation in \mathbb{C} die Rolle des jeweiligen neutralen Elements übernehmen. Man kann nachrechnen, dass $0 + z = z + 0 = z$ für jede komplexe Zahl $z \in \mathbb{C}$ gilt (das entspricht auch der Addition

des Nullvektors, den wir bereits als neutrales Element der Addition in \mathbb{R}^2 kennen), und auch, dass $1 \cdot z = z \cdot 1 = z$ für jede komplexe Zahl $z \in \mathbb{C}$ gilt.

Für jedes $a + bi \in \mathbb{C}$ lässt sich auch leicht das inverse Element bezüglich Addition angeben (genauso wie bei Vektoren in \mathbb{R}^2): Man nimmt sowohl für den Realteil als auch für den Imaginärteil das jeweilige inverse Element bezüglich der Addition in den reellen Zahlen, d.h. $(-a) + (-b)i$, und stellt fest, dass

$$(a + bi) + ((-a) + (-b)i) = ((-a) + (-b)i) + (a + bi) = 0$$

gilt.

Die einzige Eigenschaft, die nicht so einfach zu sehen ist, ist die Tatsache, dass jede komplexe Zahl $\neq 0$ auch ein Inverses bezüglich der Multiplikation besitzt. Wir wollen dafür keine allgemeine Formel angeben und stattdessen die Bestimmung eines solchen Inverses in einem Beispiel erläutern; die Methode lässt sich jedoch verallgemeinern. Wir erinnern zunächst an die dritte binomische Formel $(x - y)(x + y) = x^2 - y^2$, die wir gleich verwenden wollen.

Beispiel. Wir wollen $\frac{1}{1+i} \in \mathbb{C}$ bestimmen, also eine komplexe Zahl der Form $a + bi$, die mit $1 + i$ multipliziert genau 1 ergibt. Da wir wieder verlangen, dass man in \mathbb{C} „wie gewohnt“ rechnen kann, ist es insbesondere möglich, diesen Bruch zu erweitern. Wir multiplizieren also den Zähler und den Nenner mit dem Faktor $1 - i$ und erhalten unter Verwendung der dritten binomischen Formel und der definierenden Eigenschaft $i^2 = -1$:

$$\frac{1}{1+i} = \frac{1-i}{(1+i)(1-i)} = \frac{1-i}{1^2 - i^2} = \frac{1-i}{1 - (-1)} = \frac{1-i}{2} = \frac{1}{2} - \frac{1}{2}i.$$

Wir machen nochmal die Probe, um zu sehen, dass diese Zahl tatsächlich die gewünschte Eigenschaft hat (was mathematisch nicht notwendig ist und nur ein Gefühl für die Sache vermitteln soll). Dabei nutzen wir erneut das Distributivgesetz und die definierende Eigenschaft $i^2 = -1$.

$$(1+i) \left(\frac{1}{2} - \frac{1}{2}i \right) = \frac{1}{2} + \frac{1}{2}i - \frac{1}{2}i - \frac{1}{2}i^2 = \frac{1}{2} + \frac{1}{2} = 1.$$

Diese Vorüberlegung liefert die wesentlichen Beweisideen zu dem folgenden Satz:

Satz 9.3. *Die Addition und die Multiplikation auf den komplexen Zahlen sind jeweils kommutativ und assoziativ und erfüllen das Distributivgesetz. Die komplexe Zahl 0 ist neutrales Element der Addition, die komplexe Zahl 1 ist das neutrale Element der Multiplikation. Zu jeder komplexen Zahl $a + bi \in \mathbb{C}$ gibt es ein Inverses bezüglich der Addition, das durch $(-a) + (-b)i$ gegeben ist.*

Ferner existiert zu jeder komplexen Zahl $a + bi$, die nicht Null ist, ein Inverses bezüglich Multiplikation, d.h. eine komplexe Zahl $c + di$, sodass gilt:

$$(a + bi)(c + di) = (c + di)(a + bi) = 1.$$

Es gibt jedoch eine wichtige Eigenschaft, die wir vorher bei den reellen Zahlen hervorgehoben haben und die die komplexen Zahlen nicht haben: Es geht hierbei um einen sinnvollen Begriff der „Positivität“ bzw. der „Ordnung“. Sicherlich gibt es viele Möglichkeiten, irgendeine Ordnung auf komplexen Zahlen zu definieren, oder irgendwelche Zahlen einfach als positiv zu bezeichnen. Jedoch kann man die Ordnung nicht so gestalten, dass unsere üblichen Rechenregeln für Ungleichungen erhalten bleiben. Einige naheliegende Versuche lassen sich sehr schnell widerlegen. Zunächst könnte man etwa auf die Idee kommen, diejenigen komplexen Zahlen als positiv auszuzeichnen, die einen positiven Realteil haben; jedoch wäre mit dieser Definition das Produkt zweier positiven komplexen Zahlen nicht immer positiv. Genauso wenig funktioniert es mit dem positiven Imaginärteil oder mit der lexikographischen Ordnung. Man kann dieser Überlegung eine präzise mathematische Gestalt geben, indem man formuliert, welche Eigenschaften eine „sinnvolle“ Ordnung bzw. ein sinnvoller Begriff der Positivität haben müsste (wie etwa: Summe und Produkt zweier positiven Zahlen ist wieder positiv). Dann lässt sich beweisen, dass eine Ordnung mit diesen Eigenschaften auf \mathbb{C} nicht möglich ist.

Das führt zusätzlich zu dem folgenden Problem. In den reellen Zahlen ist $\sqrt{2}$ dadurch definiert, dass sie die Gleichung $x^2 = 2$ löst. Allerdings ist $\sqrt{2}$ nicht die einzige Zahl, die der Gleichung $x^2 = 2$ genügt: Wie meist bei quadratischen Gleichungen, hat $x^2 = 2$ zwei Lösungen, in diesem Fall $\sqrt{2}$ und $-\sqrt{2}$. In den reellen Zahlen ist jedoch $\sqrt{2}$ eindeutig dadurch festgelegt, dass es die einzige *positive* Lösung der Gleichung $x^2 = 2$ ist. Allgemeiner ist für jede positive reelle Zahl $a \in \mathbb{R}$ die Zahl \sqrt{a} als die eindeutige *positive* reelle Zahl definiert, die der Gleichung $x^2 = a$ genügt. Eine solche Definition der Wurzel ist in \mathbb{C} nicht möglich, jedenfalls nicht in einer konsistenten Art und Weise. Zwar kann man zum Beispiel bei jeder konkreten quadratischen Gleichung eine Lösung festlegen, die man bevorzugt, jedoch wird das im Allgemeinen nicht wie bei reellen Zahlen zu einem konsistenten Begriff. Man sollte also beachten, dass die komplexen Zahlen in vielerlei Hinsicht sehr praktisch und hilfreich sind, man allerdings im Vergleich zu reellen Zahlen auch Eigenschaften verloren hat, die beim Arbeiten mit den reellen Zahlen durchaus wichtig sind.

Als nächstes wollen wir uns mit der Frage beschäftigen, welche weitere Gleichungen durch Einführung der komplexen Zahlen lösbar wurden, die in den reellen Zahlen nicht lösbar waren. Einige dieser Gleichungen sind recht offensichtlich, wie etwa im folgenden Beispiel.

Beispiel. Die Gleichung $x^2 = -4$ hat keine Lösungen in reellen Zahlen, aber zwei Lösungen in komplexen Zahlen, nämlich $2i$ und $-2i$.

Es ist auch leicht zu sehen, dass nun jede quadratische Gleichung mit reellen Koeffizienten mindestens eine Lösung hat, da die $p - q$ -Formel, in der

unter Umständen Wurzeln aus negativen Zahlen vorkommen, nun stets ein Ergebnis liefert.

Um eine allgemeinere Aussage über Lösbarkeit von Gleichungen über \mathbb{C} zu formulieren, brauchen wir noch einige Begriffe. Meist ist der Begriff eines Polynoms bereits aus der Schule bekannt; wir wollen diesen nochmal kurz wiederholen.

Definition 9.4. Ein **Polynom** in der Variablen x mit reellen (bzw. komplexen) Koeffizienten ist ein Ausdruck der Form

$$a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0,$$

wobei $n \in \mathbb{N}_0$ und $a_n, a_{n-1}, \dots, a_2, a_1, a_0$ beliebige, aber feste reelle (bzw. komplexe) Zahlen sind, und man noch $a_n \neq 0$ fordert. (Die Zahlen $a_n, a_{n-1}, \dots, a_2, a_1, a_0$ werden *Koeffizienten* des Polynoms genannt.)

Man nennt die Zahl n auch den **Grad** des Polynoms.

Man nennt eine reelle (bzw. komplexe) Zahl b eine **reelle (bzw. komplexe) Nullstelle** des Polynoms, falls

$$a_n b^n + a_{n-1} b^{n-1} + \dots + a_2 b^2 + a_1 b + a_0 = 0$$

gilt.

Ähnlich wie bei der Dezimaldarstellung einer ganzen Zahlen wollen wir auch bei einem Polynom nicht, dass es mit einer 0 anfängt. Hingegen können die anderen Koeffizienten nach Belieben Null sein; in diesem Fall schreibt man den entsprechenden Term meist nicht auf. Im Prinzip kann man bei Polynomen beliebige Einschränkungen für die Koeffizienten oder für die Nullstellen, die man sucht, fordern: Etwa kann man verlangen, dass die Koeffizienten nur rationale oder ganze oder gerade ganze Zahlen sind, und man nur positiven rationalen Nullstellen suchen will. Für uns werden jedoch nur reelle bzw. komplexe Koeffizienten eine wichtige Rolle spielen.

Man beachte, dass ein Polynom zunächst nur ein Ausdruck in einer Variablen ist, man aber durch Einsetzen einer Zahl für die Variable x auch wieder eine Zahl bekommt. Ist das Ergebnis Null, so besagt die Definition, dass die Zahl, die man eingesetzt hat, Nullstelle des Polynoms genannt wird. Zeichnet man den Graphen der Polynomfunktion auf, so sind Nullstellen genau die Schnittpunkte des Graphen mit der x -Achse.

Wir betrachten nun einige Beispiele von Polynomen.

Beispiel. • Der Ausdruck $(1 + i)x + (2 + 7i)$ ist ein Polynom mit komplexen Koeffizienten vom Grad 1.

- Der Ausdruck $x^2 - 2$ ist ein Polynom mit reellen Koeffizienten und zwei reellen Nullstellen $\sqrt{2}, -\sqrt{2}$.

- Der Ausdruck $x^2 + 1$ ist ein Polynom mit reellen Koeffizienten und ohne reelle Nullstellen. Setzt man eine beliebige reelle Zahl b ein, so ist b^2 nicht-negativ und somit $b^2 + 1$ immer positiv. In den komplexen Zahlen hat dieses Polynom hingegen zwei Nullstellen, nämlich i und $-i$.

Im letzten Beispiel haben wir nochmal gesehen, dass manche Polynome über reellen Zahlen keine reelle Nullstellen besitzen. Der folgende zentrale Satz besagt, dass so was über komplexen Zahlen nicht passieren kann. Dieser Satz ist zwar leicht zu formulieren, ist jedoch nicht ganz einfach zu beweisen.

Satz 9.5 (Fundamentalsatz der Algebra). *Jedes Polynom vom Grad $n \geq 1$ mit komplexen Koeffizienten hat mindestens eine und höchstens n komplexe Nullstellen.*

Die Einschränkung an den Grad ist nötig, um Polynome vom Grad 0, die einfach nur Konstanten sind, auszuschließen. Dass Nullstellen im Prinzip existieren, gibt keinerlei Auskunft darüber, wo sie sind und wie man sie ausrechnen kann. Trotzdem ist dieser Satz von großer Wichtigkeit.

Wir wollen uns in einem konkreten Beispiel die Nullstellen von einem Polynom mit komplexen Koeffizienten bestimmen. Unser Polynom wird eine ganz einfache Form haben, sodass man informell sagen kann, wir würden „komplexe Wurzeln“ ziehen (eine präzise Formulierung von diesem Begriff ist jedoch, wie bereits erläutert, problematisch.)

Beispiel. Wir betrachten das Polynom $P(x) = x^2 - (3 - 4i)$ und suchen die komplexen Nullstellen von diesem. Zunächst bemerken wir, dass die dadurch entstehende Gleichung wie folgt äquivalent umgeformt werden kann.

$$x^2 - (3 - 4i) = 0 \Leftrightarrow x^2 = 3 - 4i.$$

Wir suchen also nach allen komplexen Zahlen x , die dieser Gleichung genügen. Dafür gibt es zunächst keine feste Methode. Wir wissen allerdings, dass jede komplexe Zahl von der Form $x = a + bi$ mit *reellen* Zahlen a, b ist. Wir wollen das in die obige Gleichung einsetzen, um nach einer weiteren Überlegung zu einer Gleichung in den reellen Zahlen zu kommen, denn Lösen von Gleichungen in den reellen Zahlen ist und insgesamt vertrauter. Wir setzen ein und nutzen die erste binomische Formel, um die Klammer aufzulösen:

$$\begin{aligned} (a + bi)^2 = 3 - 4i &\Leftrightarrow a^2 + 2abi + (bi)^2 = 3 - 4i \\ &\Leftrightarrow a^2 - b^2 + 2abi = 3 - 4i. \end{aligned}$$

Im zweiten Schritt haben wir hierbei benutzt, dass wir das Produkt $(bi)^2 = b \cdot i \cdot b \cdot i$ durch Vertauschungen zu $(bi)^2 = b^2 \cdot i^2$ umsortieren dürfen, und dass $i^2 = -1$ nach Definition von i gilt.

Nun machen wir uns klar, dass zwei komplexe Zahlen genau dann gleich sind, wenn sowohl ihre Realteile als auch ihre Imaginärteile jeweils übereinstimmen (genau wie bei Vektoren zwei Vektoren gleich sind, wenn die jeweiligen Komponenten gleich sind). Verwendet man dies in diesem Fall, so kann

man aus der obigen Gleichung ein Gleichungssystem mit zwei Gleichungen erhalten.

$$a^2 - b^2 + 2abi = 3 - 4i \Leftrightarrow \begin{cases} a^2 - b^2 = 3, \\ 2ab = -4. \end{cases}$$

Nun haben wir ein Gleichungssystem mit zwei Gleichungen in zwei Variablen. Allerdings bemerken wir sofort, dass das Gleichungssystem nicht linear ist, da beispielsweise Quadrate der Variablen darin vorkommen. Die bislang erlernten Methoden zum Lösen linearer Gleichungssysteme können wir an dieser Stelle nicht anwenden. Stattdessen werden wir die zweite Gleichung nach b auflösen. Dafür müssten wir die zweite Gleichung durch $2a$ teilen. Wann immer man eine Gleichung durch einen Ausdruck teilen will, insbesondere wenn Variablen darin vorkommen, muss man sich klarmachen, dass dieser Ausdruck (bei beliebigen zulässigen Werten von Variablen) nicht 0 sein kann. In diesem Fall allerdings ist es einfach zu sehen, dass $a \neq 0$ und somit auch $2a \neq 0$ ist: Sonst wäre $a \cdot 2b = 2ab$ ebenfalls 0 und könnte nicht gleichzeitig -4 sein. Folglich können wir durch $2a$ teilen und dann den so erreichten Ausdruck für b in die erste Gleichung einsetzen und diese umformen. Dabei bemerken wir auch, dass auch $a^2 \neq 0$ gelten muss, und damit zu multiplizieren eine Äquivalenzumformung ist, die wir auch verwenden.

$$\begin{aligned} \begin{cases} a^2 - b^2 = 3, \\ 2ab = -4 \mid : (2a) \end{cases} &\Leftrightarrow \begin{cases} a^2 - b^2 = 3, \\ b = -\frac{2}{a} \end{cases} \\ &\Leftrightarrow \begin{cases} a^2 - \left(-\frac{2}{a}\right)^2 = 3, \\ b = -\frac{2}{a} \end{cases} \\ &\Leftrightarrow \begin{cases} a^2 - \frac{4}{a^2} = 3 \mid \cdot a^2, \\ b = -\frac{2}{a} \end{cases} \\ &\Leftrightarrow \begin{cases} a^2 - \frac{4}{a^2} = 3 \mid \cdot a^2, \\ b = -\frac{2}{a} \end{cases} \\ &\Leftrightarrow \begin{cases} a^4 - 4 = 3a^2 \mid - 3a^2, \\ b = -\frac{2}{a} \end{cases} \\ &\Leftrightarrow \begin{cases} a^4 - 3a^2 - 4 = 0, \\ b = -\frac{2}{a} \end{cases}. \end{aligned}$$

Nun sind wir bereits einen Schritt weiter, denn in der ersten Gleichung haben wir nun eine Gleichung in nur einer Variable, und suchen nur reelle

Zahlen a , die dieser Gleichung genügen. Allerdings ist diese Gleichung durch ein Polynom vierten Grades gegeben. Im Allgemeinen gibt es für solche Lösungsformel, doch diese sind sehr kompliziert. Glücklicherweise ist die vorliegende Gleichung vierten Grades eine von einer besonders einfachen Sorte, denn sowohl Koeffizient von a^3 als auch von a sind Null. Solche Gleichungen vierten Grades nennt man auch *biquadratisch*, informell gesagt, weil man sie lösen kann, indem man zweimal eine quadratische Gleichung löst. Dabei ist es etwas einfacher, wenn man eine zusätzliche Variable $c = a^2$ einführt. Dann können wir anstelle von a^4 auch $c^2 = (a^2)^2 = a^4$ schreiben. Dadurch erhalten wir, wie wir gleich sehen, eine quadratische Gleichung, die wir wie gewohnt mit der $p - q$ -Formel lösen können.

$$\begin{aligned}
 \begin{cases} a^4 - 3a^2 - 4 = 0, \\ b = -\frac{2}{a} \end{cases} & \Leftrightarrow \begin{cases} c^2 - 3c - 4 = 0 \\ c = a^2, \\ b = -\frac{2}{a} \end{cases} \\
 & \stackrel{p-q\text{-Formel}}{\Leftrightarrow} \begin{cases} \left(c = \frac{3}{2} + \sqrt{\frac{9}{4} + 4} \text{ oder } c = \frac{3}{2} - \sqrt{\frac{9}{4} + 4} \right), \\ c = a^2, \\ b = -\frac{2}{a} \end{cases} \\
 & \Leftrightarrow \begin{cases} \left(c = \frac{3}{2} + \frac{5}{2} \text{ oder } c = \frac{3}{2} - \frac{5}{2} \right), \\ c = a^2, \\ b = -\frac{2}{a} \end{cases} \\
 & \Leftrightarrow \begin{cases} (c = 4 \text{ oder } c = -1), \\ c = a^2, \\ b = -\frac{2}{a} \end{cases}
 \end{aligned}$$

Nun erinnern wir uns, dass obwohl unser ursprüngliches Problem mit komplexen Zahlen gearbeitet hat, die Zahlen a, b reell sein müssen. Folglich kann das Quadrat $c = a^2$ nicht negativ sein, womit die Möglichkeit $c = -1$ ausgeschlossen ist. Also können wir diese Lösung streichen und dann c wieder eliminieren. Wir formen danach das Gleichungssystem so weit um, bis wir die Lösungen haben.

$$\begin{cases} (c = 4 \text{ oder } c = -1), \\ c = a^2, \\ b = -\frac{2}{a} \end{cases} \stackrel{a \in \mathbb{R}}{\Leftrightarrow} \begin{cases} c = 4, \\ c = a^2, \\ b = -\frac{2}{a} \end{cases}$$

$$\begin{aligned} &\Leftrightarrow \begin{cases} a^2 = 4, \\ b = -\frac{2}{a} \end{cases} \\ &\Leftrightarrow \begin{cases} (a = 2 \text{ oder } a = -2), \\ b = -\frac{2}{a} \end{cases} \\ &\Leftrightarrow \left(\begin{cases} a = 2, \\ b = -1 \end{cases} \text{ oder } \begin{cases} a = -2, \\ b = 1 \end{cases} \right). \end{aligned}$$

Kommt man zum ursprünglichen Problem zurück, so haben wir auf diese Weise gezeigt, dass das Polynom $P(x) = x^2 - (3 - 4i)$ genau zwei komplexe Nullstellen besitzt, und zwar $2 - i$ und $-2 + i$, denn a und b aus dem Gleichungssystem sind ja genau Real- bzw. Imaginärteile der Nullstellen. In Formeln lässt sich die Aussage, die wir gezeigt haben, wie folgt formulieren:

$$\{x \in \mathbb{C} \mid x^2 = 3 - 4i\} = \{2 - i, -2 + i\}.$$

Es sei angemerkt, dass man in diesem Fall gerade die im Fundamentalsatz der Algebra erwähnte Maximalanzahl der Nullstellen für ein Polynom zweiten Grades erreicht, nämlich zwei. Es sei ferner nochmal erwähnt, dass keine dieser Nullstellen als „positiv“ zu betrachten ist, da Positivität auf komplexen Zahlen kein sinnvoller Begriff ist.

10 Folgen, Reihen und Grenzwerte

Nach diesem Ausblick auf die komplexen Zahlen, über die natürlich noch viel mehr gesagt werden könnte, wenden wir uns jetzt den Grenzwerten zu. Folgen und Reihen werden unter anderem die einfachsten Fälle sein, an denen wir Grenzwerte betrachten können; diese definieren wir in Kürze.

Zunächst wollen wir uns nochmal die Notwendigkeit eines Grenzwertbegriffs vor Augen führen. Zum einen wollen wir eine präzise Definition der Ableitung geben, in der, wie wir bereits erläutert haben, die Tangentensteigung durch die Sekantensteigungen „annähern“. Diese Annäherung wollen wir präzise mathematisch formulieren können.

Ein weiteres Problem besteht in der Frage, was es heißt, eine reelle Zahl zu berechnen. Bereits an dem einfachen Beispiel von $\sqrt{2}$, also der eindeutigen positiven Lösung der Gleichung $x^2 = 2$, stellt sich die Frage, wie man die Dezimaldarstellung dieser Zahl bestimmt. Hier ist es nicht schwierig, ein Verfahren zu finden, mit dem man an die ersten Nachkommastellen drankommt (wobei die hier vorgestellte Methode bei weitem nicht die effektivste ist, allerdings recht einfach zu verstehen): Zunächst merkt man, dass eine solche Zahl irgendwo zwischen 1 und 2 liegen muss, da für positive reelle Zahlen $a, b \in \mathbb{R}$, $a, b > 0$ aus $a < b$ auch $a^2 < b^2$ folgt und

$$1^2 < x^2 = 2 < 2^2 = 4$$

gilt. Damit muss die Dezimaldarstellung vor dem Komma mit 1,?? beginnen. Dann schaut man sich die Quadrate der Zahlen 1,1; 1,2; ... an und stellt fest, dass

$$1,4^2 = 1,96 < 2 < 1,5^2 = 2,25$$

ist, sodass $\sqrt{2}$ mit 1,4 anfangen muss. Während man mit diesem langwierigen Verfahren jede beliebige Stelle von $\sqrt{2}$ im Prinzip ausrechnen kann, müssen wir auch sagen, was es insgesamt heißt, $\sqrt{2}$, die ja unendlich viele Stellen hat, anzugeben. Insbesondere wissen wir, dass die Nachkommastellen nicht periodisch werden, also kann man nicht an einer endlichen Stelle sagen, dass ab da sich ein Ziffernblock immer wiederholt, was bei rationalen Zahlen immer möglich ist. Darüber müssen wir uns Gedanken machen.

Noch schwieriger wird es beispielsweise mit der Kreiszahl π . Diese ist definiert als das Verhältnis von der Länge eines Kreises zu seinem Durchmesser, und dieses Verhältnis ist (etwas überraschend) stets gleich. Eine gewisse Näherung kann man durch Messungen erhalten, aber insgesamt ist es zunächst gar nicht klar, wie man die Zahl π auf einige Nachkommastellen annähert, oder was es heißt, π auszurechnen. Auch diese Frage müssen wir uns stellen.

Wir wollen also den Begriff des Grenzwertes präzisieren. Bevor wir damit anfangen, führen wir noch einige einfache Objekte ein, die dafür hilfreich sein werden. Wir fangen mit dem Begriff der Folge an. Folgen kann man sich in etwa wie unendliche Arrays mit reellen Zahlen vorstellen. Nun etwas genauer:

Definition 10.1. Eine **Folge** ist eine Ansammlung $(a_n)_{n \in \mathbb{N}}$ (oder $(a_n)_{n \in \mathbb{N}_0}$) von reellen Zahlen, die durch natürliche Zahlen durchnummeriert sind. Die jeweiligen Zahlen a_n heißen **Folgenglieder** der Folge.

Dieses Konzept ist nicht besonders schwierig. Wir erläutern es nochmal an einigen Beispielen.

Beispiel. • Durch $a_n = n$ für alle $n \in \mathbb{N}_0$ ist eine Folge definiert. Das ist die Folge der natürlichen Zahlen.

- Durch $a_n = \frac{1}{n}$ für alle $n \in \mathbb{N} \setminus \{0\}$ ist eine Folge definiert. Diese kann natürlich nicht für $n = 0$ definiert werden. Das ist die Folge der Kehrwerte der natürlichen Zahlen. Zur Veranschaulichung kann man sich die ersten Folgenglieder notieren, was allerdings nicht die allgemeine Angabe der Folge ersetzt:

$$\begin{aligned} a_1 &= \frac{1}{1} = 1, \\ a_2 &= \frac{1}{2}, \\ a_3 &= \frac{1}{3}, \end{aligned}$$

$$a_4 = \frac{1}{4},$$

$$\dots$$

- Durch $a_n = n^4 + 1,7n + \frac{1}{n^2} + 2^n$ für alle $n \in \mathbb{N} \setminus \{0\}$ ist eine Folge vorgegeben. Diese ist viel weniger intuitiv als die beiden vorherigen Beispiele.
- Folgen können auch rekursiv angegeben werden. Ein typisches Beispiel dafür ist die Fibonacci-Folge, die auch in der Informatik häufig betrachtet wird. Es gibt bezüglich der Nummerierung der Fibonacci-Zahlen zwei unterschiedliche, gebräuchliche Konventionen. Wir verwenden die folgende:

$$F_0 = F_1 = 1,$$

$$F_{n+2} = F_{n+1} + F_n \text{ für alle } n \geq 0.$$

Die ersten Folgenglieder sind

$$F_2 = F_1 + F_0 = 1 + 1 = 2,$$

$$F_3 = F_2 + F_1 = 2 + 1 = 3,$$

$$F_4 = F_3 + F_2 = 3 + 2 = 5,$$

$$F_5 = F_4 + F_3 = 5 + 3 = 8,$$

$$\dots$$

Durch die Vorschrift ist jedes Folgenglied eindeutig festgelegt. Allerdings muss man, um die 100-ste Fibonacci-Zahl anhand der Definition zu berechnen, alle vorherigen 100 Fibonacci-Zahlen bestimmt haben, was die Fibonacci-Folge von den bisherigen unterscheidet.

- $1, 2, 4, \dots$ ist hingegen keine Folge. Diese Angabe legt nicht eindeutig fest, wie es weitergeht, informell ausgedrückt, also welche Folgenglieder an weiteren Stellen erscheinen. Es wäre natürlich denkbar oder vielleicht sogar naheliegend, diese Zahlen mit $8, 16, \dots$ zu der Folge $b_n = 2^n$ für $n \in \mathbb{N}_0$ fortzusetzen; es sind allerdings immer auch andere Fortsetzungen möglich, hier etwa die Folge, bei der jedes Glied aus dem vorherigen durch Addition der nächsten natürlichen Zahl entsteht, also c_n , durch $c_0 = 1$ und $c_n = c_{n-1} + n$ für $n \geq 1$ festgelegt, gemeint ist. Mathematisch wollen wir also erst dann von einer Folge sprechen, wenn für alle natürlichen Zahlen n das entsprechende Folgenglied a_n festgelegt ist, auch wenn es nicht unbedingt einfach sein muss, dieses auszurechnen.

Wir werden noch einen weiteren einfachen Begriff brauchen, der eine spezielle Art von Folgen auszeichnet.

Definition 10.2. Eine **Reihe** ist eine Folge, deren Folgenglieder als Summen einer anderen, festen Folge vorgegeben sind.

Diese Definition ist eigentlich recht einfach. Um darüber besser und präziser reden zu können, erinnern wir nochmal an die Summennotation. Sei $(a_n)_{n \in \mathbb{N}}$ dazu eine feste Folge. Für die Summe der ersten k Folgenglieder dieser Folge notieren wir als $\sum_{i=1}^k a_i$. Dabei steht \sum für Summe, a_i notiert, welche Terme aufsummiert werden, der untere Index (hier $i = 1$) sagt sowohl, wie die Variable heißt, die in der Summe verändert wird, als auch, bei welchem Wert sie anfängt, und der obere Index k sagt schließlich, bei welchem Wert der Variable die Summe endet. In unserer üblichen, informellen Schreibweise erhalten wir also

$$\sum_{i=1}^k a_i = a_1 + a_2 + a_3 + \dots + a_{k-1} + a_k.$$

Die Reihe $(\sum_{i=1}^k a_i)_{k \in \mathbb{N}}$ ist also die Folge, deren erste Folgenglieder $a_1, a_1 + a_2, a_1 + a_2 + a_3, a_1 + a_2 + a_3 + a_4, \dots$ sind, also gerade die Summen der alten Folge $(a_n)_{n \in \mathbb{N}}$. Um dieses Konzept besser zu verstehen, betrachten wir einige Beispiele.

Beispiel. • Wir beginnen mit der Folge der natürlichen Zahlen, die wir bereits vorher betrachtet haben, also mit der Folge $a_n = n$ für alle $n \in \mathbb{N}$. Die dazugehörige Reihe ist die Folge b_n , die aus Summen unserer vorgegebenen Folge a_n besteht, genauer aus den Summen der Form $b_k = \sum_{i=1}^k i$. Die ersten Folgenglieder sind also

$$\begin{aligned} b_1 &= 1, \\ b_2 &= 1 + 2 = 3, \\ b_3 &= 1 + 2 + 3 = 6, \\ b_4 &= 1 + 2 + 3 + 4 = 10, \\ &\dots \end{aligned}$$

In diesem Fall können wir die Folgenglieder auch durch eine geschlossene Formel beschreiben, anstatt die Darstellung durch Summe zu benutzen: Aus Mathematischen Grundlagen 1 ist die Formel $\sum_{i=1}^k i = \frac{k(k+1)}{2}$ bekannt. Die Folge $(b_k)_{k \in \mathbb{N}}$ kann also gleichermaßen auch in der Form $b_k = \frac{k(k+1)}{2}$ angegeben werden; allerdings ist in dieser Form ihre Darstellung als Reihe nicht sichtbar.

- Diesmal starten wir mit der Folge der Kehrwerte der natürlichen Zahlen, also mit der Folge $a_n = \frac{1}{n}$ für alle $n \in \mathbb{N}$. Die dazugehörige Reihe ist die Folge H_n , die aus Summen unserer vorgegebenen Folge a_n besteht, genauer aus den Summen der Form $H_k = \sum_{i=1}^k \frac{1}{i}$. Die ersten Folgenglieder sind also

$$H_1 = \frac{1}{1} = 1,$$

$$\begin{aligned}
 H_2 &= \frac{1}{1} + \frac{1}{2} = \frac{3}{2}, \\
 H_3 &= \frac{1}{1} + \frac{1}{2} + \frac{1}{3} = \frac{3}{2} + \frac{1}{3} = \frac{11}{6}, \\
 H_4 &= \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \\
 &\dots
 \end{aligned}$$

Diese Reihe H_n , die als Summe der Kehrwerte natürlicher Zahlen definiert ist, heißt **harmonische Reihe**. Sie lässt sich nicht wie die vorherige Reihe durch eine einfache andere Formel beschreiben. Auch die Größenordnung der Zahlen, die man hier als Folgenglieder bekommt, lässt sich zunächst nicht ohne weiteres abschätzen. Dazu kommen wir später nochmal.

- Die nächste Reihe, die wir betrachten werden, ist die sogenannte **geometrische Reihe** (wir betrachten den Spezialfall für den Wert $\frac{1}{2}$). Hier starten wir mit der Folge $\left(\frac{1}{2^n}\right)_{n \in \mathbb{N}_0}$ und betrachten die Summen dieser Folge. Die ersten Glieder sehen wie folgt aus:

$$\begin{aligned}
 G_0 &= \sum_{k=0}^0 \frac{1}{2^k} = \frac{1}{2^0} = \frac{1}{1} = 1, \\
 G_1 &= \sum_{k=0}^1 \frac{1}{2^k} = \frac{1}{2^0} + \frac{1}{2^1} = \frac{1}{1} + \frac{1}{2} = \frac{3}{2}, \\
 G_2 &= \sum_{k=0}^2 \frac{1}{2^k} = \frac{1}{2^0} + \frac{1}{2^1} + \frac{1}{2^2} = \frac{1}{1} + \frac{1}{2} + \frac{1}{4} = \frac{7}{4}, \\
 &\dots
 \end{aligned}$$

Zu der geometrischen Reihe kommen wir in Kürze wieder zurück.

- Ebenfalls für den späteren Gebrauch betrachten wir das folgende Beispiel. Sei $(a_n)_{n \in \mathbb{N}}$ diesmal eine beliebige Folge. Sei x zunächst eine reelle Zahl. Dann können wir die Folge $(a_n x^n)_{n \in \mathbb{N}}$ betrachten, also jedes Folgenglied mit der entsprechenden Potenz unserer Zahl x multiplizieren. Zu dieser Folge betrachten wir nun die zugehörige Reihe, also Folge der Summen $\sum_{k=0}^n a_k x^k$. Betrachtet man x nun nicht mehr als feste Zahl, sondern als Variable, so erhält man gewissermaßen eine Folge von Polynomen. Diese heißt dann **Potenzreihe** in der Variablen x (mit Koeffizienten a_n). Auf Potenzreihen kommen wir später nochmal zu sprechen.

Nun wenden wir uns dem eigentlichen Ziel dieses Kapitels zu: Der Definition des Grenzwertes. Zunächst beschreiben wir die Idee dahinter. Als erstes möchten wir, dass eine Folge $(a_n)_{n \in \mathbb{N}}$ eine reelle Zahl a approximiert. Wenn wir also eine bestimmte Anzahl der Nachkommastellen von a wissen wollen, dann sollten wir einfach jedes hinreichend späte Folgenglied anschauen

können und eine Approximation von a in gewünschter Genauigkeit erhalten können. Beispielsweise kann es so sein, dass alle Folgenglieder a_n ab dem Tausendstem dieselben ersten drei Nachkommastellen haben, und zwar dieselben wie a , dann ab dem 2000-ten die ersten vier Nachkommastellen, dann beispielsweise ab dem 10000-ten sollen die ersten fünf Nachkommastellen übereinstimmen, und so weiter. Wichtig dabei ist, dass man zu jeder vorgegebenen Genauigkeit eine Position in der Folge finden, ab der alle Folgenglieder mit eben dieser gewünschten Genauigkeit mit der Zahl a übereinstimmen.

Die Notwendigkeit einer kleinen Änderung wird schnell deutlich: Wir betrachten als Beispiel die Folge

$$c_n = 0, \underbrace{99 \dots 9}_n,$$

n Neunen

wo das n -te Folgenglied genau n Neunen an den n Nachkommastellen besitzt. Sicherlich sollte unsere Definition so formuliert sein, dass diese Folge die Zahl 1 approximiert, auch wenn bereits die Einerstelle der Zahlen nie übereinstimmen wird. Deswegen betrachten wir stattdessen die Differenz der Folgenglieder und des potentiellen Grenzwertes a , und verlangen, dass diese Differenz möglichst nah bei 0 ist.

Nun zur präzisen Formulierung.

Definition 10.3. Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge und a eine reelle Zahl. Man sagt, a sei der **Grenzwert** der Folge $(a_n)_{n \in \mathbb{N}}$, falls gilt:

Zu jeder natürlichen Zahl $k \in \mathbb{N}$ gibt es eine Stelle $N \in \mathbb{N}$, sodass für alle $n > N$ gilt:

$$|a_n - a| < 10^{-k}.$$

Diese Definition ist recht abstrakt, und wir kommentieren diese zunächst, bevor wir einige Beispiele von Grenzwerten sehen. Die Zahl k in der Definition sollte man sich wie die Genauigkeit vorstellen, also in etwa wie die Anzahl der Nachkommastellen, die wir übereinstimmend haben wollen. Die Folge $(a_n)_{n \in \mathbb{N}}$ hat also den Grenzwert a , wenn es zu jeder solchen Genauigkeit eine Stelle N in der Folge gibt, ab der alle späteren Folgenglieder (also alle a_n für $n > N$) mit a bis auf Genauigkeit k übereinstimmen. Man bemerke, dass der Betrag in der Definition erlaubt, dass die Folgenglieder mal größer und mal kleiner als a sind, jedoch die Abweichung nie zu groß werden darf.

Nicht jede Folge hat einen Grenzwert. Recht klar ist das beispielsweise bei der Folge $a_n = n$ der natürlichen Zahlen: Diese wächst über alle Grenzen und „stabilisiert“ sich überhaupt nicht in der Nähe einer reellen Zahl. Wenn eine Folge einen Grenzwert hat, ist dieser eindeutig, mit anderen Worten kann eine Folge nicht mehrere reelle Zahlen approximieren.

Ist a der Grenzwert der Folge $(a_n)_{n \in \mathbb{N}}$, so sagt man dazu auch alternativ, $(a_n)_{n \in \mathbb{N}}$ **konvergiert** gegen a und schreibt dafür $\lim_{n \rightarrow \infty} a_n = a$ oder, etwas weniger formal, $(a_n) \rightarrow a$.

Nun betrachten wir ein Beispiel, in dem wir formal zeigen werden, dass eine Folge gegen 0 konvergiert.

Proposition 10.4. *Die Folge $a_n = \frac{1}{n}$, $n \in \mathbb{N}$, der Kehrwerte der natürlichen Zahlen konvergiert gegen 0.*

Beweis. Wir wollen die formale Definition vom Grenzwert benutzen. Wir müssen also zu jeder Genauigkeit 10^{-k} mit $k \in \mathbb{N}$, die uns vorgegeben ist, eine Position $N \in \mathbb{N}$ finden, sodass für alle späteren Folgenglieder a_n (also für alle mit $n > N$) gilt: $|a_n - a| < 10^{-k}$. In unserem Fall ist $a = 0$ und $a_n = \frac{1}{n}$.

Als kleine Vorüberlegung bemerken wir, dass der Kehrwert einer Zahl umso kleiner wird, je größer die Zahl ist. In Formeln heißt es für beliebige positive reelle Zahlen $x, y \in \mathbb{R}$, $x, y > 0$: Ist $x < y$, so ist $\frac{1}{y} < \frac{1}{x}$.

Sei nun ein k wie oben vorgegeben. Für das N dürfen wir jetzt eine Zahl einfach auswählen und hier setzen wir $N = 10^k$. Übrigens sagt die Definition nicht, dass wir die früheste Position finden müssen, ab der die Ungleichung erfüllt ist, sie darf gerne auch vor der Position N erfüllt sein: Das Wichtige ist, dass die Ungleichung an allen späteren Positionen erfüllt ist.

Nun gilt für alle natürlichen Zahlen n , die größer sind als 10^k , nach unserer Vorüberlegung und nach Definition des Betrags:

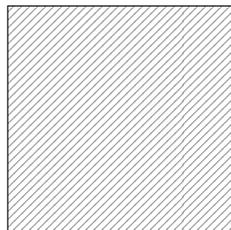
$$\left| \frac{1}{n} - 0 \right| = \left| \frac{1}{n} \right| = \frac{1}{n} < \frac{1}{10^k} = 10^{-k}.$$

Somit haben wir bewiesen, dass die Folge $\left(\frac{1}{n}\right)_{n \in \mathbb{N}}$ gegen 0 konvergiert. \square

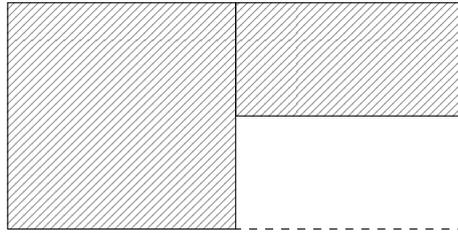
In unserem zweiten Beispiel wollen wir einen weniger formalen Beweis für die Konvergenz einer Folge liefern.

Beispiel. Wir betrachten nun wieder die geometrische Reihe, also die Folge der Summen $\left(\sum_{k=0}^n \frac{1}{2^k}\right)_{n \in \mathbb{N}}$. Wir haben bereits einige Folgenglieder dieser Folge ausgerechnet; nun wollen wir sie zusätzlich graphisch darstellen und dabei erläutern, warum diese geometrische Reihe gegen 2 konvergiert.

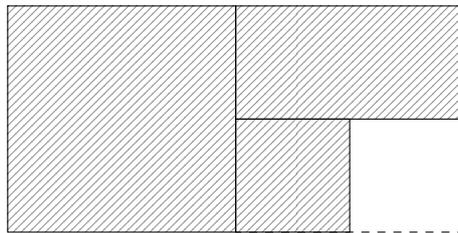
Für $G_0 = 1$ zeichnen wir ein Quadrat mit der Seitenlänge 1 ein.



Für $G_1 = 1 + \frac{1}{2} = \frac{3}{2}$ fügen wir nun ein Rechteck mit Seitenlängen 1 und $\frac{1}{2}$, also von der Gesamtfläche $\frac{1}{2}$, hinzu. Wir bemerken, dass dann noch $\frac{1}{2}$ Flächeneinheiten bis zum Rechteck der Größe 1×2 fehlen.



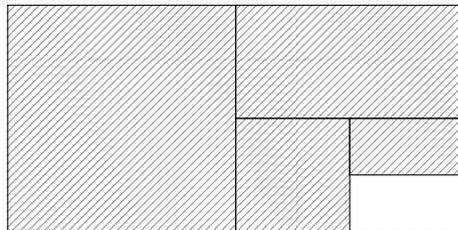
Nun kommt genau die Hälfte der fehlenden Fläche, nämlich $\frac{1}{4}$ Flächeneinheiten hinzu, und zwar veranschaulichen wir das als ein Quadrat mit Seitenlängen $\frac{1}{2}$. Wir erhalten die Gesamtfläche von $1 + \frac{1}{2} + \frac{1}{4} = \frac{7}{4} = 2 - \frac{1}{4}$.



Als nächstes kommt $\frac{1}{8}$ hinzu, also wieder die Hälfte von dem noch fehlenden $\frac{1}{4}$, und wir erhalten

$$G_3 = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{15}{8} = 2 - \frac{1}{8},$$

also



Es lässt sich bereits erahnen, und wir werden uns im Rahmen dieser Veranstaltung mit dieser Begründung zufrieden geben, dass die weiteren Folgenglieder der Folge G_n , also die Summe der Reihe, immer näher an 2 drankommen. Es ist also plausibel (und kann auch präzise gemacht werden), dass die geometrische Reihe $\sum_{k=0}^n \frac{1}{2^k}$ gegen 2 konvergiert. Um den Grenzwert einer Reihe zu notieren, schreibt man auch

$$\sum_{k=0}^{\infty} \frac{1}{2^k} \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{2^k} = 2.$$

Es ist bemerkenswert, dass die Folge G_n konvergiert, obwohl die Folgenglieder immer größer werden. Entscheidend ist hier, dass die Folge aber nicht über alle Schranken wächst. Informell kann man sagen, dass die Folge zwar immer weiter wächst, doch der Zuwachs immer kleiner wird. In solchen Fällen muss genauer untersucht werden, ob die Folge konvergiert.

Wir betrachten einige weitere Beispiele von konvergenten und nicht - konvergenten Folgen.

Beispiel. • Die Folge $b_n = 1$, $n \in \mathbb{N}$, die als jedes Folgenglied 1 hat, konvergiert auch gegen 1. Tatsächlich gilt für jedes $k \in \mathbb{N}$ und jedes Folgenglied b_n :

$$|b_n - 1| = |1 - 1| = 0 < 10^{-k}.$$

- Die Folge der natürlichen Zahlen $a_n = n$, $n \in \mathbb{N}$, wächst unbeschränkt über alle Grenzen. Sie kann somit keinen Grenzwert haben, da die Folgenglieder beliebig viel größer als jede reelle Zahl werden. Diese Folge konvergiert also nicht.
- Ein weiteres Beispiel einer nicht-konvergenten Folge ist von ganz anderer Art. Wir betrachten die Folge $c_n = (-1)^n$ für $n \in \mathbb{N}$. Die ersten Folgenglieder dieser Folge sind $c_1 = -1$, $c_2 = 1$, $c_3 = -1$, $c_4 = 1$. Auch weiterhin gibt es nur zwei mögliche Werte für c_n , nämlich $+1$ und -1 , und diese werden abwechselnd angenommen: Die Folgenglieder mit der geraden Nummer sind stets 1 , mit ungerader Nummer stets -1 . Diese Folge hat keinen Grenzwert: Informell gesprochen, hieß Konvergenz, dass man eine Zahl durch die Folge immer besser approximiert. Hier würden nur 1 oder -1 als Grenzwert in Frage kommen. Wäre allerdings 1 dieser Grenzwert, so wäre jede zweite Näherung, nämlich bei geraden Folgengliedern, perfekt, allerdings wäre jede darauffolgende Näherung um 2 falsch. Dieses Argument kann man nun zu einem präzisen Beweis dafür machen, dass diese Folge nicht konvergiert.
- Wir betrachten wieder die Folge $b_n = 1$, $n \in \mathbb{N}$, allerdings beschäftigen wir uns diesmal nicht mit der Konvergenz der Folge selbst, sondern mit der Konvergenz der dazugehörigen Reihe. Wir bezeichnen die Folgenglieder dieser Reihe mit $s_n = \sum_{i=1}^n b_i$. Das n -te Folgenglied besteht aus der Summe von n Einsen, also gilt

$$s_n = \sum_{i=1}^n b_i = \sum_{i=1}^n 1 = n,$$

und die Folge s_n stimmt mit der Folge der natürlichen Zahlen überein. Wir haben uns bereits überlegt, dass diese Folge nicht konvergiert, also konvergiert die Reihe, die zu der Folge $b_n = 1$ gehört, nicht.

Es ist manchmal aufwendig, direkt mit der Definition der Konvergenz zu arbeiten. Häufig ist es sinnvoll, stattdessen den folgenden Satz zu verwenden, den wir nicht beweisen werden. Der erste Teil sagt im Wesentlichen aus, dass man mit konvergenten Folgen rechnen kann, also dass Summe, Differenz,

Produkt und mit einer kleinen Einschränkung der Quotient zweier konvergenter Folgen wieder konvergent ist, und der Grenzwert sich durch die gleiche Grundrechenart aus den Grenzwerten der Folgen bestimmt. Der zweite Teil wird zu einem späteren Zeitpunkt gebraucht. Wir notieren ihn schon, obwohl dieser nicht ganz so intuitiv ist.

Dabei müssen wir uns einmal überlegen, wie man mit Folgen rechnet. Hier ist es so, dass man an jeder Position n die jeweiligen n -ten Folgenglieder addiert, subtrahiert, multipliziert oder dividiert und dadurch das n -te Folgenglied der neuen Folge erhält. Mit dieser Festlegung können wir nun unseren Satz formulieren.

Satz 10.5. 1. Seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ zwei Folgen, die jeweils gegen $a \in \mathbb{R}$ bzw. $b \in \mathbb{R}$ konvergieren. Dann sind die Folgen $(a_n + b_n)_{n \in \mathbb{N}}$, $(a_n - b_n)_{n \in \mathbb{N}}$, $(a_n \cdot b_n)_{n \in \mathbb{N}}$ konvergent und es gilt:

$$\begin{aligned}\lim_{n \rightarrow \infty} (a_n + b_n) &= a + b, \\ \lim_{n \rightarrow \infty} (a_n - b_n) &= a - b, \\ \lim_{n \rightarrow \infty} (a_n \cdot b_n) &= a \cdot b.\end{aligned}$$

Ist ferner $b_n \neq 0$ für alle $n \in \mathbb{N}$ und $b \neq 0$, so ist auch $\left(\frac{a_n}{b_n}\right)_{n \in \mathbb{N}}$ eine konvergente Folge und es gilt

$$\lim_{n \rightarrow \infty} \left(\frac{a_n}{b_n}\right) = \frac{a}{b}.$$

2. Seien nun $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ zwei Folgen, die gegen dieselbe Zahl $a \in \mathbb{R}$ konvergieren. Sei $(c_n)_{n \in \mathbb{N}}$ eine weitere Folge, und es gelte für die Folgenglieder der drei Folgen die Ungleichung

$$a_n \leq c_n \leq b_n$$

für alle $n \in \mathbb{N}$. Dann konvergiert auch die Folge (c_n) gegen den Grenzwert a der Folgen (a_n) und (b_n) .

Für den Quotienten zweier Folgen brauchen wir eine Zusatzbedingung, da es wie immer nicht möglich ist, durch 0 zu teilen, und wir diesen Fall ausschließen müssen. Nun wollen wir einige Beispiele für die Anwendung dieses Satzes sehen.

Beispiel. • Genauso wie die Folge, die konstant 1 ist, konvergiert auch die Folge $a_n = 7$, $n \in \mathbb{N}$, allerdings ist der Grenzwert in diesem Fall die Zahl 7. Ferner haben wir gesehen, dass die Folge $b_n = \frac{1}{n}$, $n \in \mathbb{N}$, gegen 0 konvergiert. Nun können wir daraus mit Hilfe des Satzes beispielsweise schließen, dass die Folge $\left(\frac{7}{n}\right)_{n \in \mathbb{N}}$, die ja die Produktfolge $(a_n b_n)_{n \in \mathbb{N}}$ ist, gegen das Produkt der Grenzwerte konvergiert, also gegen $7 \cdot 0 = 0$. Mit dem Satz erhalten wir also

$$\lim_{n \rightarrow \infty} \left(\frac{7}{n}\right) = 0.$$

- Kombiniert man die Folgen $(1)_{n \in \mathbb{N}}$ und $(\frac{1}{n})_{n \in \mathbb{N}}$, die jeweils gegen 1 bzw. 0 konvergieren, so erhält man mit dem Satz beispielsweise

$$\lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right) = 1 - 0 = 1 \text{ und}$$

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right) = 1 + 0 = 1.$$

Man bemerke, dass die erste Folge wieder die Eigenschaft hat, dass die Folgenglieder dieser Folge immer größer werden, denn wir ziehen immer weniger von der 1 ab. Das können wir uns an den ersten Folgengliedern veranschaulichen: Diese sind

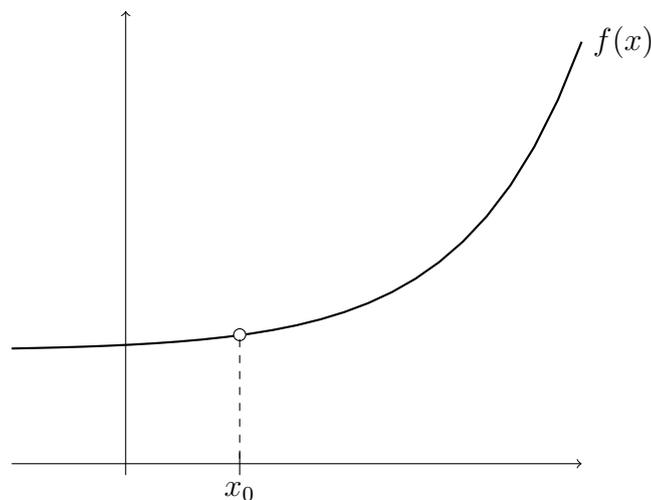
$$0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \frac{5}{6}, \dots$$

Grenzwerte von Funktionen

Nun wollen wir einen weiteren Schritt in Richtung unseres momentanen Ziels, nämlich der Definition der Ableitung, machen und die Grenzwerte von Funktionen definieren. Dabei bauen wir auf dem Begriff des Grenzwertes einer Folge auf.

Der Definitionsbereich unserer Funktionen ist meist ein offenes Intervall. Wir erinnern uns daran, dass das ein Intervall der Form (a, b) oder $(-\infty, b)$ oder (a, ∞) oder $I = \mathbb{R} = (-\infty, \infty)$ ist, wobei $a, b \in \mathbb{R}$ sind. Sei also ein offenes Intervall I vorgegeben, und sei x_0 ein Punkt in I . Wir nehmen nun an, dass unsere Funktion zunächst auf I ohne den Punkt x_0 definiert ist, also haben wir eine Funktion der Form $F: I \setminus \{x_0\} \rightarrow \mathbb{R}$.

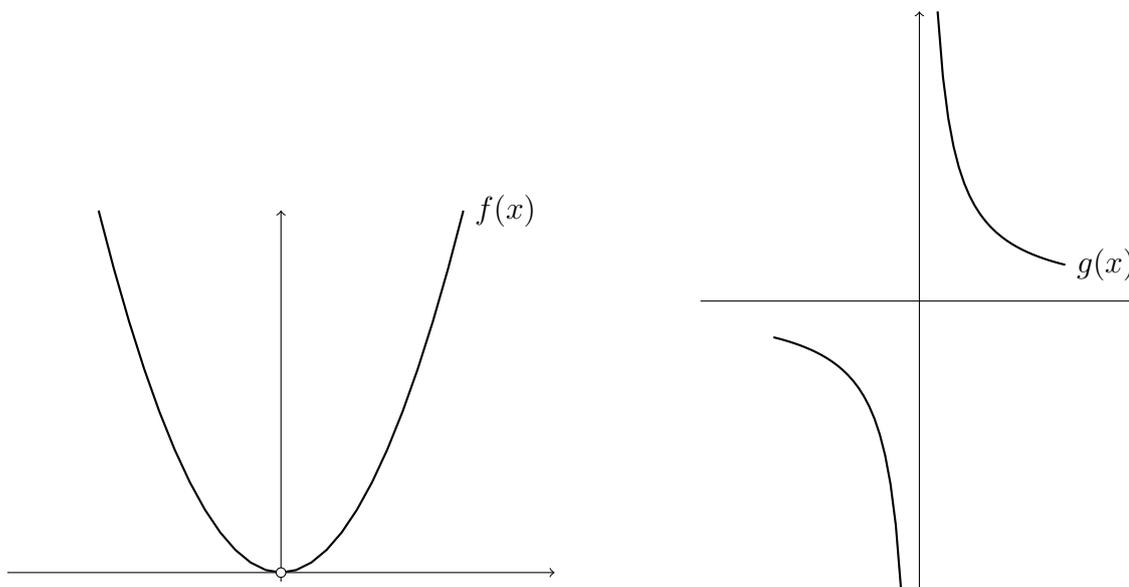
Wir fragen uns, ob diese Funktion in einer sinnvollen Art und Weise im Punkt x_0 fortgesetzt werden kann, also in etwa wie im folgenden Bild.



Wie schon im Schulunterricht üblich, markieren wir mit einem unausgefüllten kleinen Kreis am Funktionsgraphen einen Punkt, an dem der Wert der

Funktion nicht definiert ist. Häufig, wie auch in diesem Bild, hat man eine klare Vermutung, welcher Wert an der Stelle x_0 gewählt werden muss. Hingegen kennen wir auch Beispiele, wo keine sinnvolle Wahl des Wertes für f an der Stelle x_0 möglich ist.

Wir schauen uns dazu die Funktionen $f, g: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$, die durch $f(x) = x^2$ bzw. $g(x) = \frac{1}{x}$ definiert ist. Die Funktionengraphen sehen in etwa wie folgt aus:



Dabei merkt man, dass der Wert $0^2 = 0$ perfekt das linke Bild vervollständigen würde, während man im Graphen auf der rechten Seite nicht wüßte, welchen Wert man für 0 wählen sollte. Uns geht es nun darum, dieses „perfekt passen“ präzise zu fassen.

Definition 10.6. Sei I ein offenes Intervall in \mathbb{R} und $x_0 \in I$. Sei ferner $f: I \setminus \{x_0\} \rightarrow \mathbb{R}$ eine Funktion. Der **Grenzwert** von f für $x \rightarrow x_0$ ist der übereinstimmende Grenzwert aller Folgen $(f(a_n))_{n \in \mathbb{N}}$, wobei $(a_n)_{n \in \mathbb{N}}$ eine beliebige Folge ist, die gegen x_0 konvergiert und dessen alle Folgenglieder im Definitionsbereich $I \setminus \{x_0\}$ von f liegen (also $a_n \in I \setminus \{x_0\}$ für alle $n \in \mathbb{N}$), falls dieser übereinstimmende Grenzwert existiert. Dieser Grenzwert wird mit $\lim_{x \rightarrow x_0} f(x)$ bezeichnet.

Die Idee dahinter ist in etwa die folgende. Wir können im Allgemeinen den Wert x_0 nicht in die Funktion f einsetzen, oder dürfen es jedenfalls nicht. Stattdessen wählt man sich also eine Anzahl von „Testpunkten“, die immer näher an x_0 dran liegen, und überprüft, ob die Werte von f auf diesen Testpunkten eine „Tendenz“ aufzeigen, also einen Wert für $f(x_0)$ vermuten lassen, der mit den „Testdaten“ im Einklang steht. Nun will man sich nicht nur auf einen Satz solcher „Testdaten“, also auf eine Folge $(a_n)_{n \in \mathbb{N}}$, die in

$I \setminus \{x_0\}$ liegt und gegen x_0 konvergiert, beschränken. Stattdessen wollen wir, dass der von uns gewählte Wert für *alle* möglichen Wahlen einer solchen Folge derselbe ist. Das wird nicht bei jeder Funktion möglich sein: Manchmal ist ein Grenzwert $\lim_{x \rightarrow x_0} f(x)$ nicht definiert, etwa, wenn einsetzen mancher Folgen gegen x_0 konvergenter Folgen in f keine konvergente Folge liefert, oder wenn man durch verschiedene Approximationen von x_0 verschiedene potentielle Werte für $f(x_0)$ erhält. Das wird in unseren Überlegungen jedoch der Ausnahmefall sein.

Wir wollen die Definition anhand eines bereits angekündigten Beispiels etwas besser verstehen.

Beispiel. Sei $f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ erneut die Funktion, die durch $f(x) = x^2$ definiert ist. Wir wollen nun anhand der Definition einsehen, dass der Grenzwert der Funktion f für $x \rightarrow 0$ gerade 0 ist, wie wir uns vorher veranschaulicht haben.

Dafür müssen wir prüfen, dass für jede Folge $(a_n)_{n \in \mathbb{N}}$, die gegen 0 konvergiert und für die $a_n \neq 0$ für alle $n \in \mathbb{N}$ gilt, auch der Grenzwert der Folge $(f(a_n))_{n \in \mathbb{N}}$ gerade 0 ist. Dabei erhalten wir durch Anwendung von f die Folge $(a_n^2)_{n \in \mathbb{N}}$, in der das Folgenglied an der Stelle n aus dem Folgenglied a_n der ursprünglichen Folge an selber Stelle durch Quadrieren erhalten wurde.

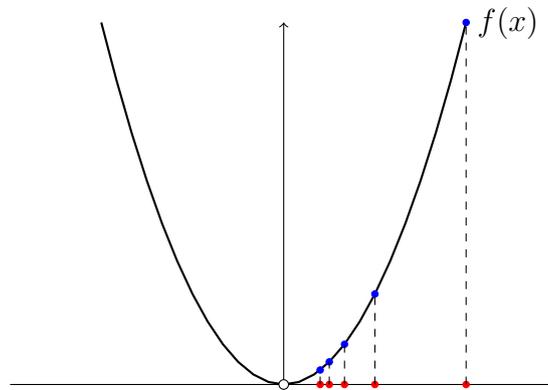
Bevor wir das allgemein machen, wählen wir uns eine konkrete „Testfolge“ und überprüfen, ob die Werte von f immer näher an 0 sind, wenn wir die Folgenglieder dieser Folge in f einsetzen. Wir haben bereits gesehen, dass die Folge $\left(\frac{7}{n}\right)_{n \in \mathbb{N}}$ gegen 0 konvergiert. Ferner ist kein Folgenglied dieser Folge 0, also ist sie als „Testfolge“ in diesem Fall geeignet. Wir setzen also jedes Folgenglied in die Funktion f ein und erhalten die Folge

$$\left(f\left(\frac{7}{n}\right)\right)_{n \in \mathbb{N}} = \left(\left(\frac{7}{n}\right)^2\right)_{n \in \mathbb{N}} = \left(\frac{7}{n} \cdot \frac{7}{n}\right)_{n \in \mathbb{N}}.$$

Das ist das Produkt unserer ursprünglichen Folge $\left(\frac{7}{n}\right)_{n \in \mathbb{N}}$ mit sich selbst. Der Satz 10.5 sagt uns nun, dass das Produkt zweier Folgen, die gegen 0 konvergieren, ebenfalls gegen $0 \cdot 0 = 0$ konvergiert. Folglich konvergiert die Folge der Werte von f , die wir durch Einsetzen der Folge $\left(\frac{7}{n}\right)_{n \in \mathbb{N}}$ erhalten, gegen 0.

Das reicht allerdings nicht, um den Grenzwert von f für $x \rightarrow 0$ zu finden, obwohl das ein Hinweis darauf ist, dass dieser Grenzwert 0 sein sollte. Allerdings haben wir das bislang für eine einzige „Testfolge“ überprüft, und wir würden gerne wissen, dass jede Testfolge, also etwa die Folge $\left(\frac{1}{n}\right)_{n \in \mathbb{N}}$ oder $\left(\frac{8}{n}\right)_{n \in \mathbb{N}}$ oder eben jede andere, dasselbe Ergebnis liefert.

Deshalb müssen wir nun eine ganz ähnliche Argumentation für jede Folge $(a_n)_{n \in \mathbb{N}}$, die gegen 0 konvergiert und die $a_n \neq 0$ für alle $n \in \mathbb{N}$ erfüllt, führen. Das Problem können wir uns wie folgt bildlich vergegenwärtigen:



Hier markieren die roten Punkte, die wir in die Funktion f einsetzen, und die y -Werte der blauen Punkte bilden gerade die Folge $(f(a_n))_{n \in \mathbb{N}}$. Im Bild ist es plausibel, dass die y -Werte dieser Folge immer näher an Null liegen, wenn die eingesetzten Werte immer näher bei Null sind. Das wollen wir nun beweisen.

Sei also $(a_n)_{n \in \mathbb{N}}$ eine Folge, die gegen 0 konvergiert und die $a_n \neq 0$ für alle $n \in \mathbb{N}$ erfüllt. Wir bemerken, dass auch hier die Folge der Werte, die wir durch Anwenden von f erhalten, sich als

$$(f(a_n))_{n \in \mathbb{N}} = (a_n^2)_{n \in \mathbb{N}} = (a_n \cdot a_n)_{n \in \mathbb{N}}$$

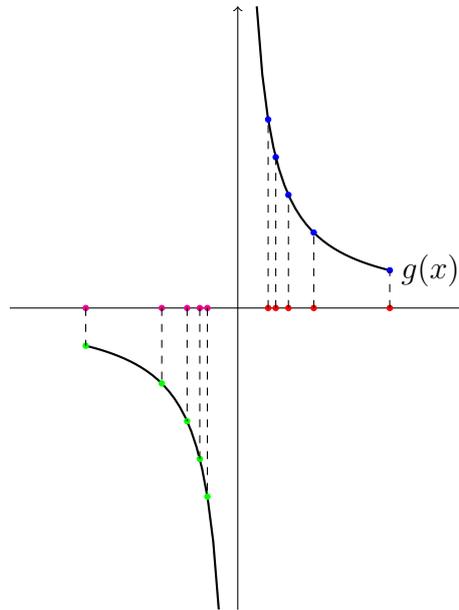
schreiben lässt. Also ist die Folge, die wir so erhalten, das Produkt der ursprünglichen Folge $(a_n)_{n \in \mathbb{N}}$ mit sich selbst. Wir nutzen nun erneut den Satz 10.5 und beweisen dadurch, dass die neue Folge gegen $0 \cdot 0 = 0$ konvergiert. Nun hing das nicht von der Wahl der „Testfolge“ ab, also stimmt der Grenzwert von $(f(a_n))_n$ für alle „Testfolgen“ a_n überein und ist stets 0. Wir haben also bewiesen, dass der Grenzwert der Funktion $f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$, $f(x) = x^2$ für $x \rightarrow 0$ existiert und 0 ist, in Formeln:

$$\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} x^2 = 0.$$

Im folgenden Beispiel wollen wir nochmal kurz erläutern, dass Grenzwerte von Funktionen manchmal nicht existieren.

Beispiel. Sei nun $g: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ die Funktion, die durch $g(x) = \frac{1}{x}$ definiert ist. Für diese existiert der Grenzwert für $x \rightarrow 0$ nicht. Wir wollen das nur informell erläutern. Setzt man positive Argumente x ein, die sehr klein sind, so werden die Grenzwerte immer größer. Für $x \rightarrow 0$ wächst die Funktion g über alle Schranken. Das würde schon reichen, damit g kein Grenzwert für $x \rightarrow 0$ hat. Wir wollen außerdem noch beobachten, dass wir für negative Argumente x vom sehr kleinen Betrag wieder betragsmäßig unbeschränkt wachsende Werte der Funktion g erhalten, die allerdings negativ sind. Auch das würde ausreichen, um zu sehen, dass die Funktion g für $x \rightarrow 0$ keinen Grenzwert besitzt.

Beides wollen wir in einem Bild veranschaulichen:



Hier markieren die roten bzw. die magenta Punkte die positiven bzw. negativen Folgenglieder, die in g eingesetzt werden, und die blauen bzw. grünen Punkte sind die entsprechenden Punkte auf dem Funktionsgraphen. Es lässt sich bereits erahnen, dass die y -Koordinaten der blauen Punkte beliebig wachsen, wenn die roten Punkte näher am Nullpunkt liegen. Ähnlich werden auch die y -Koordinaten der grünen Punkte betragsmäßig beliebig groß, bleiben aber stets negativ.

Wieder ist es recht aufwendig, ohne weitere Hilfsmittel zu prüfen, ob der Grenzwert einer Funktion existiert und welchen Wert er annimmt. Jedoch lässt sich auch hier mit Grenzwerten rechnen, und wir erhalten das folgende Analogon vom Satz 10.5:

Satz 10.7. Sei $I \subseteq \mathbb{R}$ ein Intervall und $x_0 \in I$ ein Punkt in I . Seien $f, g: I \setminus \{x_0\} \rightarrow \mathbb{R}$ zwei Funktionen, die jeweils den Grenzwert y bzw. z für $x \rightarrow x_0$ haben, in Formeln:

$$\lim_{x \rightarrow x_0} f(x) = y \text{ und } \lim_{x \rightarrow x_0} g(x) = z.$$

Dann haben die Funktionen $f + g: I \setminus \{x_0\} \rightarrow \mathbb{R}$, definiert durch $(f + g)(x) = f(x) + g(x)$, außerdem $f - g: I \setminus \{x_0\} \rightarrow \mathbb{R}$, definiert durch $(f - g)(x) = f(x) - g(x)$ und $f \cdot g: I \setminus \{x_0\} \rightarrow \mathbb{R}$, definiert durch $(f \cdot g)(x) = f(x) \cdot g(x)$, jeweils einen Grenzwert für $x \rightarrow x_0$, und es gilt:

$$\begin{aligned} \lim_{x \rightarrow x_0} (f(x) + g(x)) &= y + z, \\ \lim_{x \rightarrow x_0} (f(x) - g(x)) &= y - z, \\ \lim_{x \rightarrow x_0} (f(x) \cdot g(x)) &= y \cdot z. \end{aligned}$$

Ist ferner $g(x) \neq 0$ für alle $x \in I \setminus \{x_0\}$ und $z \neq 0$, so hat auch die Funktion $\frac{f}{g}: I \setminus \{x_0\} \rightarrow \mathbb{R}$, gegeben durch $\frac{f}{g}(x) = \frac{f(x)}{g(x)}$ einen Grenzwert und es gilt

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{y}{z}.$$

Der Grenzwert einer Funktion f für $x \rightarrow x_0$ soll einen geeigneten Wert für $f(x_0)$ liefern, der mit den anderen Werten der Funktion zusammenpasst. Manchmal, wie etwa im Beispiel der Funktion $f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$, $f(x) = x^2$, kann man den fraglichen x -Wert $x_0 = 0$ einfach in die Funktionsvorschrift einsetzen und erhält den Wert 0, den wir auch als Grenzwert von f für $x \rightarrow 0$ bestimmt haben. Das trifft auf fast alle Funktionen zu, die üblicherweise im Schulunterricht behandelt werden. Eine präzise Aussage lässt sich besonders gut für ganz-rationale Funktionen, also Funktionen, die als Quotient zweier Polynome definiert sind, formulieren.

Satz 10.8. Sei I ein offenes Intervall und a ein Punkt in I . Seien ferner P, Q Polynome und es gelte $Q(x) \neq 0$ für alle $x \in I$. Wir betrachten die Funktion

$$\begin{aligned} \frac{P}{Q}: I \setminus \{a\} &\rightarrow \mathbb{R}, \\ x &\mapsto \frac{P(x)}{Q(x)}. \end{aligned}$$

Der Grenzwert dieser Funktion für $x \rightarrow a$ ist gegeben gerade durch den Wert, den wir durch Einsetzen von a erhalten, also in Formeln:

$$\lim_{x \rightarrow a} \frac{P(x)}{Q(x)} = \frac{P(a)}{Q(a)}.$$

Die Bedingung $Q(a) \neq 0$ ist wie immer notwendig, um sicherzustellen, dass man nicht durch 0 teilt.

Wir betrachten einige Beispiele für die Anwendung dieses Satzes.

Beispiel. • Wir betrachten die Funktion von $\mathbb{R} \setminus \{2\}$ nach \mathbb{R} , die durch $x \mapsto \frac{x^2-4}{x^2+2}$ gegeben ist. Wir stellen fest, dass das eine ganzrationale Funktion ist, also ein Quotient zweier Polynome ist. Ferner bemerken wir, dass der Nenner nie Null wird, da $x^2 \geq 0$ für alle $x \in \mathbb{R}$ und $x^2 + 2 > 0$ für alle $x \in \mathbb{R}$ gilt. Nun können wir den Satz anwenden, um den Grenzwert dieser Funktion für $x \rightarrow 2$ zu bestimmen:

$$\lim_{x \rightarrow 2} \frac{x^2 - 4}{x^2 + 2} = \frac{2^2 - 4}{2^2 + 2} = \frac{0}{6} = 0.$$

Insbesondere bemerken wir, dass der Nenner in der untersuchten Stelle $x_0 = 2$ keine Nullstelle haben darf, damit der Satz anwendbar ist, aber eine Nullstelle im Zähler gar kein Problem darstellt. Insgesamt besagt der Satz 10.8, dass die Grenzwertbestimmung in solchen Fällen gar kein Problem ist.

- Interessanter gestaltet sich die Grenzwertbestimmung in dem Fall, wo wir zwar eine ganzrationale Funktion haben, deren Nenner allerdings an der fraglichen Stelle eine Nullstelle hat. In diesem Fall muss genauer geprüft werden, ob der Grenzwert existiert und welchen Wert er gegebenenfalls hat. Wir betrachten zunächst die Funktion $\mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}$, gegeben durch $x \mapsto \frac{x^2 - 2x + 1}{x - 1}$, für $x \rightarrow 1$. Der Nenner hat für $x = 1$ eine Nullstelle, also kann $x = 1$ nicht einfach eingesetzt werden. Wir formen den Term, durch den die Funktion gegeben ist, jedoch weiter umformen, sodass wir dann leichter den Grenzwert bestimmen können. Dafür nutzen wir zunächst die zweite binomische Formel:

$$\lim_{x \rightarrow 1} \frac{x^2 - 2x + 1}{x - 1} = \lim_{x \rightarrow 1} \frac{(x - 1)^2}{x - 1}.$$

Nun stellen wir fest, dass wir für alle $x \neq 1$ den Faktor $x - 1$ aus dem Zähler und Nenner rauskürzen können und erhalten

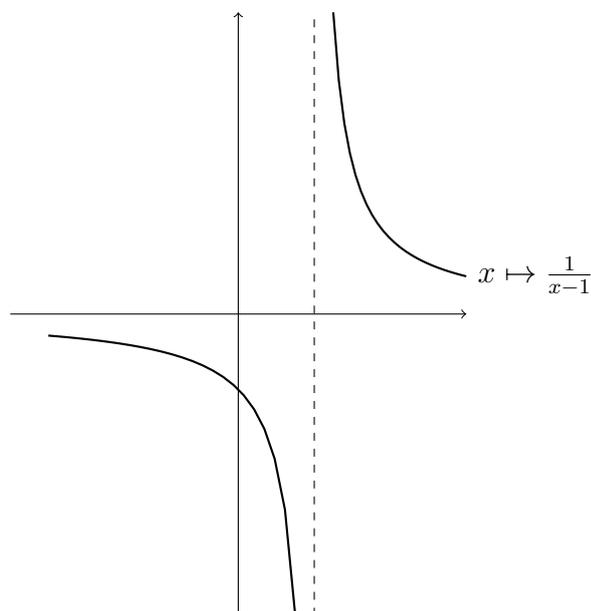
$$\lim_{x \rightarrow 1} \frac{(x - 1)^2}{x - 1} = \lim_{x \rightarrow 1} (x - 1).$$

Hier ist der Nenner nun 1, und wir können den Satz 10.8 anwenden, um den Grenzwert zu bestimmen. Wir erhalten also:

$$\lim_{x \rightarrow 1} \frac{x^2 - 2x + 1}{x - 1} = \lim_{x \rightarrow 1} (x - 1) = 1 - 1 = 0.$$

In diesem Fall hat die Funktion also einen Grenzwert für $x \rightarrow 1$, und zwar den Wert 0.

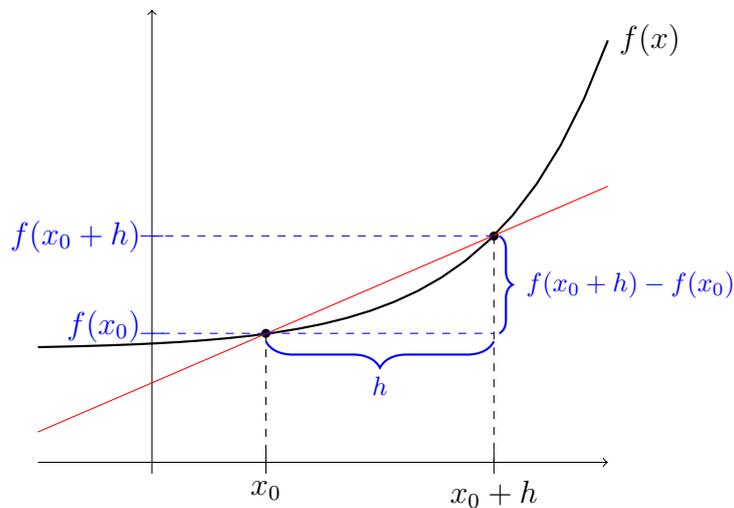
- Nun betrachten wir die Funktion $\mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}$, gegeben durch $x \mapsto \frac{1}{x-1}$, und suchen wieder den Grenzwert für $x \rightarrow 1$. Dieser existiert jedoch nicht, da die Funktion für $x \rightarrow 1$, $x > 1$ über alle Schranken wächst und für $x \rightarrow 1$, $x < 1$, zwar negativ ist, aber betragsmäßig über alle Schranken wächst. Im Bild sieht es fast genauso aus wie für die Funktion $x \rightarrow \frac{1}{x}$, allerdings um einen verschoben, also in etwa:



11 Die Ableitung einer Funktion

Nun können wir Grenzwerte von Funktionen benutzen, um eine präzise Definition der Ableitung zu geben.

Zur Motivation erinnern wir uns wieder daran, dass die Ableitung einer Funktion f im Punkt x_0 die Steigung der Tangente im Punkt x_0 angeben soll, und wir diese Steigung durch die Steigung der Sekante annähern wollen. Wir betrachten nun die Sekante durch die Punkte des Graphens mit den x -Koordinaten x_0 bzw. $x_0 + h$, wobei h eine kleine, positive oder negative reelle Zahl sein soll. Um die Steigung der Sekante zu bestimmen, müssen wir ein Steigungsdreieck an der Sekante einzeichnen und darin die Änderung der y - und x -Werte bestimmen. Der Quotient dieser Änderungen ist die Steigung der Gerade. Wir markieren die entsprechenden Größen im Bild.



Die Differenz der y -Werte ist also $f(x_0 + h) - f(x_0)$, während die Differenz der x -Werte nach Definition gerade h ist, sodass die Steigung der Sekante durch den *Differenzenquotienten*

$$\frac{f(x_0 + h) - f(x_0)}{h}$$

gegeben ist. Um die Steigung der Tangente in x_0 zu erhalten, wollen wir h gegen 0 laufen lassen. Dafür benötigen wir unseren neuen Begriff des Grenzwertes. Das motiviert die folgende Definition der Ableitung einer Funktion f an einer Stelle x_0 .

Definition 11.1. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $x_0 \in I$. Eine Funktion $f: I \rightarrow \mathbb{R}$ heißt **differenzierbar** im Punkt $x_0 \in I$, falls der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

in \mathbb{R} existiert. In diesem Fall wird der Wert dieses Grenzwertes $f'(x_0)$ bezeichnet und die **Ableitung** von f im Punkt x_0 genannt. Die Funktion f wird **differenzierbar** auf I genannt, falls sie in jedem Punkt $y \in I$ differenzierbar ist. In diesem Fall kann man die **Ableitungsfunktion** $f': I \rightarrow \mathbb{R}$ betrachten, die jedem $y \in I$ den Wert der Ableitung in y zuordnet.

Es sei angemerkt, dass bei der obigen Grenzwertbildung x_0 als ein Parameter betrachtet werden sollte; in manchen Fällen hat man sogar x_0 ganz explizit vorgegeben. Der Differenzenquotient

$$\frac{f(x_0 + h) - f(x_0)}{h}$$

wird also als Funktion der Variable h aufgefasst, in der x_0 als eine Konstante betrachtet wird.

Wir wollen nun mit dieser Definition die Ableitungsfunktionen einiger einfacher Funktionen bestimmen. Das Ergebnis ist meist bereits aus der Schule bekannt. (Manchmal trifft das auch auf die Herleitung zu.)

Beispiel. • Wir beginnen mit der Ableitung einer sehr einfachen Funktion. Wir betrachten die Funktion $\varphi: \mathbb{R} \rightarrow \mathbb{R}$, $\varphi(x) = x$. Unser Ziel ist es, die Ableitung dieser Funktion an einer beliebigen Stelle $x_0 \in \mathbb{R}$ zu bestimmen. Dafür setzen wir zunächst die Definition der Ableitung ein:

$$\varphi'(x_0) = \lim_{h \rightarrow 0} \frac{\varphi(x_0 + h) - \varphi(x_0)}{h}.$$

Nun setzen wir auch die Definition unserer konkreten Abbildung φ ein:

$$\lim_{h \rightarrow 0} \frac{\varphi(x_0 + h) - \varphi(x_0)}{h} = \lim_{h \rightarrow 0} \frac{(x_0 + h) - x_0}{h}.$$

Den Term formen wir nun um:

$$\lim_{h \rightarrow 0} \frac{(x_0 + h) - x_0}{h} = \lim_{h \rightarrow 0} \frac{h}{h}.$$

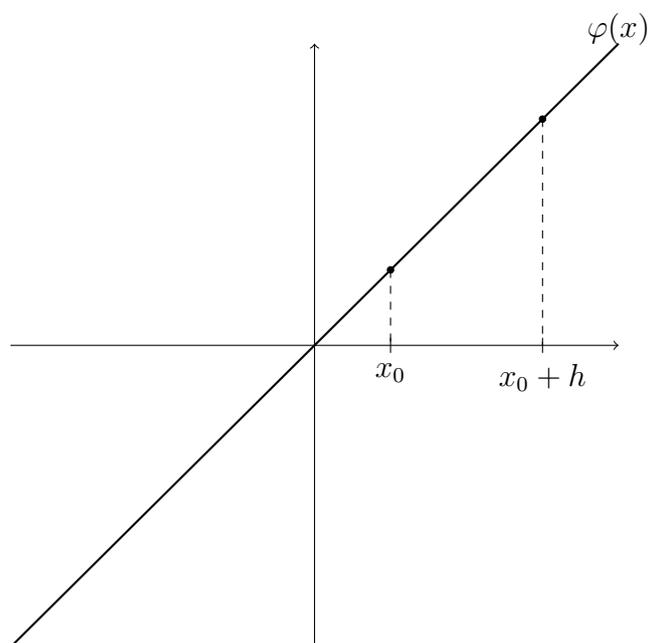
Für alle $h \neq 0$ können wir diesen Bruch kürzen und erhalten

$$\lim_{h \rightarrow 0} \frac{h}{h} = \lim_{h \rightarrow 0} 1.$$

Nun hängt diese Funktion ($h \mapsto 1$) gar nicht von h ab, und deren Grenzwert für $h \rightarrow 0$ ist, genau wie man vermuten würde, ebenfalls 1. (Alternativ kann hier der Satz 10.8 angewandt werden.) Also erhalten wir insgesamt: $\varphi'(x_0) = 1$ für alle $x_0 \in \mathbb{R}$. Die Tangentensteigung ist also 1 an jedem Punkt, und hängt nicht von der Wahl des Punktes ab. Die Ableitungsfunktion der Funktion φ ist also gegeben durch

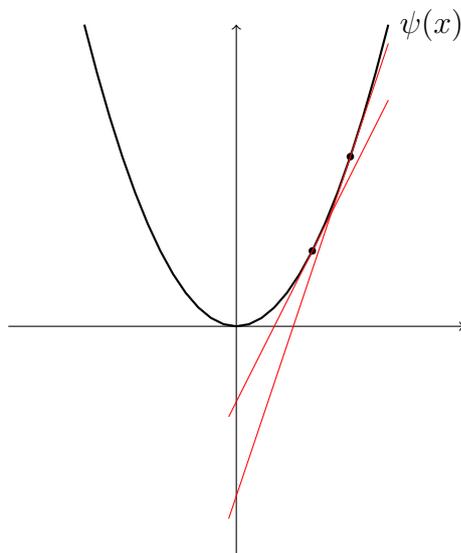
$$\begin{aligned} \varphi': \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto 1. \end{aligned}$$

Betrachtet man die Sache nochmal an dem Funktionsgraphen, so merkt man, dass der Begriff „Tangente“ in diesem Fall rein formal zu verstehen ist. Der Graph der Funktion φ ist nämlich eine Gerade, nämlich die Diagonale des ersten und dritten Quadranten:



In diesem Fall ist insbesondere die Sekante durch zwei Punkte des Graphen genau der Funktionsgraph, und das muss dann auch für die Tangenten stimmen. Alle diese Geraden sind also in diesem langweiligen Spezialfall dieselbe Gerade mit Steigung 1, unabhängig von dem Punkt des Graphen, den wir betrachten.

- Als nächstes betrachten wir ein etwas spannenderes Beispiel. Wir wollen die Ableitung der Funktion $\psi: \mathbb{R} \rightarrow \mathbb{R}$, gegeben durch $\psi(x) = x^2$, für alle $x_0 \in \mathbb{R}$ bestimmen. Hier sieht man bereits am Funktionsgraphen, der ja eine Parabel ist, dass die Tangenten abhängig von der Stelle, an der sie eingezeichnet werden, unterschiedliche Steigungen haben:



Wir bestimmen nun die Ableitung. Dafür setzen wir wieder zunächst die Definition der Ableitung ein:

$$\varphi'(x_0) = \lim_{h \rightarrow 0} \frac{\psi(x_0 + h) - \psi(x_0)}{h}.$$

Nun setzen wir auch die Definition unserer konkreten Abbildung ψ ein:

$$\lim_{h \rightarrow 0} \frac{\psi(x_0 + h) - \psi(x_0)}{h} = \lim_{h \rightarrow 0} \frac{(x_0 + h)^2 - x_0^2}{h}.$$

Den Term formen wir nun um. Hier sieht man auch, dass man nicht einfach $\psi(h)$ im Zähler erhält, da die Abbildung, die wir ableiten wollen, im Allgemeinen Summen nicht auf Summen abbildet. Wir verwenden die erste binomische Formel:

$$\lim_{h \rightarrow 0} \frac{(x_0 + h)^2 - x_0^2}{h} = \lim_{h \rightarrow 0} \frac{x_0^2 + 2x_0h + h^2 - x_0^2}{h} = \lim_{h \rightarrow 0} \frac{2x_0h + h^2}{h}.$$

Für alle $h \neq 0$ können wir diesen Bruch kürzen (beim Kürzen aus der Summe müssen wir aus *jedem* Summanden einen Faktor h herauskürzen) und erhalten

$$\lim_{h \rightarrow 0} \frac{2x_0h + h^2}{h} = \lim_{h \rightarrow 0} (2x_0 + h).$$

Hier kann wieder der Satz 10.8 angewandt werden, und wir erhalten

$$\lim_{h \rightarrow 0} (2x_0 + h) = 2x_0.$$

Also erhalten wir insgesamt: $\psi'(x_0) = 2x_0$ für alle $x_0 \in \mathbb{R}$. Die Tangensteigung ist hier also durchaus von der Wahl des Punktes abhängig, an dem die Tangente angelegt wird. Die Ableitungsfunktion der Funktion ψ ist gegeben durch

$$\begin{aligned} \psi': \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto 2x, \end{aligned}$$

in Übereinstimmung mit den uns bereits bekannten Ableitungsregeln.

- Wir wollen ein weiteres Beispiel sehen, wo die Umformungen, die beim Berechnen des Grenzwertes durchgeführt werden, Wir betrachten die Funktion $\gamma: (0, \infty) \rightarrow \mathbb{R}$, die durch die Vorschrift $\gamma(x) = \frac{1}{x+1}$ festgelegt ist. Der Definitionsbereich ist etwas willkürlich gewählt, allerdings ist es wichtig, dass -1 nicht darin enthalten ist, weil an diesem Punkt die Funktion nicht definiert ist (und, nebenbei bemerkt, auch keinen Grenzwert für $x \rightarrow -1$ besitzt). Zur Bestimmung der Ableitung an

einem beliebigen Punkt $x_0 \in (0, \infty)$ gehen wir wie zuvor vor. Wir setzen wieder die Definition der Ableitung ein:

$$\varphi'(x_0) = \lim_{h \rightarrow 0} \frac{\gamma(x_0 + h) - \gamma(x_0)}{h}.$$

Nun setzen wir auch die Definition unserer konkreten Abbildung γ ein:

$$\lim_{h \rightarrow 0} \frac{\gamma(x_0 + h) - \gamma(x_0)}{h} = \lim_{h \rightarrow 0} \frac{\frac{1}{x_0+h+1} - \frac{1}{x_0+1}}{h}.$$

Den Term formen wir nun um. Dafür bringen wir beide Brüche im Zähler auf einen gemeinsamen Nenner und erinnern uns dann, dass man einen Bruch durch eine Zahl h teilt, indem man den Nenner dieses Bruchs mit h multipliziert:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\frac{1}{x_0+h+1} - \frac{1}{x_0+1}}{h} &= \lim_{h \rightarrow 0} \frac{\frac{x_0+1}{(x_0+h+1)(x_0+1)} - \frac{x_0+h+1}{(x_0+h+1)(x_0+1)}}{h} \\ &= \lim_{h \rightarrow 0} \frac{\frac{(x_0+1)-(x_0+h+1)}{(x_0+h+1)(x_0+1)}}{h} \\ &= \lim_{h \rightarrow 0} \frac{(x_0+1) - (x_0+h+1)}{(x_0+h+1)(x_0+1)h} \\ &= \lim_{h \rightarrow 0} \frac{-h}{(x_0+h+1)(x_0+1)h}. \end{aligned}$$

Für alle $h \neq 0$ können wir diesen Bruch kürzen:

$$\lim_{h \rightarrow 0} \frac{-h}{(x_0+h+1)(x_0+1)h} = \lim_{h \rightarrow 0} \frac{-1}{(x_0+h+1)(x_0+1)}.$$

Hier kann wieder der Satz 10.8 angewandt werden, und wir erhalten

$$\lim_{h \rightarrow 0} \frac{-1}{(x_0+h+1)(x_0+1)} = -\frac{1}{(x_0+1)^2}.$$

Also sehen wir insgesamt, dass $\gamma'(x_0) = -\frac{1}{(x_0+1)^2}$ für alle $x_0 \in (0, \infty)$. Die Ableitungsfunktion der Funktion γ ist gegeben durch

$$\begin{aligned} \gamma' : (0, \infty) &\rightarrow \mathbb{R} \\ x &\mapsto -\frac{1}{(x+1)^2}. \end{aligned}$$

Das ist genau dasselbe Ergebnis, wie das, was man erhält, wenn man die Funktion nach Kettenregel mit äußerer Funktion $u: (0, \infty) \rightarrow \mathbb{R}$, $u(x) = \frac{1}{x}$ und innerer Funktion $v: (0, \infty) \rightarrow (0, \infty)$, $v(x) = x+1$, ableitet.

Damit wollen wir die erste Behandlung der Ableitungen abschließen. Aus der nun präzise gemachten Definition lassen sich die aus der Schule bekannten Ableitungsregeln, wie etwa Kettenregel und Produktregel, herleiten, und auch Ableitungen weiterer Funktionen. Wir wollen diese Beweise nur noch in Einzelfällen ausführen.

12 Exponentialfunktion und Logarithmus

Aus der Schule ist häufig bereits die Eulersche Zahl e bekannt. Wir wollen eine Definition dieser Zahl geben und auch einige Eigenschaften der zugehörigen Exponentialfunktion studieren. Leider werden wir aus Zeitgründen nicht den historischen Zugang zu der Zahl e wählen können. Eine wesentliche historische Motivation, um die Zahl e zu studieren, war die Untersuchung vom Zinseszins. Dabei beschäftigt man sich in etwa mit der folgenden Frage: Man hat 1000 Euro auf einem Konto angelegt, das mit 1% jährlich verzinst wird und wo die Zinsen stets am 1. Januar ausgezahlt werden. Nach dem ersten Jahr erhält man also 10 Euro auf sein Konto ausgezahlt, und man lässt dieses Geld auf dem Konto. Nun erhält man im Folgejahr bereits mehr Zinsen, da diese 10 Euro mitverzinst werden, und man erhält 10,1 Euro dazu. Wie viel Geld hat man dann nach 100 Jahren auf seinem Konto, bei gleichbleibenden Bedingungen? Es stellt sich raus, dass dieser Betrag, circa 2704 Euro, recht nah am 1000-fachem der Zahl e liegt.

Wir werden allerdings einen etwas anderen Zugang zu der Eulerschen Zahl wählen, der mit der Ableitungsfunktion zu tun hat. Dafür fragen wir uns, welche Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$ der Gleichung $f' = f$ genügen.

Zunächst bemerken wir, dass kein Polynom dieser Gleichung genügen kann. Betrachtet man ein Polynom vom Grad n , also etwa

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0,$$

so ist die Ableitung von diesem gegeben durch

$$P'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \dots + 2 a_2 x + a_1.$$

Diese ist zwar wieder ein Polynom, doch vom Grad $n-1$, und kann somit nicht gleich P sein.

Das Problem bei Polynomen war, dass man vom höchsten Grad n zwar einen Term in P hat, aber keinen in P' bekommt. Nun machen wir den *Potenzreihenansatz* und tun so, informell gesprochen, als ob wir ein „unendliches“ Polynom hätten, also eine Potenzreihe der Form $\sum_{n=0}^{\infty} a_n x^n$ für eine feste Folge $(a_n)_{n \in \mathbb{N}}$.

Diese Potenzreihe leiten wir nun ab, und zwar so, wie wir auch ein Polynom ableiten würden. Danach erhalten wir wieder ein Potenzreihe, die wir mit der ursprünglichen vergleichen. Dabei gehen wir wie bei Polynomen vor und vergleichen die Koeffizienten der jeweiligen Potenzen von x . Dabei verwenden wir sowohl die formale als auch die informellere Schreibweise, um die Argumentation transparenter zu gestalten. Dabei benutzen wir die Summen-

regel für die Ableitung sowie $(a_n x^n)' = n a_n x^{n-1}$.

$$\begin{aligned}\varphi(x) &= \sum_{n=0}^{\infty} a_n x^n &= a_0 &+ a_1 x^1 &+ a_2 x^2 &+ a_3 x^3 &+ a_4 x^4 &+ \dots \\ \varphi'(x) &= \sum_{n=0}^{\infty} n a_n x^{n-1} &= a_1 &+ 2a_2 x^1 &+ 3a_3 x^2 &+ 4a_4 x^3 &+ 5a_5 x^4 &+ \dots\end{aligned}$$

Über den Koeffizienten a_0 können wir aus diesem Koeffizientenvergleich keinerlei Informationen erhalten. Deshalb setzen wir, etwas willkürlich, $a_0 = 1$, was sich als eine hilfreiche Festlegung erweist. Vergleicht man nun die Koeffizienten von x^0 , also die Terme ohne x , in beiden Potenzreihen, so erhält man:

$$a_1 = a_0,$$

und da $a_0 = 1$ gesetzt wurde, muss auch $a_1 = 1$ gelten. Wir fahren fort und vergleichen als nächstes die Koeffizienten von x^1 . Daraus erhalten wir

$$2a_2 = a_1.$$

Diese Gleichung können wir nun nach a_2 auflösen und den bereits ermittelten Wert $a_1 = 1$ darin einsetzen. Das liefert

$$a_2 = \frac{a_1}{2} = \frac{1}{2}.$$

Für den Koeffizienten von x^2 erhalten wir genauso:

$$3a_3 = a_2.$$

Diese Gleichung können wir wiederum nach a_3 auflösen und den bereits ermittelten Wert $a_2 = \frac{1}{2}$ darin einsetzen. Das liefert

$$a_3 = \frac{a_2}{3} = \frac{1}{2 \cdot 3} = \frac{1}{6}.$$

Wir sehen uns noch x^3 an, damit das Muster besser erkennbar wird:

$$\begin{aligned}4a_4 &= a_3 \\ \Rightarrow a_4 &= \frac{a_3}{4} = \frac{1}{2 \cdot 3 \cdot 4} = \frac{1}{24}.\end{aligned}$$

Nun lässt sich bereits erahnen, dass im Nenner von a_n jeweils das Produkt von allen natürlichen Zahlen von 1 bis n stehen wird. Für dieses haben wir bereits in Mathematischen Grundlagen 1 bereits eine Notation festgelegt, $n!$ für n -Fakultät, also

$$n! \stackrel{\text{Def}}{=} 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-1) \cdot n.$$

Verwendet man diese Notation, so liegt die Vermutung nahe, dass hier $a_n = \frac{1}{n!}$ gelten muss. Um das zu formalisieren, würde man einen Beweis mittels vollständiger Induktion führen müssen. Das Herzstück des Beweises, also der Induktionsschritt, würde dann aus der Beobachtung bestehen, dass aus $a_{n-1} = \frac{1}{(n-1)!}$ und dem Koeffizientenvergleich für x^{n-1} folgt:

$$a_n = \frac{a_{n-1}}{n} = \frac{1}{(n-1)! \cdot n} = \frac{1}{n!},$$

was gerade das gewünschte Ergebnis ist.

Folglich muss die Funktion $\varphi(x) = \sum_{n=0}^{\infty} a_n x^n$, falls wir $\varphi'(x) = \varphi(x)$ und $a_0 = 1$ verlangen wollen, die Form $\varphi(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ haben.

Nun erinnern wir uns daran, was das genau heißt. Für jede reelle Zahl x , beispielsweise für $x = 2$, soll $\varphi(x)$ genau der Grenzwert der Folge $\left(\sum_{n=0}^k \frac{x^n}{n!} \right)_{k \in \mathbb{N}}$. In Formeln heißt das, allgemein und für $x = 2$ eingesetzt:

$$\begin{aligned} \varphi(x) &= \sum_{n=0}^{\infty} \frac{x^n}{n!} = \lim_{k \rightarrow \infty} \sum_{n=0}^k \frac{x^n}{n!} \\ \varphi(2) &= \sum_{n=0}^{\infty} \frac{2^n}{n!} = \lim_{k \rightarrow \infty} \sum_{n=0}^k \frac{2^n}{n!} \end{aligned}$$

Wir wissen bereits, dass der Grenzwert einer Reihe nicht unbedingt existieren muss, aber wir haben auch am Beispiel der geometrischen Reihe gesehen, dass ein solcher Grenzwert existieren kann, obwohl die Folgenglieder immer größer werden. Auch hier ist es der Fall, uns zwar für beliebige Werte von $x \in \mathbb{R}$. Ein wesentlicher Grund dafür ist die Tatsache, dass die Fakultäten sehr schnell wachsen, und dementsprechend die Summanden der Form $\frac{x^n}{n!}$ schnell sehr klein werden. Somit erhalten wir tatsächlich für jedes $x \in \mathbb{R}$ eine reelle Zahl als Wert für $\varphi(x)$, und es lassen sich einfach Approximationen für $\varphi(x)$ berechnen (und zwar nur mit den Grundrechenarten), indem man etwa $\sum_{n=0}^{20} \frac{x^n}{n!}$ oder $\sum_{n=0}^{30} \frac{x^n}{n!}$ berechnet.

Der folgende Satz fasst zusammen, was wir bis jetzt erreicht haben:

Satz 12.1. *Die Vorschrift*

$$\begin{aligned} \varphi: \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \sum_{n=0}^{\infty} \frac{x^n}{n!} \end{aligned}$$

definiert eine Abbildung. Diese Abbildung hat die Eigenschaft $\varphi'(x) = \varphi(x)$ für alle $x \in \mathbb{R}$.

Nun haben wir also eine Funktion, deren Ableitungsfunktion die Funktion selbst ist, wissen allerdings noch überhaupt nicht, warum diese Funktion eine Exponentialfunktion sein soll, also warum $\varphi(x)$ gerade dasselbe wie die Potenz e^x irgendeiner mysteriösen Zahl e sein soll. Das wollen wir als nächstes skizzieren. Dafür werden wir allerdings einige Vorüberlegungen anstellen müssen.

Zunächst machen wir die Beobachtung, dass man zwar im Allgemeinen für $\varphi(x)$ gut Näherungswerte erhalten kann, die präzisen Werte allerdings häufig schwer zu erhalten sind. Dabei gibt es eine wesentliche Ausnahme, nämlich bei der Berechnung von $\varphi(0)$. Wir betrachten einige der ersten Glieder der Folge $\left(\sum_{n=0}^k \frac{x^n}{n!}\right)_{k \in \mathbb{N}}$ für $x = 0$, um danach den Grenzwert dieser Folge zu bestimmen. Dabei erhalten wir:

$$\begin{aligned} \sum_{n=0}^0 \frac{x^n}{n!} &= \frac{x^0}{0!} = 1, \\ \sum_{n=0}^1 \frac{x^n}{n!} &= 1 + \frac{0^1}{1!} = 1 + 0 = 1, \\ \sum_{n=0}^2 \frac{x^n}{n!} &= 1 + \frac{0^2}{2!} = 1 + 0 = 1, \\ \sum_{n=0}^3 \frac{x^n}{n!} &= 1 + \frac{0^3}{3!} = 1 + 0 = 1, \\ &\dots \end{aligned}$$

Wir bemerken, dass das k -te Folgenglied sich für beliebiges x vom $(k-1)$ -ten durch den Summanden $\frac{x^k}{k!}$ unterscheidet. Für $x = 0$ ist jeder dieser Summanden für $k > 0$ bereits Null, also erhalten wir eine Folge, deren alle Folgenglieder gleich 1 sind. Wir haben uns bereits überlegt, dass diese Folge gegen 1 konvergiert, also gilt:

$$\varphi(0) = \lim_{k \rightarrow \infty} \sum_{n=0}^k \frac{0^n}{n!} = \lim_{k \rightarrow \infty} 1 = 1.$$

Folglich ist $\varphi(0) = 1$.

Das passt bereits ganz gut zu der Vermutung, dass $\varphi(x)$ genau die x -te Potenz irgendeiner positiven reellen Zahl ist, denn 0-te Potenz einer positiven reellen Zahl ist stets 1. Da $a^1 = a$ für jede positive reelle Zahl a gilt, ist es sinnvoll, sich im besonderen für $\varphi(1)$ zu interessieren. Diesen Wert können wir nicht so einfach wie $\varphi(0)$ angeben, spielt allerdings große Rolle, denn es stellt sich heraus, dass dies gerade die Zahl e ist. Wir legen das als die Definition von e fest.

Definition 12.2. Die **Eulersche Zahl** e ist definiert als der Grenzwert

$$e = \sum_{n=0}^{\infty} \frac{1^n}{n!} = \lim_{k \rightarrow \infty} \sum_{n=0}^k \frac{1}{n!}.$$

Es sei angemerkt, dass e keine rationale Zahl ist, ihre Dezimaldarstellung also weder abbricht noch periodisch wird; das ist allerdings nicht ganz offensichtlich und bedarf eines Beweises. Durch Berechnen der endlichen Folgenglieder, etwa $\sum_{n=0}^{20} \frac{1}{n!}$, kann man aber gute Näherungen für e erhalten. Wir

bemerken, dass in der Folge $\left(\sum_{n=0}^k \frac{1}{n!}\right)_{k \in \mathbb{N}}$ immer ein positiver Summand zu dem vorherigen Folgenglied hinzukommt, um das nächste Folgenglied zu erhalten, und die Folgenglieder hier immer größer werden. Berechnet man also die ersten Folgenglieder, so weiß man definitiv, dass die Zahl e größer sein wird als diese. Dadurch erhält man die folgende einfache Abschätzung:

$$e > \sum_{n=0}^2 \frac{1}{n!} = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} = 1 + 1 + \frac{1}{2} = 2,5.$$

Durch Aufsummieren von weiteren Summanden, etwa mit dem Taschenrechner, kann man weitere Stellen von e ausrechnen. Auf 5 Nachkommastellen genau gilt: $e \approx 2,71828$.

Nun wollen wir uns wieder der Frage widmen, warum $\varphi(x)$ gerade mit der Funktion übereinstimmt, die jeder reellen Zahl $x \in \mathbb{R}$ die Potenz e^x zuordnet. Dafür erinnern wir uns zunächst an eine wesentliche Eigenschaft der Exponentialfunktionen, nämlich an die Potenzgesetze.

Potenzgesetze. Für jede positive reelle Zahl $a \in \mathbb{R}$ und beliebige reelle Zahlen $x, y \in \mathbb{R}$ gilt:

- Die 0-te Potenz einer beliebigen positiven reellen Zahl ist 1, in Zeichen:

$$a^0 = 1.$$

- Potenzen mit gleicher Basis werden multipliziert, indem man die Exponenten addiert, in Zeichen:

$$a^x \cdot a^y = a^{x+y}.$$

- Potenzen werden potenziert, indem man die Exponenten multipliziert, in Zeichen:

$$(a^x)^y = a^{x \cdot y}.$$

Es stellt sich heraus, dass die ersten beiden Eigenschaften im Wesentlichen bereits eine Funktion als eine Exponentialfunktion festlegen. Mit anderen Worten ist jede „gute“ Funktion, die Summen auf Produkte schickt und 0 auf 1 schickt, bereits eine Exponentialfunktion. Die richtige Forderung für „gut“ wäre „stetig“. In der Schule wird Stetigkeit etwas vereinfacht mit „lässt sich mit dem Stift durchzeichnen“ beschrieben. Wir könnten es in Termen von Grenzwerten von Funktionen formulieren; in etwa heißt es, dass der

Grenzwert der Funktion für $x \rightarrow a$ mit dem Wert der Funktion in a übereinstimmt, falls sie in a definiert ist. Das trifft, wie wir bereits erwähnt haben, auf fast alle Funktionen zu, die wir kennen.

Präziser lässt sich die obige Aussage im folgenden Satz formulieren:

Satz 12.3. *Sei $\psi: \mathbb{R} \rightarrow \mathbb{R}$ eine „gute“ (präziser: stetige) Funktion, die die folgenden beiden Eigenschaften besitzt:*

- $\psi(0) = 1$,
- Für alle reellen Zahlen x, y gilt: $\psi(x + y) = \psi(x) \cdot \psi(y)$.

Dann ist die Funktion ψ bereits eine Exponentialfunktion, d.h., es gibt eine reelle Zahl b , sodass für alle reellen Zahlen $x \in \mathbb{R}$ gilt: $\psi(x) = b^x$. Die Zahl b lässt sich als $b = \psi(1)$ bestimmen.

Der vollständige Beweis von diesem Satz ist etwas komplizierter. Was sich jedoch relativ einfach durch vollständige Induktion zeigen lässt, ist die Tatsache, dass $\psi(n) = \psi(1)^n$ ist. Wir führen auch das nicht aus, skizzieren allerdings das Vorgehen an einigen Beispielen. Sei also $\psi: \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion, die die beiden obigen Eigenschaften besitzt. Wir wollen zunächst $\psi(2)$ ausrechnen. Dabei können wir $2 = 1 + 1$ nutzen und die zweite Eigenschaft von ψ anwenden; wir erhalten so:

$$\psi(2) = \psi(1 + 1) = \psi(1) \cdot \psi(1) = \psi(1)^2.$$

Will man nun weiter $\psi(3)$ ausrechnen, so verwendet man erneut die zweite Eigenschaft, jetzt zusammen mit der Aufteilung $3 = 2 + 1$ und mit der Tatsache, dass $\psi(2) = \psi(1)^2$ bereits gezeigt wurde:

$$\psi(3) = \psi(2 + 1) = \psi(2) \cdot \psi(1) = \psi(1)^2 \cdot \psi(1) = \psi(1)^3.$$

Dabei merkt man bereits, wie auch der Induktionsschritt funktionieren würde: Weiß man bereits, dass $\psi(n) = \psi(1)^n$ ist, so kann man auf $n + 1$ die zweite Eigenschaft von ψ anwenden und erhält:

$$\psi(n + 1) = \psi(n) \cdot \psi(1) \stackrel{\text{IV}}{=} \psi(1)^n \cdot \psi(1) = \psi(1)^{n+1}.$$

Man kann durch ähnliches Vorgehen, zusammen mit der ersten Eigenschaft von ψ , zeigen, dass $\psi(q) = \psi(1)^q$ für jede rationale Zahl $q \in \mathbb{Q}$ ist. Die Stetigkeit benötigt man dann, um die Aussage auch auf beliebige reelle Exponenten zu übertragen.

Nun kehren wir zu dem Hauptthema dieses Kapitels zurück. Wir haben die Funktion $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ durch $\varphi(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ definiert, die die Eigenschaft $\varphi'(x) = \varphi(x)$ für alle $x \in \mathbb{R}$ hat, und wollen nun sehen, dass diese Funktion andererseits gerade die Vorschrift $x \mapsto e^x$ hat, wobei $e = \varphi(1)$ ist. Dafür

wollen wir im Wesentlichen den obigen Satz verwenden. Wie bereits erläutert, wollen wir uns nicht mit Stetigkeit beschäftigen, aber wir überzeugen uns informell davon, dass die beiden Eigenschaften aus dem Satz für φ erfüllt sind. Die erste Eigenschaft haben wir bereits eingesehen.

Es geht also darum zu beweisen, dass $\varphi(x) \cdot \varphi(y) = \varphi(x + y)$ für alle reellen Zahlen x, y gilt. Wir werden diesen Beweis informell führen. Um den Beweis präzise zu machen, muss man an einigen Stellen die Konvergenz der betrachteten Folgen prüfen; jedoch enthält der Teil, den wir hier vorstellen, die meisten wesentlichen Ideen im Beweis.

Wir fangen mit der linken Seite, $\varphi(x) \cdot \varphi(y)$, an, und formen diese um. Jeder der Terme $\varphi(x)$ bzw. $\varphi(y)$ ist als eine „unendliche“ Summe definiert. Wir stellen uns für die Beweisidee vor, dass es sich hierbei um endliche Summen handelt. Will man das Produkt zweier Summen anders ausdrücken, so multipliziert man aus. Das tun wir in diesem Fall, und dabei sortieren wir gleich die erhaltenen Summanden so, dass wir die Behauptung leichter sehen können. Dabei verwenden wir wieder die informelle Schreibweise anstelle der Summenschreibweise. Wir wollen also

$$\begin{aligned} & \varphi(x) \cdot \varphi(y) \\ &= \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) \cdot \left(\sum_{n=0}^{\infty} \frac{y^n}{n!} \right) \\ &= \left(\frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots \right) \left(\frac{y^0}{0!} + \frac{y^1}{1!} + \frac{y^2}{2!} + \frac{y^3}{3!} + \frac{y^4}{4!} + \dots \right) \\ &= \left(1 + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots \right) \left(1 + \frac{y^1}{1!} + \frac{y^2}{2!} + \frac{y^3}{3!} + \frac{y^4}{4!} + \dots \right) \end{aligned}$$

ausmultiplizieren.

Zunächst bemerken wir, dass der einzige Term nach dem Ausmultiplizieren, der weder einen Faktor x noch einen Faktor y enthält, als $1 \cdot 1 = 1$ entsteht.

Als nächstes fragen wir uns, was alle Terme sind, die nur ein x oder nur ein y enthalten. Das kann nur passieren, wenn wir in der ersten Klammer 1 und in der zweiten Klammer $\frac{y^1}{1!}$ ausgewählt haben, oder wenn wir umgekehrt in der ersten Klammer $\frac{x^1}{1!}$ ausgewählt haben, denn alle anderen Summanden würden höhere Potenzen von x oder y liefern. Diese Summanden ergeben

$$1 \cdot \frac{y^1}{1!} + \frac{x^1}{1!} \cdot 1 = y + x = \frac{(x + y)^1}{1!}.$$

(Der letzte Ausdruck ist zwar nur eine kompliziertere Schreibweise für $x + y$; diese Betrachtungsweise wird sich jedoch als hilfreich erweisen.)

Nun suchen wir nach allen Termen, die x^2 oder y^2 oder xy enthalten, also allen Termen, die in der ersten binomischen Formel vorkommen. Dabei erhalten wir genau die Terme $1 \cdot \frac{y^2}{2!} + \frac{x^2}{2!} \cdot 1$ und $\frac{x^1}{1!} \cdot \frac{y^1}{1!}$. Wir bringen diese

Summanden auf den gemeinsamen Nenner und fassen diese zusammen:

$$1 \cdot \frac{y^2}{2!} + \frac{x^2}{2!} \cdot 1 + \frac{x^1}{1!} \cdot \frac{y^1}{1!} = \frac{y^2 + 2xy + x^2}{2!} = \frac{(x + y)^2}{2!}.$$

Als nächstes suchen wir die Summanden, die zu den Potenzen x^3 und y^3 passen. Dazu gehören noch die Terme, die x^2y und xy^2 enthalten, denn das sind genau die Terme, die in $(x + y)^3$ nach dem Ausmultiplizieren vorkommen. Tatsächlich wissen wir dank dem binomischen Lehrsatz, wie $(x + y)^3$ ausmultipliziert aussieht:

$$\begin{aligned} (x + y)^3 &= \binom{3}{0}x^3 + \binom{3}{1}x^2y^1 + \binom{3}{2}x^1y^2 + \binom{3}{3}y^3 \\ &= x^3 + 3x^2y + 3xy^2 + y^3. \end{aligned}$$

Nun suchen wir also in dem Term für $\varphi(x) \cdot \varphi(y)$ die erwähnten Summanden. Dabei erhalten wir:

$$1 \cdot \frac{y^3}{3!} + \frac{x^1}{1!} \cdot \frac{y^2}{2!} + \frac{x^2}{2!} \cdot \frac{y^1}{1!} + \frac{x^3}{3!} \cdot 1.$$

Wir bringen die Summanden auf den gemeinsamen Nenner $3!$, indem wir die beiden mittleren Summanden mit dem Faktor $3!$ erweitern:

$$1 \cdot \frac{y^3}{3!} + \frac{x^1}{1!} \cdot \frac{y^2}{2!} + \frac{x^2}{2!} \cdot \frac{y^1}{1!} + \frac{x^3}{3!} \cdot 1 = \frac{y^3 + \frac{3!}{1!2!}xy^2 + \frac{3!}{2!1!}x^2y + x^3}{3!}.$$

Nun erinnern wir uns an die aus Mathematischen Grundlagen 1 (und manchmal bereits aus der Schule) bekannte Formel für Binomialkoeffizienten:

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

Für $n = 3$ und $k = 1$ bzw. $k = 2$ erhalten wir gerade

$$\binom{3}{1} = \frac{3!}{1!(3 - 1)!} \text{ und } \binom{3}{2} = \frac{3!}{2!(3 - 2)!}.$$

Das sind die Koeffizienten, die wir soeben beim Ausmultiplizieren von $\varphi(x) \cdot \varphi(y)$ erhalten haben, also bekommen wir aus den Summanden zu x^3 , y^3 , x^2y , xy^2 :

$$\begin{aligned} \frac{y^3 + \frac{3!}{1!2!}xy^2 + \frac{3!}{2!1!}x^2y + x^3}{3!} &= \frac{\binom{3}{0}y^3 + \binom{3}{1}xy^2 + \binom{3}{2}x^2y + \binom{3}{3}x^3}{3!} \\ &= \frac{(x + y)^3}{3!}. \end{aligned}$$

Nun kann man sich vorstellen, wie man mit Hilfe des allgemeinen binomischen Lehrsatzes, den wir in Mathematischen Grundlagen 1 kennengelernt haben,

weiter Summanden zusammenfassen kann. Der Term der Form $x^k y^m$ kommt beim ausmultiplizieren nur einmal vor, und zwar in dem Produkt

$$\frac{x^k}{k!} \cdot \frac{y^m}{m!} = \frac{1}{k!m!} x^k y^m.$$

Diesen Bruch kann man nun wieder mit $(k+m)!$ erweitern und erhält:

$$\frac{1}{k!m!} x^k y^m = \frac{1}{(k+m)!} \frac{(k+m)!}{k!} \frac{1}{m!} x^k y^m = \frac{1}{(k+m)!} \binom{k+m}{k} x^k y^m.$$

Fasst man nun alle Summanden beim Ausmultiplizieren von $\varphi(x) \cdot \varphi(y)$ zusammen, bei denen die Summe der Exponenten von x und y dieselbe ist, beispielsweise n (in obigen Beispielen war diese Summe 0, 1, 2 bzw. 3), so erhält man gerade das Produkt von $\frac{1}{n!}$ mit dem Term aus dem binomischen Lehrsatz, nämlich

$$\frac{1}{n!} \cdot \left(y^n + \binom{n}{1} x^1 y^{n-1} + \binom{n}{2} x^2 y^{n-2} + \dots + \binom{n}{n-1} x^{n-1} y^1 + x^n \right).$$

Also lassen sich diese Summanden gerade zu

$$\frac{(x+y)^n}{n!}$$

zusammenfassen. Multipliziert man $\varphi(x) \cdot \varphi(y)$ also aus, so erhält man

$$\varphi(x) \cdot \varphi(y) = 1 + \frac{(x+y)^1}{1!} + \frac{(x+y)^2}{2!} + \frac{(x+y)^3}{3!} + \frac{(x+y)^4}{4!} + \dots$$

Auf der rechten Seite steht aber gerade die Definition von $\varphi(x+y)$ (in der informellen Schreibweise).

Die obige Rechnung liefert die grundlegende Idee dafür, warum die Funktion φ Summen in Produkte überführt. Präzise gemacht, zeigt dieses Argument also

$$\varphi(x+y) = \varphi(x) \cdot \varphi(y)$$

für alle reellen Zahlen $x, y \in \mathbb{R}$.

Zusammen mit dem Satz 12.3 impliziert diese Tatsache (zusammen damit, dass die Funktion φ stetig ist, und dass wir bereits $\varphi(0) = 1$ wissen) nun, dass die Funktion φ tatsächlich eine Exponentialfunktion ist, und mit $e = \varphi(1)$ gilt: $\varphi(x) = e^x$ für alle $x \in \mathbb{R}$. Ausgeschrieben heißt das:

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

für alle $x \in \mathbb{R}$. Insbesondere haben wir damit gezeigt, dass die Funktion $x \rightarrow e^x$ sich selbst als Ableitungsfunktion hat.

Wir wollen nun weitere Eigenschaften der Exponentialfunktion mit Basis e erforschen. Dafür bemerken wir zunächst, dass in der Summe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ für eine positive reelle Zahl x jeder Summand positiv ist (und für eine negative reelle Zahl x die Summanden abwechselnd positiv und negativ sind). Somit ist der Grenzwert der Reihe für positives x größer als jede endliche Summe $\sum_{n=0}^k \frac{x^n}{n!}$, da die Folgenglieder dieser Reihe immer größer werden. Insbesondere gilt für jedes positive x :

$$\varphi(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} > \sum_{n=0}^0 \frac{x^n}{n!} = 1.$$

Das passt auch damit zusammen, dass die positiven Potenzen der Zahl $e > 1$ stets größer als 1 sein müssen.

Ferner sind beliebige Potenzen der positiven Zahl e wieder positiv, also gilt $e^x > 0$ für alle $x \in \mathbb{R}$. Man kann also die Exponentialfunktion zur Basis e als eine Abbildung von \mathbb{R} in die positiven reellen Zahlen auffassen, in Zeichen

$$\begin{aligned} \mathbb{R} &\rightarrow (0, \infty) \\ x &\mapsto e^x. \end{aligned}$$

Diese Abbildung hat sehr gute Eigenschaften. Insbesondere ist sie bijektiv, wie wir im nächsten Satz festhalten wollen:

Satz 12.4. *Die Abbildung*

$$\begin{aligned} \mathbb{R} &\rightarrow (0, \infty) \\ x &\mapsto e^x. \end{aligned}$$

ist bijektiv.

Beweis der Injektivität. Wir erinnern uns, dass eine Abbildung nach Definition genau dann bijektiv ist, wenn sie surjektiv und injektiv ist. Der Nachweis der Surjektivität dieser Abbildung ist etwas schwieriger, und darauf wollen wir nicht weiter eingehen. Hingegen lässt sich die Injektivität sehr gut mit unseren Mitteln beweisen.

Wir müssen zeigen, dass aus $x \neq y$ für zwei reelle Zahlen x, y auch $e^x \neq e^y$ folgt. Seien nun x, y zwei verschiedene reelle Zahlen. Davon muss also eine größer sein als die andere, wir legen $x > y$ fest. Wir führen einen Widerspruchsbeweis und nehmen an, dass $e^x = e^y$ gilt. Nun wollen wir diese Annahme zum Widerspruch führen. Wir erinnern uns, dass e^{-y} eine reelle Zahl $\neq 0$ ist (sogar positiv). Also ist es eine Äquivalenzumformung, eine Gleichung mit e^{-y} zu multiplizieren, und wir erhalten unter Verwendung der Potenzge-

setze:

$$\begin{aligned}
 & e^x = e^y \quad | \cdot e^{-y} \\
 \Leftrightarrow & e^x \cdot e^{-y} = e^y \cdot e^{-y} \\
 \Leftrightarrow & e^{x-y} = e^{y-y} \\
 \Leftrightarrow & e^{x-y} = e^0 \\
 \Leftrightarrow & e^{x-y} = 1.
 \end{aligned}$$

(Es sei nochmal daran erinnert, dass e^{-y} gerade $\frac{1}{e^y}$ ist, was also eine weitere Erklärung für $e^y \cdot e^{-y} = 1$ liefert.)

Da wir festgelegt haben, dass $x > y$ ist, ist die Differenz $x - y$ eine positive Zahl. Wir haben uns aber soeben überlegt, dass für diese positive Zahl $e^{x-y} > 1$ gelten muss. Das ist ein Widerspruch zu $e^{x-y} = 1$. Also muss unsere Annahme $e^x = e^y$ falsch gewesen sein. Den verschiedenen reellen Zahlen $x \neq y$ werden also unterschiedliche Werte $e^x \neq e^y$ zugeordnet, und die Abbildung ist somit injektiv, was auch zu beweisen war. \square

Nun haben wir festgestellt, dass die Abbildung

$$\begin{aligned}
 \varphi: \mathbb{R} &\rightarrow (0, \infty) \\
 x &\mapsto e^x.
 \end{aligned}$$

bijektiv ist. Wir wissen, dass eine Abbildung genau dann bijektiv ist, wenn sie eine Umkehrabbildung besitzt. Folglich hat die Exponentialfunktion zur Basis e eine Umkehrabbildung, und diese wird **natürlicher Logarithmus** genannt und mit \ln bezeichnet. Der natürliche Logarithmus ordnet also jeder positiven reellen Zahl eine reelle Zahl zu; für negative reelle Zahlen und für 0 ist der natürliche Logarithmus nicht definiert. Wir erinnern uns daran, dass eine Umkehrabbildung zu einer Abbildung φ dadurch definiert ist, dass die Anwendung der Umkehrabbildung den Effekt von φ aufhebt. In diesem Fall heißt es konkret in Formeln:

$$\begin{aligned}
 e^{\ln x} &= x \text{ für alle } x \in (0, \infty) \text{ und} \\
 \ln(e^y) &= y \text{ für alle } y \in \mathbb{R}.
 \end{aligned}$$

Wendet man also auf eine positive reelle Zahl x den natürlichen Logarithmus an und setzt das Ergebnis als Exponenten für e ein, so erhält man die ursprüngliche Zahl x zurück. Berechnet man umgekehrt für eine beliebige reelle Zahl y die Potenz e^y und wendet den natürlichen Logarithmus auf das Ergebnis an, so erhält man wieder die ursprüngliche Zahl y als Ergebnis.

Aus der Definition vom natürlichen Logarithmus als Umkehrabbildung der Exponentialfunktion zur Basis e lassen sich bereits einige Eigenschaften des natürlichen Logarithmus herleiten. Da wir bereits $e^0 = 1$ wissen, liefert die Definition von \ln :

$$\ln(1) = 0.$$

Ferner lässt sich aus dem Potenzgesetz $e^x \cdot e^y = e^{x+y}$ für alle $x, y \in \mathbb{R}$ die folgende Gesetzmäßigkeit für den natürlichen Logarithmus herleiten:

$$\ln(a \cdot b) = \ln(a) + \ln(b) \text{ für alle } a, b \in (0, \infty).$$

Man kann also sagen, dass die Exponentialfunktion aus Summen Produkte macht, während die Logarithmusfunktion umgekehrt aus Produkten Summen macht.

Zum Abschluss dieses Kapitels wollen wir noch die Ableitung der Logarithmusfunktion $\ln: (0, \infty) \rightarrow \mathbb{R}$ bestimmen. Dabei werden wir die Funktion

$$\begin{aligned} (0, \infty) &\rightarrow \mathbb{R} \\ x &\mapsto e^{\ln x} \end{aligned}$$

auf zwei verschiedene Weisen ableiten und werden daraus die Ableitung des Logarithmus bestimmen können.

Zunächst wenden wir auf diese Funktion einfach die Kettenregel an und bedenken dabei, dass nach unserer Konstruktion die Ableitungsfunktion der Funktion $x \mapsto e^x$ wieder dieselbe Funktion ist. Wenden wir also die Kettenregel hier auf die äußere Funktion $u(y) = e^y$ und innere Funktion $v(x) = \ln(x)$ an, so erhalten wir die Ableitungsfunktion:

$$(e^{\ln x})' = (u \circ v)'(x) = u'(v(x)) \cdot v'(x) = e^{\ln(x)} \cdot \ln'(x).$$

Nun gilt aber nach Definition vom natürlichen Logarithmus $e^{\ln x} = x$ für alle $x \in (0, \infty)$. Die Funktion, die wir ableiten, ist also eigentlich die Funktion $x \mapsto x$, und deren Ableitungsfunktion ist, wie wir uns bereits überlegt haben, die Funktion $x \mapsto 1$. Also haben wir für alle $x \in (0, \infty)$:

$$e^{\ln(x)} \cdot \ln'(x) = 1.$$

Setzt man hier nochmal die Identität $e^{\ln x} = x$ für alle $x \in (0, \infty)$ ein, so erhält man für alle $x \in (0, \infty)$:

$$\begin{aligned} x \cdot \ln'(x) &= 1 \mid : x \neq 0 \\ \Leftrightarrow \ln'(x) &= \frac{1}{x}. \end{aligned}$$

Also ist die Ableitungsfunktion der Logarithmusfunktion gegeben durch die Funktion

$$\begin{aligned} (0, \infty) &\rightarrow \mathbb{R} \\ x &\mapsto \frac{1}{x}. \end{aligned}$$

13 Taylor-Polynome

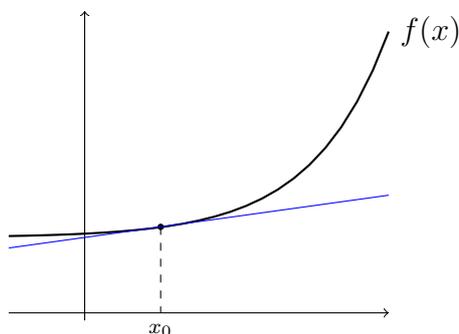
In dem letzten Kapitel haben wir Exponentialfunktion und Logarithmusfunktion kennengelernt. Während die Exponentialfunktion sich durch ihre Definition als Reihe leicht näherungsweise approximieren lässt, wissen wir zunächst nicht, wie wir Näherungen für den Logarithmus berechnen können. Dafür eignen sich beispielsweise Approximationen durch Polynome. Werte von Polynomen beim Einsetzen von reellen Zahlen lassen sich besonders gut ausrechnen, da man nur die Grundrechenarten hierbei braucht.

Dafür gibt es viele verschiedene Methoden. Wir wollen im Rahmen unserer Veranstaltung nur die sogenannten Taylor-Polynome als Approximationen von einer vorgegebenen, beliebig oft differenzierbaren Funktion betrachten. Wir werden nicht näher untersuchen, wie man die Qualität einer Approximation im Allgemeinen oder im Spezialfall der Taylor-Polynome beschreibt.

Das Grundproblem ist also das Folgende. Es ist ein offenes Intervall $I \subseteq \mathbb{R}$ vorgegeben und eine Funktion $f: I \rightarrow \mathbb{R}$, die beliebig oft differenzierbar ist. (Das heißt, dass die Ableitungsfunktion $f': I \rightarrow \mathbb{R}$ existiert und wieder differenzierbar ist, folglich selbst eine Ableitungsfunktion besitzt, die zweite Ableitung $f'': I \rightarrow \mathbb{R}$ von f , und diese wiederum besitzt eine Ableitungsfunktion $f''': I \rightarrow \mathbb{R}$ und so weiter.) Wir fragen uns also, wie man ein Polynom P von einem vorgegebenen Grad n findet, das die Funktion möglichst gut in einer Umgebung einer vorgegebenen Stelle x_0 approximiert.

Bevor wir die Definition der Taylor-Polynome geben, stellen wir zwei motivierende Vorüberlegungen an. Zunächst bemerken wir, dass wir dieser Fragestellung in ganz ähnlicher Form für den Fall $n = 1$ begegnet sind, also für die Approximation durch affin-lineare Funktionen.

Das wollen wir nun etwas genauer ausführen. Gesucht wir also eine Gerade in der Ebene, die an einer festgelegten Stelle $x_0 \in I$ genau denselben Wert wie die Funktion f hat, und dabei die Funktion in der Nähe von x_0 möglichst gut annähert. Es stellt sich heraus, dass eine solche Gerade (in einem präzisen Sinne) durch die Tangente an den Funktionsgraphen im Punkt x_0 vorgegeben ist. Anschaulich ist es ebenfalls plausibel, dass die Tangente eine bessere Approximation als jede Sekante liefert. (Übrigens kann das in ähnlicher Form auch als Definition der Tangente genutzt werden.) Wir wollen uns überlegen, welche Gleichung die Tangente an den Funktionsgraphen im Punkt x_0 hat, denn das stellt sich später als Spezialfall der Taylor-Polynome heraus:



Wir kennen bereits die Steigung der Tangente, denn das war gerade nach Definition die Ableitung der Funktion f im Punkt x_0 , also $f'(x_0)$. Ferner wissen wir, dass die Tangente durch den Punkt $(x_0, f(x_0))$ geht. Daraus können wir auch den Achsenabschnitt b der Geraden bestimmen, denn für die Tangentengleichung der Form

$$y = f'(x_0) \cdot x + b$$

muss insbesondere gelten:

$$f(x_0) = f'(x_0) \cdot x_0 + b.$$

Wir können diese Gleichung nach b auflösen und erhalten: $b = f(x_0) - f'(x_0) \cdot x_0$. Es stellt sich als praktisch heraus, die Geradengleichung etwas umzuformen:

$$\begin{aligned} y &= f'(x_0) \cdot x + (f(x_0) - f'(x_0) \cdot x_0) \\ \Leftrightarrow y &= f(x_0) + f'(x_0) \cdot (x - x_0). \end{aligned}$$

Es sei nochmal angemerkt, dass x_0 eine feste Stelle ist, an der wir den Wert der Funktion f und ihrer Ableitung brauchen, um dann für einen Wert x in der Nähe von x_0 eine affin-lineare Näherung auszurechnen.

Nun kommen wir zu einer zweiten Vorüberlegung zu Taylor-Polynomen. Wir wollen die Koeffizienten von einem Polynom bestimmen, der möglichst nah an unserer Funktion f ist. Dafür betrachten wir zuerst den Fall, dass unsere Funktion selbst bereits ein Polynom ist. Sei also $P(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \dots + a_2 x^2 + a_1 x + a_0$ ein Polynom vom Grad n . Wir werden gleich sehen, wie man durch Ableiten und Einsetzen des speziellen Wertes $x = 0$ alle Koeffizienten von P bestimmen können.

Wir haben bereits gesehen, dass Einsetzen der Null direkt in das Polynom uns den Koeffizienten vom Term x^0 liefert, also

$$P(0) = a_n \cdot 0^n + a_{n-1} \cdot 0^{n-1} + \dots + a_2 \cdot 0^2 + a_1 \cdot 0 + a_0 = a_0.$$

Nun betrachten wir die Ableitung von dem Polynom P . Diese bestimmt sich zu

$$P'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + (n-2) a_{n-2} x^{n-3} + \dots + 2 a_2 x + a_1.$$

Man beachte, dass der konstante Term a_0 des Polynoms P beim Ableiten wegfällt, da die Ableitung einer konstanten Funktion die Nullfunktion ist. Setzt man nun Null in die Ableitungsfunktion von P ein, so erhält man dadurch gerade den Koeffizienten a_1 :

$$P'(0) = na_n \cdot 0^{n-1} + (n-1)a_{n-1} \cdot 0^{n-2} + (n-2)a_{n-2} \cdot 0^{n-3} + \dots + 2a_2 \cdot 0 + a_1 = a_1.$$

Die Vermutung liegt nahe, dass man mit der nächsten Ableitung den Wert des Koeffizienten a_2 des x^2 -Terms vom Polynom P bestimmen kann. Das ist auch richtig, allerdings erhalten wir hier einen zusätzlichen Faktor. Die zweite Ableitung von P ist nämlich gegeben durch

$$P''(x) = n(n-1)a_n \cdot x^{n-2} + (n-1)(n-2)a_{n-1} \cdot x^{n-3} + \dots + 3 \cdot 2a_3x + 2a_2.$$

Setzt man hier $x = 0$ ein, so erhält man diesmal $P''(0) = 2a_2$. Folglich lässt sich a_2 aus $P''(0)$ bestimmen, indem wir diese Gleichung auflösen und dabei

$$a_2 = \frac{P''(0)}{2}$$

erhalten. Bei der nächsten Ableitung, $P'''(0)$, erhalten wir analog $P'''(0) = 3 \cdot 2a_3$ und

$$a_3 = \frac{P'''(0)}{3 \cdot 2}.$$

Um weitere Ableitungen sinnvoll zu bezeichnen, ohne noch mehr Striche hinzuzufügen, führen wir die folgende Notation ein.

Notation. Sei $I \subset \mathbb{R}$ ein offenes Intervall und $f: I \rightarrow \mathbb{R}$ eine beliebig oft differenzierbare Funktion. Wir schreiben $f^{(k)}(x)$ für die k -te Ableitung der Funktion f an der Stelle $x \in I$. (Das sollte nicht mit dem Exponenten

$$f^k(x) = \underbrace{f(x) \cdot f(x) \cdot f(x) \cdot \dots \cdot f(x)}_{k \text{ Faktoren}}$$

verwechselt werden.) Es gilt also beispielsweise:

$$\begin{aligned} f^{(1)}(x) &= f'(x), \\ f^{(2)}(x) &= f''(x), \\ f^{(3)}(x) &= f'''(x). \end{aligned}$$

Es stellt sich ferner als eine sinnvolle Konvention heraus, die Funktion selbst als ihre 0-te Ableitung zu betrachten. Also setzen wir

$$f^{(0)}(x) = f(x).$$

Mit dieser Notation ausgestattet, kommen wir nun zu der Frage zurück, wie wir die weiteren Koeffizienten von P als Auswertungen von höheren Ableitungen von P an der Stelle 0 bestimmen können. Bei der vierten Ableitung stellen wir etwa fest, dass

$$a_4 = \frac{P^{(4)}(0)}{4 \cdot 3 \cdot 2}$$

gilt. Allgemein stellen wir fest (und das kann mittels vollständiger Induktion bewiesen werden), dass wir hier - genau wie bei der Exponentialfunktion - im Nenner stets die Zahlen $k!$ erhalten; genauer gesagt:

$$a_k = \frac{P^{(k)}(0)}{k!}.$$

Man bemerke, dass die Koeffizienten a_k Null werden, wenn $k > n$ ist, also größer als der Grad des Polynoms. Das passt allerdings ausgezeichnet damit zusammen, dass die $(n+1)$ -te Ableitungsfunktion eines Polynoms n -ten Grades die Nullfunktion ist (und somit auch jede k -te Ableitung für $k > n$).

Nun ist der Gedanke bei den Taylor-Polynomen für eine beliebige Funktion $f: I \rightarrow \mathbb{R}$, dass man dieselbe Formel für die Koeffizienten benutzen kann. Das wird uns eine Näherung der Funktion f an der Stelle 0 liefern, sofern diese in 0 definiert ist. Nun hat man allerdings das Problem, dass man manchmal die Näherung an einer anderen Stelle bestimmen will, etwa für den Logarithmus, der in $x_0 = 0$ gar nicht definiert ist. Diese allgemeinere Form der Taylor-Polynome ist etwas komplizierter, weswegen wir sie erst als zweites angeben. Der Grundgedanke der Modifikation, die wir hier durchführen, ist allerdings bereits aus dem Schulunterricht bekannt. Da hat man bereits gelernt, dass während $y = x^2$ die Standardparabel beschreibt, die Gleichung $y = (x - a)^2$ die Parabel beschreibt, die durch horizontale Verschiebung der Standardparabel um a Einheiten nach rechts entsteht. Ähnlich werden wir die Formel für Taylor-Polynome abändern, um eine Approximation an einer beliebigen Stelle zu erhalten.

Definition 13.1. Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $f: I \rightarrow \mathbb{R}$ eine beliebig oft differenzierbare Funktion. Ist $0 \in I$, so definieren wir das n -te **Taylor-Polynom** $T_0^n f$ der Funktion f an der Stelle 0 durch die Vorschrift (formal und informell)

$$\begin{aligned} T_0^n f(x) &= \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k \\ &= \frac{f^{(0)}(0)}{0!} x^0 + \frac{f^{(1)}(0)}{1!} x^1 + \frac{f^{(2)}(0)}{2!} x^2 + \dots + \frac{f^{(n)}(0)}{n!} x^n \\ &= f(0) + f'(0)x + \frac{f''(0)}{2} x^2 + \dots + \frac{f^{(n)}(0)}{n!} x^n. \end{aligned}$$

Dabei bezieht sich die Zahl n auf den Grad des betrachteten Polynoms.

Das n -te **Taylor-Polynom** $T_{x_0}^n f$ der Funktion f an einer beliebigen Stelle $x_0 \in I$ ist gegeben durch die Vorschrift (formal und informell)

$$\begin{aligned} T_{x_0}^n f(x) &= \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \\ &= \frac{f^{(0)}(x_0)}{0!} (x - x_0)^0 + \frac{f^{(1)}(x_0)}{1!} (x - x_0)^1 + \frac{f^{(2)}(x_0)}{2!} (x - x_0)^2 \\ &\quad + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \\ &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2} (x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n. \end{aligned}$$

Wir bemerken insbesondere, dass, wenn wir bei $n = 1$ aufhören, im zweiten Fall genau die Geradengleichung der Tangente entsteht.

Nun wollen wir die Taylor-Polynome für den natürlichen Logarithmus ausrechnen.

Beispiel. Wir betrachten die Funktion $\ln: (0, \infty) \rightarrow \mathbb{R}$. Zu dieser bestimmen wir zunächst das Taylor-Polynom vom Grad 4 an der Stelle $x_0 = 1$ für den natürlichen Logarithmus. Dieser Wert ist besonders gut für unsere Zwecke geeignet, da wir sowohl den genauen Wert des Logarithmus an dieser Stelle angeben können (tatsächlich haben wir uns bereits überlegt, dass $\ln(1) = 0$ ist) und, wie wir gleich sehen, auch Einsetzen in die Ableitungsfunktionen besonders einfach sein wird.

Wir haben bereits die erste Ableitung der Logarithmusfunktion bestimmt:

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

Leiten wir diese Funktion nun wieder ab (und benutzen dabei $\frac{1}{x} = x^{-1}$), so erhalten wir als zweite Ableitungsfunktion

$$\ln''(x) = \left(\frac{1}{x}\right)' = -\frac{1}{x^2} \text{ für alle } x \in (0, \infty).$$

Im nächsten Schritt erhalten wir

$$\ln^{(3)}(x) = \left(-\frac{1}{x^2}\right)' = \frac{2}{x^3} \text{ für alle } x \in (0, \infty)$$

und danach

$$\ln^{(4)}(x) = \left(\frac{2}{x^3}\right)' = -\frac{2 \cdot 3}{x^4} \text{ für alle } x \in (0, \infty)$$

Setzt man nun $x_0 = 1$ ein, so erhält man

$$\ln'(1) = 1, \quad \ln''(1) = -1, \quad \ln^{(3)}(1) = 2, \quad \ln^{(4)}(1) = -2 \cdot 3 = -6.$$

Wir setzen diese Werte in die allgemeine Formel für Taylor-Polynome ein und markieren die eingesetzten Werte entsprechend farbig. Das vierte Taylor-Polynom an der Stelle $x_0 = 1$ für den natürlichen Logarithmus ist also gegeben durch

$$T_1^4 \ln(x) = 0 + \frac{1}{1!} \cdot (x-1)^1 + \frac{(-1)}{2} \cdot (x-1)^2 + \frac{2}{3!} \cdot (x-1)^3 + \frac{(-6)}{4!} \cdot (x-1)^4$$

Durch Kürzen erhalten wir

$$T_1^4 \ln(x) = (x-1) - \frac{1}{2}(x-1)^2 + \frac{1}{3}(x-1)^3 - \frac{1}{4}(x-1)^4$$

(Dieser Ausdruck kann jetzt nach Bedarf auch ausmultipliziert werden, um ein Polynom in der üblicheren Form zu erhalten, wir sind allerdings daran nicht weiter interessiert.)

Wir skizzieren nun, wie man allgemeiner das Taylor-Polynom vom vorgegebenen Grad n für den natürlichen Logarithmus an der Stelle $x_0 = 1$ bestimmt. Beim Bestimmen weiterer Ableitungen der Logarithmusfunktion stellt man nämlich fest, dass in jedem Schritt ein Faktor dazukommt, und dass diese Faktoren ganz ähnlich sind wie etwa beim Ableiten von Polynomfunktion, allerdings das Vorzeichen sich jedes Mal ändert und außerdem der Vorfaktor um einen „verschoben“ ist. Genauer kann man per Induktion zeigen, dass

$$\ln^{(k)}(x) = (-1)^{k+1} \frac{(k-1)!}{x^k}$$

für alle $x \in (0, \infty)$ und alle $k \geq 1$ gilt. Insbesondere gilt an der Stelle $x_0 = 1$:

$$\ln^{(k)}(x) = (-1)^{k+1} (k-1)!$$

Setzt man das nun in die Formel für die Taylor-Polynome ein, so kürzen sich einige Faktoren heraus, denn wir bemerken, dass nach Definition gilt: $k! = k \cdot (k-1)!$. Damit erhalten wir beim Einsetzen:

$$\begin{aligned} T_1^n \ln(x) &= \sum_{k=1}^n \frac{\ln^{(k)}(1)}{k!} (x-1)^k \\ &= \sum_{k=1}^n \frac{(-1)^{k+1} (k-1)!}{k!} (x-1)^k \\ &= \sum_{k=1}^n \frac{(-1)^{k+1}}{k} (x-1)^k. \end{aligned}$$

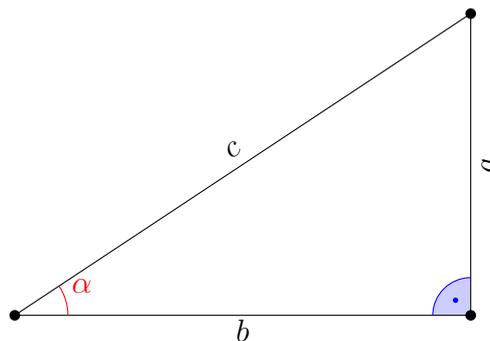
Das gibt uns also die Möglichkeit, den Logarithmus in der Nähe der 1 durch Polynome anzunähern.

Es sei nochmal hervorgehoben, dass der Parameter x_0 festgemacht wird und wir für Taylor-Polynome die präzisen Werte der Funktion f an dieser Stelle x_0 brauchen, während x die Stellen beschreibt, an denen wir die Funktion f approximieren. Meist darf x nicht zu weit von x_0 sein, damit wir einen sinnvollen Näherungswert erhalten.

14 Trigonometrische Funktionen

Das Ziel dieses Kapitels ist es, trigonometrische Funktionen, die zum Teil bereits aus der Schule bekannt sind, aufzugreifen und weiter zu untersuchen. Das Wort „trigonometrisch“ bedeutet, aus dem Griechischen übersetzt, in etwa „Dreiecksvermessung“. Die Funktionen, die wir hier betrachten, werden genutzt, um Längenverhältnisse im Dreieck zu bestimmen. Insgesamt definiert man üblicherweise 6 trigonometrische Funktionen; wir werden uns hier auf Sinus- und Kosinusfunktion konzentrieren, und nur am Rande noch die Tangensfunktion betrachten.

Aus der Schule ist meistens bereits die Definition vom Sinus und Kosinus für spitze Winkel bekannt; wir hatten diese bereits bei orthogonalen Abbildungen gebraucht: Der **Kosinus** von einem spitzen Winkel α in einem rechtwinkligen Dreieck als das Verhältnis der Längen der Ankathete und der Hypotenuse definiert ist. Wir veranschaulichen das nochmal im Bild:



In einem solchen rechtwinkligen Dreieck ist Kosinus von α gegeben durch

$$\cos(\alpha) = \frac{b}{c},$$

wobei a, b die Längen der im Bild markierten Katheten und c die Hypotenusenlänge sind.

Ähnlich ist **Sinus** des spitzen Winkels α definiert als das Verhältnis der Längen der Gegenkathete zu α und der Hypotenuse, mit den Bezeichnungen aus dem Bild also

$$\sin(\alpha) = \frac{a}{c}.$$

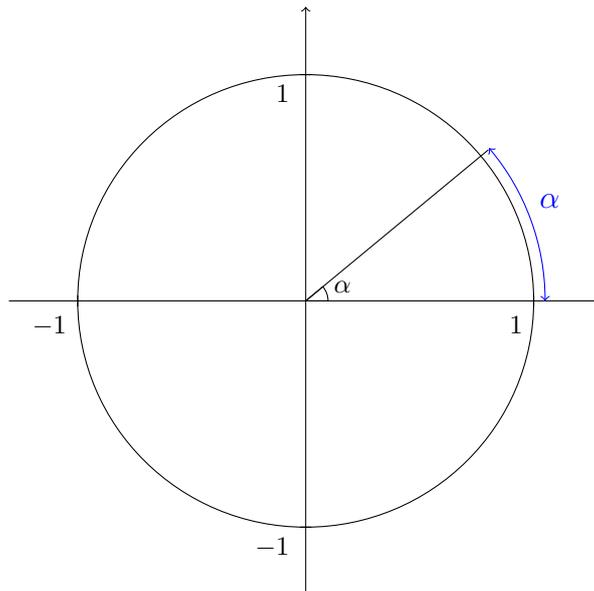
Wir bemerken, dass man dann aus der Länge der Hypotenuse und der Größe des Winkels α die Längen der beiden Katheten im rechtwinkligen Dreieck bestimmen kann, indem man die obigen Formeln nach a bzw. b auflöst. Dann erhalten wir für die Länge der Ankathete zu α :

$$b = c \cdot \cos(\alpha)$$

und für die Länge der Gegenkathete von α :

$$a = c \cdot \sin(\alpha).$$

Als nächstes ist unser Ziel, die trigonometrische Funktionen für beliebige (und nicht nur spitze) Winkel zu definieren. Dafür ist es allerdings zunächst sinnvoll, die Art, Winkel zu messen, zu verändern. In der Mathematik ist es häufig praktischer, Winkel in Bogenmaß zu messen; dieses Konzept ist manchmal bereits aus der Schule bekannt. Eine volle Umdrehung im üblichen Gradmaß entspricht einem Winkel von 360° . Die Festlegung auf 360° ist sehr alt - sie geht vermutlich schon auf alte Babylonier zurück - und recht willkürlich. Man weiß nicht genau, warum gerade 360° gewählt wurden, man vermutet aber eine Verbindung zur Anzahl der Tage im Jahr. Man könnte allerdings auch den Kreis beispielsweise in 400 gleiche Teile einteilen - dieses System wird vereinzelt tatsächlich verwenden, und entspricht der „GRAD“-Taste, die einige Taschenrechner besitzen. Beim Bogenmaß ist die Festlegung allerdings weniger willkürlich, und verbindet Längen und Winkel. Zu jedem Winkel können wir ein zugehöriges Stück vom Einheitskreis betrachten, also vom Kreis mit Mittelpunkt im Koordinatenursprung und vom Radius 1. Die Länge des entsprechenden Kreisbogens wird nun als Maß für den Winkel benutzt.



Dabei erinnern wir uns, dass die Länge einer Kreislinie mit Radius r durch $2\pi \cdot r$ gegeben ist. Die Länge des Einheitskreises ist also 2π , sodass der Winkel von 360° in Bogenmaß 2π misst. So lässt sich mit Dreisatz die Größe eines Winkels in Bogenmaß zu Gradmaß und umgekehrt vom Gradmaß ins Bogenmaß umrechnen, denn es gilt:

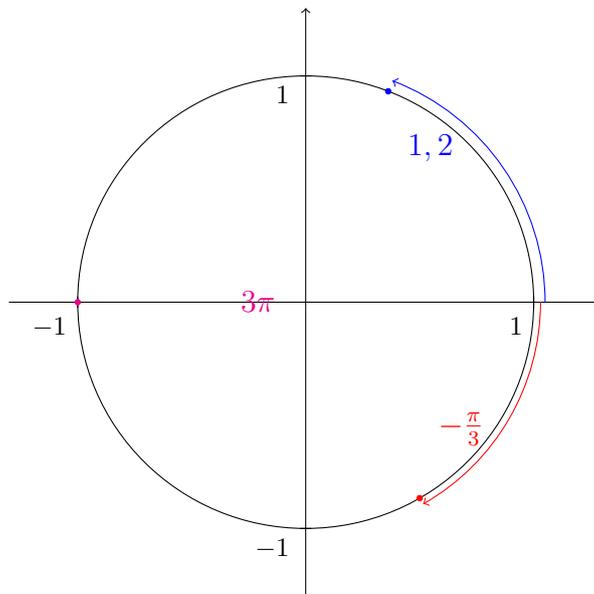
$$\frac{\alpha \text{ in Bogenmaß}}{2\pi} = \frac{\alpha \text{ in Gradmaß}}{360^\circ}.$$

Dazu geben wir einige Beispiele:

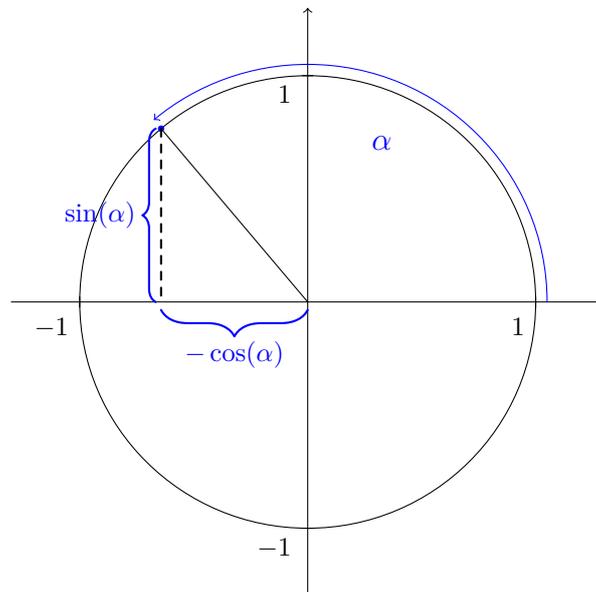
Beispiel. Hier sind einige Beispiele für die Umrechnung:

360°	entspricht in Bogenmaß	2π
180°	entspricht in Bogenmaß	π
90°	entspricht in Bogenmaß	$\frac{\pi}{2}$
45°	entspricht in Bogenmaß	$\frac{\pi}{4}$

Nun gehört zu jeder reellen Zahl x ein eindeutiger Punkt auf dem Kreis, den man erhält, indem man vom Punkt $(1, 0)$ gegen den Uhrzeigersinn die Strecke x entlang des Kreises zurücklegt und den Endpunkt betrachtet. Bei Winkeln zwischen 0 und 2π erhalten wir gerade den gewohnten Punkt auf dem Einheitskreis. Für negative reelle Zahlen x misst man die Länge $|x|$ ab, diesmal im Uhrzeigersinn vom Punkt $(1, 0)$ aus. Hier sind einige Beispiele.



Nun definieren wir den **Kosinus** der reellen Zahl (oder eines Winkels) α als die x -Koordinate des zugehörigen Punktes auf dem Einheitskreis und den **Sinus** als die y -Koordinate von diesem Punkt. Bildlich können wir das wie folgt darstellen:



Man beachte, dass nun Sinus bzw. Kosinus sowohl negative als auch positive Werte (und den Wert 0) annehmen. In dem Bild ist der Kosinus des Winkels α negativ, daher ist die Länge der Kathete des Koordinatendreiecks, die auf der x -Achse liegt, gerade $-\cos(\alpha)$, da die Länge stets eine positive (oder allenfalls nicht-negative) Zahl ist. In dem konkreten Beispiel ist Sinus von α positiv, also ist die Länge der anderen Kathete gerade $\sin(\alpha)$.

Beispiel. In dem obigen Bild haben wir gesehen, dass der zu 3π gehörende Punkt gerade der Punkt $(-1, 0)$ ist, da 3π eineinhalb Umdrehungen entspricht. Folglich gilt:

$$\begin{aligned}\sin(3\pi) &= 0 \\ \cos(3\pi) &= -1.\end{aligned}$$

Als nächstes betrachten wir nochmal das Koordinatendreieck von dem Punkt auf dem Einheitskreis, der zu dem Punkt α gehört. Die Kathetenlängen in diesem Dreieck sind $|\sin(\alpha)|$ und $|\cos(\alpha)|$, während die Hypotenuse ein Radius des Einheitskreises bildet und somit die Länge 1 hat. Wir erhalten also die folgende zentrale Identität aus dem Satz des Pythagoras:

$$(\sin(\alpha))^2 + (\cos(\alpha))^2 = 1.$$

Dabei haben wir benutzt, dass für alle reellen Zahlen $t \in \mathbb{R}$ gilt: $(-t)^2 = t^2 = (|t|)^2$.

Ferner bemerken wir, dass die x - bzw. y -Koordinate von einem Punkt auf dem Einheitskreis stets im Bereich zwischen -1 und 1 liegt, also gilt

$$\begin{aligned}-1 &\leq \sin(\alpha) \leq 1, \\ -1 &\leq \cos(\alpha) \leq 1\end{aligned}$$

für alle $\alpha \in \mathbb{R}$.

Nun haben wir die Funktionen $\sin: \mathbb{R} \rightarrow \mathbb{R}$ und $\cos: \mathbb{R} \rightarrow \mathbb{R}$ definiert. (Wir könnten auch das Ziel auf $[-1, 1]$ einschränken, allerdings wird es für unsere Zwecke nicht nötig sein.) Unser nächstes großes Ziel besteht darin, die Ableitungen dieser Funktionen zu bestimmen. Im Falle der Sinusfunktion wollen wir dazu einige Details geben. Dafür ist jedoch etwas Vorbereitung notwendig. Als Motivation erinnern wir uns, dass nach Definition die Ableitung der Sinusfunktion im Punkt x_0 gegeben ist durch

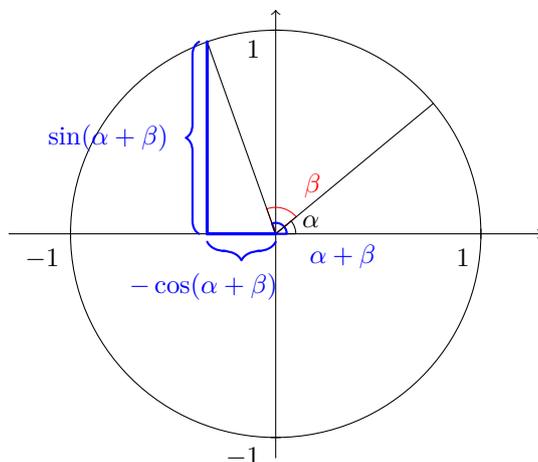
$$\sin'(x_0) = \lim_{h \rightarrow 0} \frac{\sin(x_0 + h) - \sin(x_0)}{h}.$$

Für die Beispiele der Ableitungen, die wir bis jetzt gesehen haben, war es entscheiden, den ersten Summanden im Nenner sinnvoll umformen zu können. Auch bei der Exponentialfunktion war es sehr wichtig zu wissen, dass diese Summen auf Produkte abbildet, also $e^{x+y} = e^x \cdot e^y$ gilt. Unser Ziel ist es also, Sinusfunktion der Summe von zwei Winkel anders auszudrücken, am liebsten in Termen der Sinuswerte der Summanden. Das klappt allerdings nicht ganz, doch der Unterschied ist nicht „zu schlimm“, denn man kann $\sin(x + y)$ in Termen von $\sin(x)$, $\sin(y)$ und zusätzlich $\cos(x)$, $\cos(y)$ ausdrücken. Das ist der Gegenstand der sogenannten Additionstheoreme, die wir nun beweisen wollen.

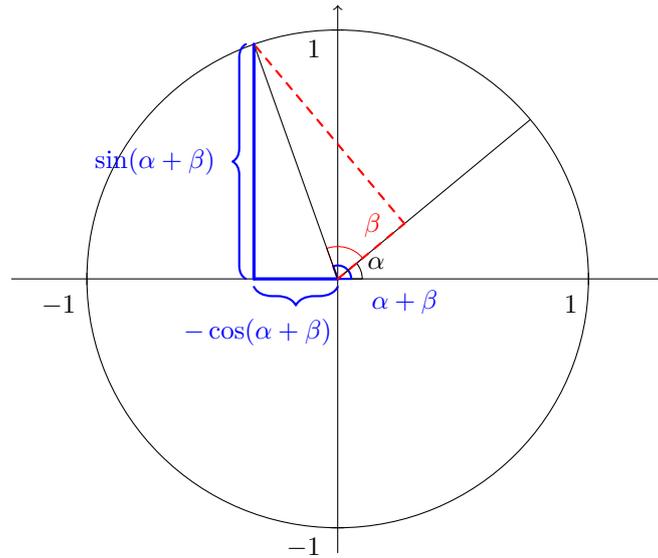
Satz 14.1 (Additionstheoreme). *Für alle reellen Zahlen $\alpha, \beta \in \mathbb{R}$ gilt:*

$$\begin{aligned}\cos(\alpha + \beta) &= \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta), \\ \sin(\alpha + \beta) &= \sin(\alpha) \cos(\beta) + \cos(\alpha) \sin(\beta).\end{aligned}$$

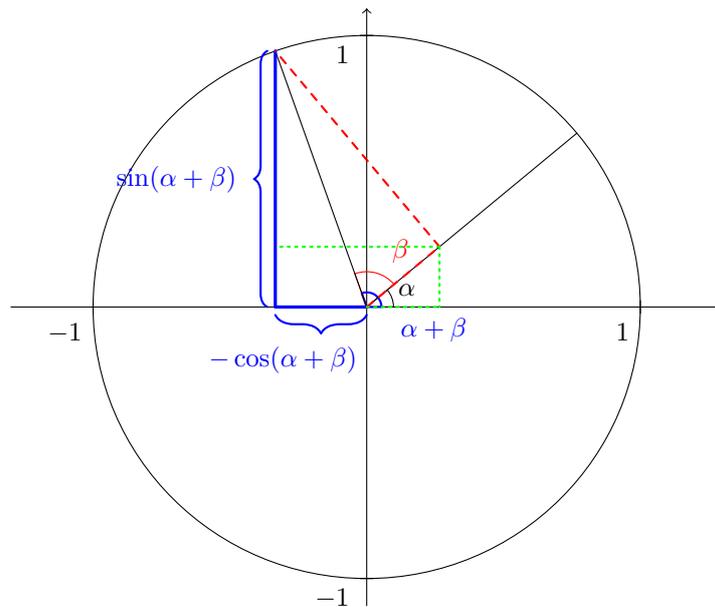
Beweis. Der Beweis dieses Satzes basiert auf einigen geometrischen Überlegungen. Wir zeichnen zuerst das Bild für die Ausgangssituation. Wir haben also zwei reelle Zahlen α, β vorgegeben und zeichnen ihre Summe in den Einheitskreis ein, indem wir zuerst den Winkel α einzeichnen und dann einen weiteren Radius in Kreis, der zum vorherigen im Winkel β steht. Wir markieren gleich die Längen, durch die $\sin(\alpha + \beta)$ und $\cos(\alpha + \beta)$ definiert sind:



Nun zeichnen wir einige Hilfslinien in dieses Bild ein. Von dem Punkt des Einheitskreises aus, der zum Winkel $\alpha + \beta$ gehört, fallen wir das Lot auf die Gerade, die mit der x -Achse den Winkel α einschließt. Dieses Lot sowie die Verbindung des Lotfußpunktes mit dem Ursprung zeichnen wir als rote, gestrichelte Linien ein:

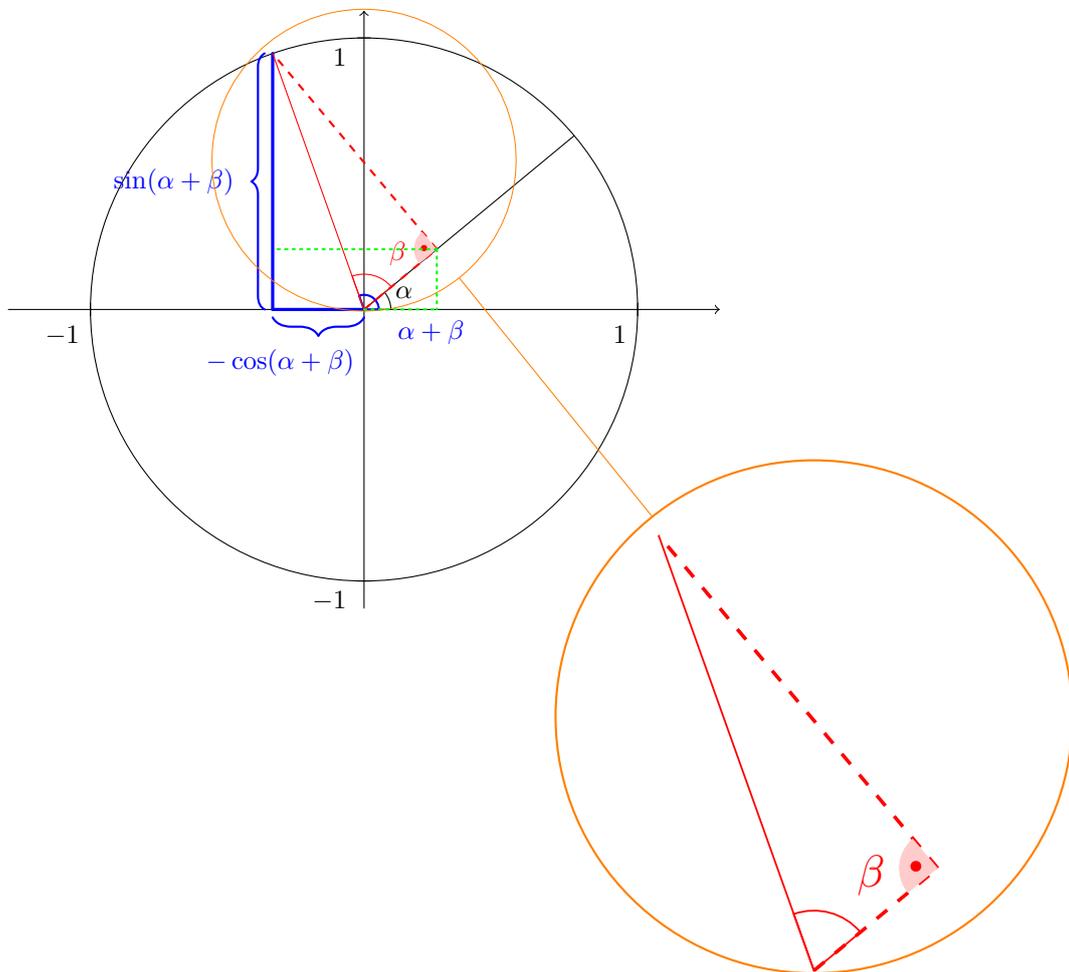


Außerdem zeichnen wir nun von unserem Lotfußpunkt aus eine waagerechte und eine senkrechte Strecke ein (grün-gepunktet), sodass ein Rechteck mit den blauen Abschnitten der Seiten des definierenden Dreiecks für $\sin(\alpha + \beta)$ und $\cos(\alpha + \beta)$ und einem Abschnitt der x -Achse, den wir ebenfalls grün markieren, entsteht.



Nun wissen wir, dass gegenüberliegende Seiten in einem Rechteck immer gleich lang sind. So sind in dem soeben markierten Rechteck beispielsweise beide waagerechte Seiten gleich lang. In der unteren waagerechten Seite kommt $-\cos(\alpha + \beta)$ als Länge eines Teilabschnittes vor. Wir werden nun die Länge des anderen, grün markierten Teilabschnittes sowie die Länge der oberen waagerechten Seite des Rechtecks bestimmen; dadurch werden wir $\cos(\alpha + \beta)$ bestimmen können und das erste Additionstheorem daraus herleiten können.

Bevor wir das erreichen können, müssen wir die Längen der Seiten im rot markiertem Dreieck bestimmen:

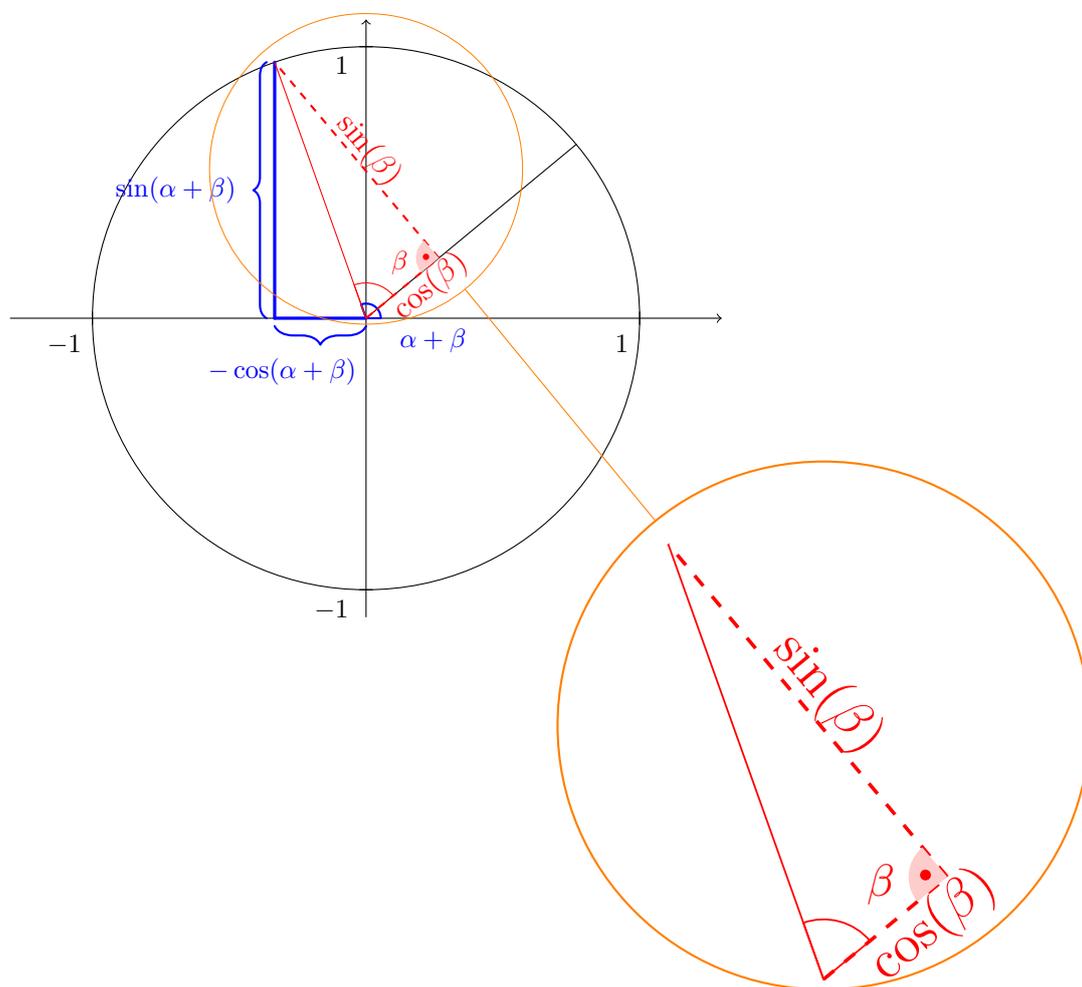


Dieses Dreieck ist rechtwinklig, da es gerade dadurch entstanden ist, dass wir einen Lot von einem Punkt des Einheitskreises auf den Strahl gefällt haben, der im Winkel α auf die x -Achse steht. Ferner ist die Hypotenuse in diesem Dreieck gleichzeitig ein Radius des Einheitskreises, hat also die Länge 1. Schließlich kennen wir auch nach Konstruktion die Größe eines der beiden spitzen Winkel, und diese ist nach Definition β , wie im Bild markiert.

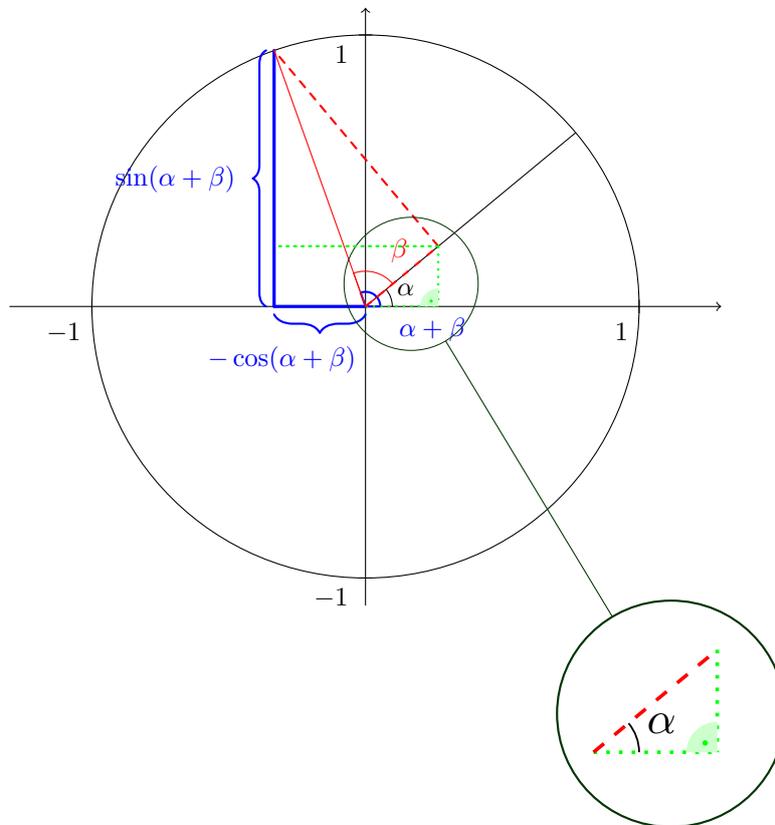
Nun haben wir uns vorher überlegt, dass wir aus der Länge der Hypotenuse und der Größe eines Winkels im rechtwinkligen Dreieck mit Hilfe der trigonometrischen Funktionen auch die Kathetenlängen in diesem Dreieck bestimmen können. Die Ankathete von β hat nämlich die Länge

$$1 \cdot \cos(\beta) = \cos(\beta),$$

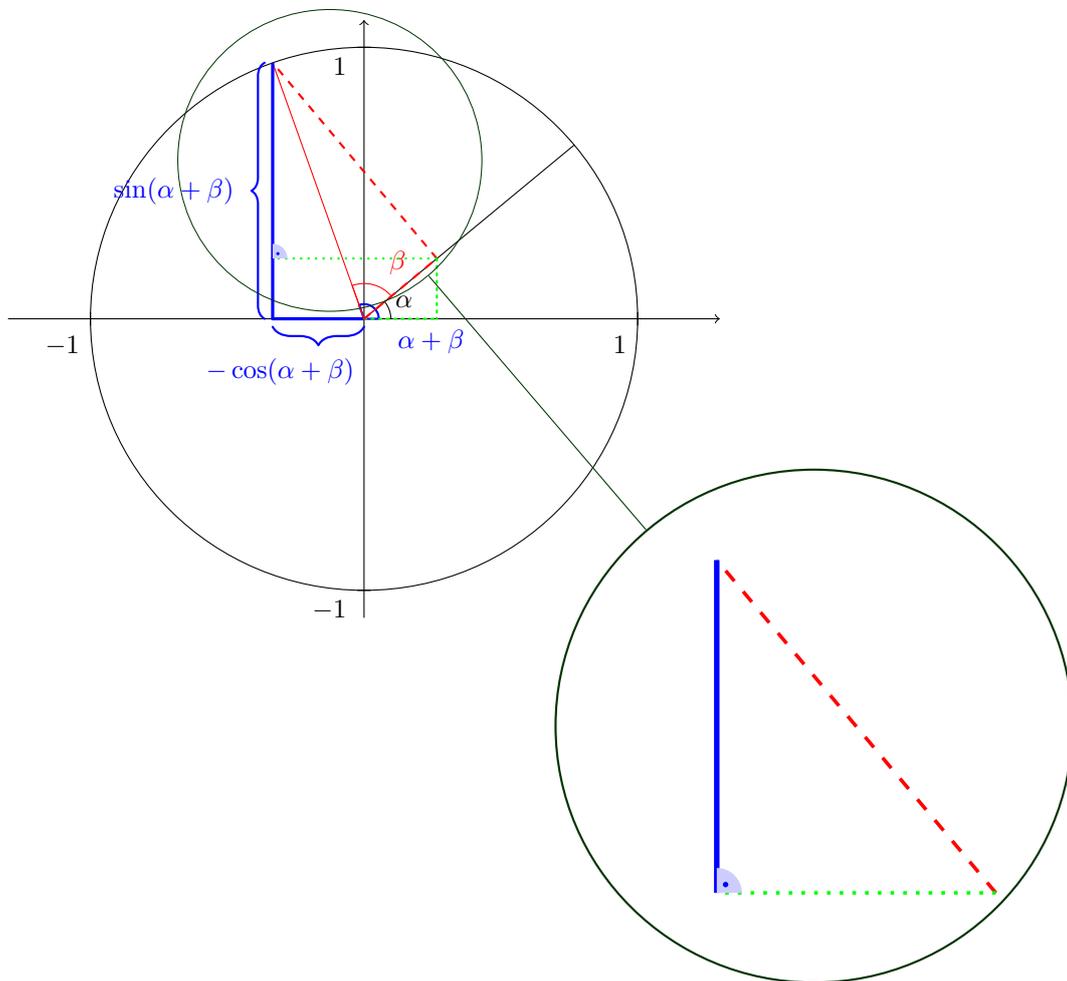
während die Gegenkathete von β die Länge $1 \cdot \sin(\beta) = \sin(\beta)$ hat. Wir markieren das nochmal im Bild:



Nun erinnern wir uns an unseren Plan, die Längen der grünen waagerechten Strecken zu bestimmen, und fangen mit der kürzeren, unteren waagerechten Strecke an.



Wir betrachten also das Dreieck, das von der unteren Kathete des vorigen, roten Dreiecks, ferner einer senkrechten Strecke, die aus dem einen Ende der roten Seite bis zur x -Achse eingezeichnet wurde, sowie aus dem entsprechenden Abschnitt der x -Achse gebildet wird. In diesem Dreieck hat man wieder einen rechten Winkel, und zwar zwischen der waagerechten und der senkrechten grünen Strecke. Der Winkel zwischen der Hypotenuse und der x -Achse ist gerade nach Konstruktion der Winkel α . Ferner haben wir soeben festgestellt, dass die Hypotenusenlänge in diesem Dreieck $\cos(\beta)$ ist, denn das ist gerade die Ankathete von β in dem roten Dreieck, das wir vorher betrachtet haben. Nun können wir nach derselben Methode die Kathetenlängen im rot-grünen Dreieck bestimmen: Die waagerechte grüne Seite, die die Ankathete des Winkels α ist, hat die Länge $\cos(\beta) \cdot \cos(\alpha)$, während die senkrechte grüne Seite, die die Gegenkathete des Winkels α ist, hat die Länge $\cos(\beta) \cdot \sin(\alpha)$. Wir markieren das wieder im Bild:

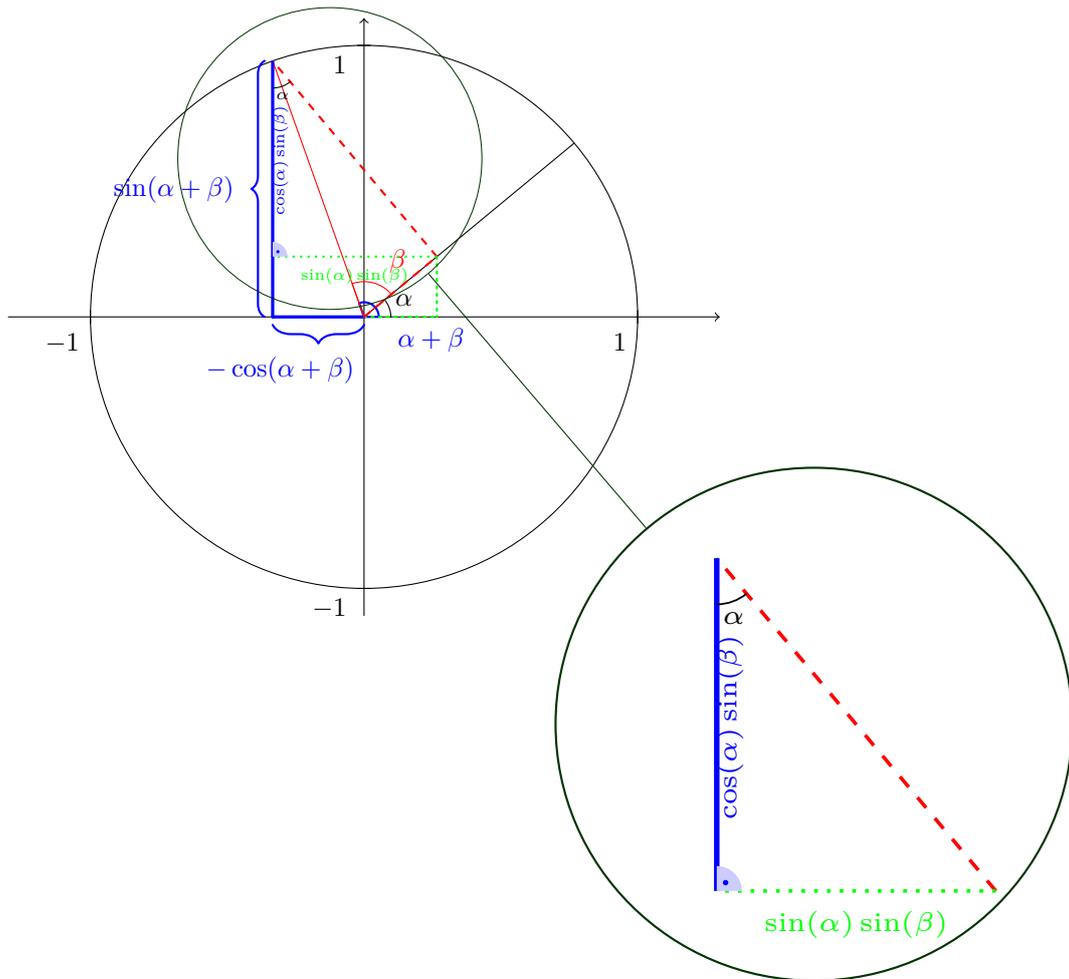


In diesem Dreieck würden wir gerne dieselbe Methode anwenden, wie auch in beiden vorherigen Dreiecken. Auch hier haben wir wieder einen rechten Winkel zwischen der blauen senkrechten und grünen waagerechten Seite. Außerdem wissen wir, dass die Länge der Hypotenuse $\sin(\beta)$ ist, da das gerade die Länge der Gegenkathete von β im roten Hilfsdreieck ist. Allerdings kennen wir ohne weiteres nicht die Größe der spitzen Winkel. Also führen wir als nächstes eine kleine Überlegung durch, um den Winkel zwischen der roten Hypotenuse und der blauen Kathete in diesem Dreieck zu bestimmen.

Dafür entscheidend ist die Tatsache, dass zwei Winkel, deren Seiten paarweise senkrecht zueinander sind, dieselbe Größe haben müssen. In diesem Fall ist die blaue, senkrechte Seite des gesuchten Winkels senkrecht zu der x -Achse, die wie üblich waagrecht ist. Ferner ist die rote Seite des gesuchten Winkels senkrecht zu der Geraden, die den Winkel α mit der x -Achse einschließt, denn die rote Strecke wurde ja als Lot auf diese Gerade konstruiert. Also stehen die blaue Kathete und die rote Hypotenuse jeweils senkrecht auf die x -Achse bzw. auf die Gerade, die damit den Winkel α einschließt. Daher bilden die blaue Kathete und die rote Hypotenuse ebenfalls den Winkel α .

Nun sind wir wieder in der Situation, in dem rechtwinkligen grün-blau-

rotem Dreieck sowohl die Hypotenusenlänge als auch einen Winkel zu kennen. Die grüne waagerechte Seite ist die Gegenkathete des Winkels α , hat also die Länge $\sin(\beta) \sin(\alpha)$, während die blaue Kathete die Länge $\sin(\beta) \cos(\alpha)$ hat. Wir markieren das wieder im Bild:

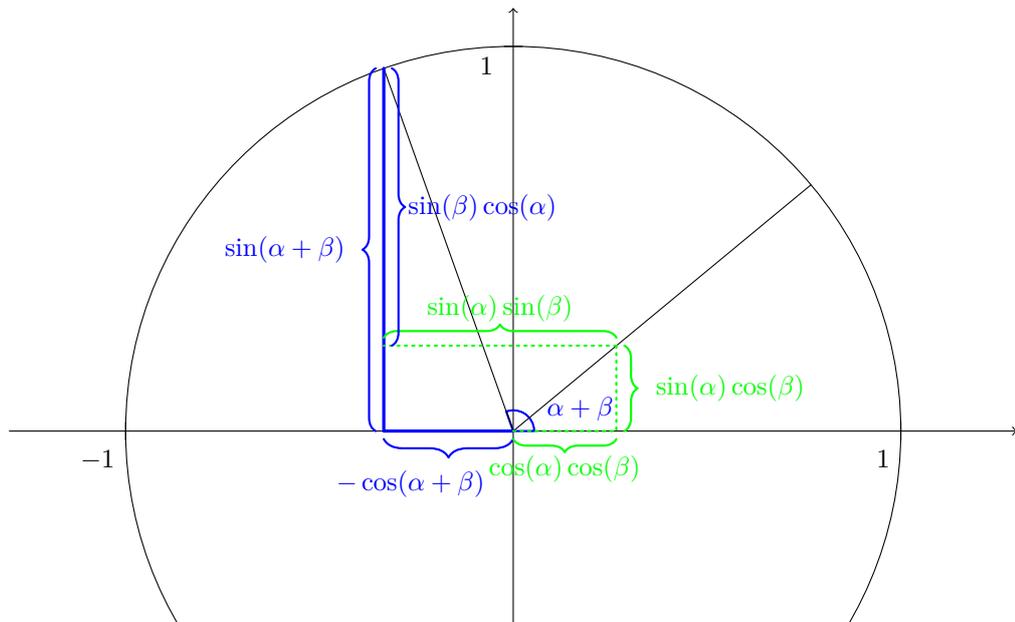


Nun haben wir festgestellt, dass die obere grüne waagerechte Rechtecksseite die Länge $\sin(\alpha) \sin(\beta)$ hat, während die untere sich aus zwei Teilstrecken zusammensetzt, von denen eine die Länge $-\cos(\alpha + \beta)$ und eine die Länge $\cos(\alpha) \cos(\beta)$ besitzt. Da die beiden Seiten des Rechtecks gleich lang sind, folgern wir also

$$\begin{aligned} \sin(\alpha) \sin(\beta) &= -\cos(\alpha + \beta) + \cos(\alpha) \cos(\beta) \\ \Leftrightarrow \cos(\alpha + \beta) &= \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta). \end{aligned}$$

(Dabei haben wir im letzten Schritt nur die Terme umsortiert.) Das ist allerdings gerade die Aussage des Additionstheorems für den Kosinus, die wir beweisen wollten. Das veranschaulichen wir nochmal im Bild (die roten Hilfs-

linien werden dabei nicht mehr gebraucht). Wir tragen auch gleich die relevanten Größen für das zweite Additionstheorem ein.



Nun kommen wir zum Beweis des Additionstheorems für Sinus. Dafür teilen wir, wie bereits im Bild angedeutet, die blaue Strecke, deren Länge nach Definition $\sin(\alpha + \beta)$ ist, in zwei Teile auf. Der untere Teil ist die senkrechte Seite des bereits betrachteten Rechtecks. Da ihre Länge gleich der Länge der anderen senkrechten Seite in diesem Rechteck, und diese haben wir bereits im grün-rotem Dreieck zu $\cos(\beta)\sin(\alpha)$ bestimmt. Der zweite Teil ist die senkrechte Kathete im rot-grün-blauen Dreieck. Auch diese Länge haben wir bereits bestimmt, und haben $\sin(\beta)\cos(\alpha)$ erhalten. Da die Strecke der Länge $\sin(\alpha + \beta)$ sich aus diesen beiden Teilen zusammensetzt, erhalten wird

$$\sin(\alpha + \beta) = \cos(\beta)\sin(\alpha) + \sin(\beta)\cos(\alpha).$$

Das ist gerade die Aussage des Additionstheorems für Sinus. Damit ist der Beweis der Additionstheoreme im Wesentlichen abgeschlossen. Es sollte noch angemerkt werden, dass wir von der konkreten Lage der Winkel α und β in unserem Bild, insbesondere davon, dass $\alpha + \beta$ im zweiten Quadranten liegt, Gebrauch gemacht haben. Allerdings lässt sich der Beweis ohne größere Schwierigkeiten vervollständigen. \square

Die Additionstheoreme sind wichtige Aussagen beim Arbeiten mit trigonometrischen Funktionen. Insbesondere wollen wir sie, wie ursprünglich angekündigt, zur Bestimmung der Ableitungsfunktion von Sinus heranziehen. Mit Hilfe der Additionstheoreme können wir nun einen Schritt dabei weitergehen und erhalten aus der Definition, indem wir die Terme im Differenzenquotienten umsortieren und bei der zweiten Gleichheit das Additionstheorem für Sinus verwenden:

$$\begin{aligned} \sin'(x_0) &= \lim_{h \rightarrow 0} \frac{\sin(x_0 + h) - \sin(x_0)}{h} \\ &= \lim_{h \rightarrow 0} \frac{\sin(x_0) \cos(h) + \cos(x_0) \sin(h) - \sin(x_0)}{h} \\ &= \lim_{h \rightarrow 0} \left(\frac{\sin(x_0) \cos(h) - \sin(x_0)}{h} + \frac{\cos(x_0) \sin(h)}{h} \right) \\ &= \lim_{h \rightarrow 0} \left(\sin(x_0) \cdot \frac{\cos(h) - 1}{h} + \cos(x_0) \cdot \frac{\sin(h)}{h} \right). \end{aligned}$$

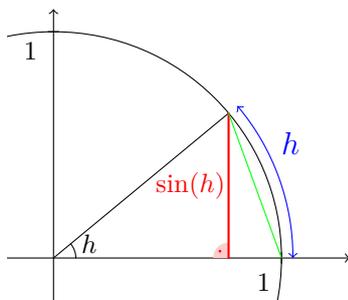
Wir erinnern uns, dass $x_0 \in \mathbb{R}$ fest gewählt ist. Wenn wir also die beiden Grenzwerte

$$\lim_{h \rightarrow 0} \frac{\cos(h) - 1}{h} \quad \text{und} \quad \lim_{h \rightarrow 0} \frac{\sin(h)}{h}$$

bestimmen können, wären wir fertig. Das bedarf einiger weiterer Überlegungen.

Wir fangen mit dem zweiten Grenzwert an. Dafür erinnern wir uns an den Satz 10.5. In dem zweiten Teil dieses Satzes haben wir notiert, dass eine Folge, die größtmäßig zwischen zwei gegen dieselbe Zahl konvergierenden Folgen liegt, selbst gegen dieselbe Zahl konvergieren muss. Dieser Satz überträgt sich auch auf Grenzwerte von Funktionen. Wir werden also zwei Ungleichungen für die Funktion $\frac{\sin(h)}{h}$ von h finden, sodass sowohl die obere als auch die untere Schranke für $h \rightarrow 0$ denselben Grenzwert haben.

Das machen wir nun etwas präziser, zunächst mit der Beobachtung, dass für kleine positive Werte von h stets $\sin(h) < h$ gilt. Dafür betrachten wir den entsprechenden Ausschnitt des Einheitskreises.



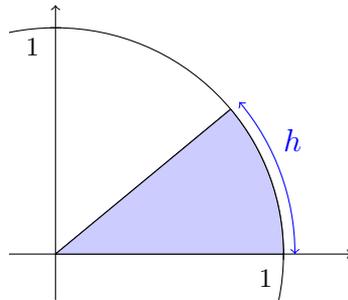
Nach Definition der Bogenlänge entspricht der Winkel h einem Bogen der Länge h auf dem Einheitskreis. Dieser Bogen ist sicherlich länger als die

direkte Verbindung seiner beiden Endpunkten durch eine Strecke, die im Bild grün markiert ist. Diese aber ist wiederum länger als das Lot von dem Punkt des Einheitskreises, der zum Winkel h gehört, auf die x -Achse. Allerdings ist die Länge dieses Lots nach Definition gerade $\sin(h)$. Für alle $0 < h < \frac{\pi}{2}$ gilt also $\sin(h) \leq h$ und folglich

$$\frac{\sin(h)}{h} < 1.$$

Ähnlich kann man diese Identität auch für negative Werte von h beweisen. Nun werden wir zeigen, dass $\frac{\sin(h)}{h}$ auch von unten durch eine Funktion beschränkt ist, die für $h \rightarrow 0$ den Grenzwert 1 hat, und auf diese Weise den Grenzwert von $\frac{\sin(h)}{h}$ für $h \rightarrow 0$ bestimmen.

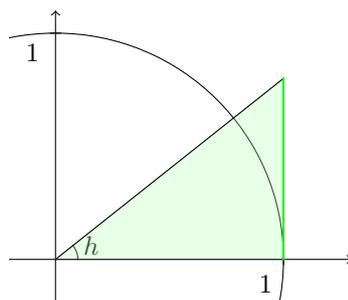
Dafür wird es nützlich sein, zunächst die Fläche vom „Tortenstück“ zu bestimmen, der durch den Winkel h gegeben ist:



Wir wissen, dass die Fläche des Kreises mit Radius r durch die Formel $A = \pi \cdot r^2$ berechnet werden kann. Also beträgt die Fläche des Einheitskreises π . Jetzt müssen wir uns also überlegen, welchen Anteil des Kreises das „Tortenstück“ ausmacht. (So hätte ja ein Halbkreis die Fläche $\frac{\pi}{2}$ und ein Viertelkreis $\frac{\pi}{4}$.) Doch das können wir gerade als $\frac{h}{2\pi}$ bestimmen, denn wir haben ein Winkel h mit dem Vollwinkel 2π zu vergleichen. Also ist die Fläche des blau eingefärbten „Tortenstücks“ gegeben durch

$$\frac{h}{2\pi} \cdot \pi = \frac{h}{2}.$$

Diese Fläche ist sicherlich kleiner als die Fläche des rechtwinkligen Dreiecks, das von der x -Achse, einer senkrechten Gerade durch den Punkt $(1, 0)$ sowie der Geraden, die im Winkel h zur x -Achse steht, gebildet wird:



Dieses Dreieck ist nach Konstruktion rechtwinklig. Die Fläche eines rechtwinkligen Dreiecks bestimmt man als die Hälfte des Produktes der Kathetenlängen. Die Länge der waagerechten Kathete ist nach Definition 1. Nun geht es darum, die Länge der senkrechten Kathete zu bestimmen. Dabei brauchen wir als Hilfsgröße die Länge der Hypotenuse in diesem Dreieck, die wir mit c bezeichnen. Nun sind wir wieder in der Situation, ein rechtwinkliges Dreieck zu haben, in dem manche Seitenlängen und die Größe eines spitzen Winkels bekannt ist. Da die waagerechte Kathete die Ankathete des Winkels h ist und die Länge 1 hat, schließen wir aus der Definition des Kosinus für spitze Winkel:

$$\frac{1}{c} = \cos(h).$$

Also kann die Länge der Hypotenuse durch

$$c = \frac{1}{\cos(h)}$$

bestimmt werden. Will man jetzt daraus die Länge der Gegenkathete von h bestimmen, so erhält man:

$$c \cdot \sin(h) = \frac{\sin(h)}{\cos(h)}.$$

Es sei angemerkt, dass dieser Quotient auch *Tangens* von h bezeichnet wird. Wir werden diese Bezeichnung nicht weiter brauchen.

Nun können wir die Fläche des markierten Dreiecks bestimmen:

$$\frac{1}{2} \cdot 1 \cdot \frac{\sin(h)}{\cos(h)} = \frac{\sin(h)}{2 \cos(h)}.$$

Da die Fläche des Dreiecks für jedes positive h größer ist als die Fläche des „Tortenstücks“, der darin ganz enthalten ist, erhalten wir die Ungleichung

$$\frac{h}{2} < \frac{\sin(h)}{2 \cos(h)}.$$

Für Werte $-\frac{\pi}{2} < h < \frac{\pi}{2}$ ist $\cos(h)$ eine positive Zahl, also kann man beide Seiten der Ungleichung mit $2 \cos(h)$ multiplizieren und wieder eine gültige Ungleichung erhalten:

$$\begin{aligned} \frac{h}{2} < \frac{\sin(h)}{2 \cos(h)} & \quad | \cdot 2 \cos(h) \\ \Leftrightarrow \cos(h) \cdot h < \sin(h). \end{aligned}$$

Für positive h erhalten wir unmittelbar, indem wir beide Seiten durch h teilen:

$$\cos(h) < \frac{\sin(h)}{h}.$$

Für $h < 0$ kann diese Ungleichung ebenfalls mit einem geringen Zusatzaufwand gezeigt werden.

Nun haben wir im Wesentlichen für alle kleinen Werte von $h \neq 0$ gezeigt, dass

$$\cos(h) < \frac{\sin(h)}{h} < 1$$

gilt. Der Wert $\cos(0)$ ist definiert und es gilt: $\cos(0) = 1$, da der Punkt auf dem Einheitskreis, der zum Winkel 0 gehört, gerade der Punkt $(1, 0)$ ist, sodass wir

$$\cos(0) = 1 \text{ und } \sin(0) = 0$$

schließen können. Ferner ist es so, dass die Werte von $\cos(h)$ immer näher an 1 sind, je näher h an 0 sind. Das kann man zwar präziser zeigen, aber wir wollen uns das vor allem bildlich veranschaulichen: Je kleiner der Winkel h ist, desto größer ist die Kathete des zugehörigen Dreiecks und desto näher ist ihre Länge an der des Radius von dem Kreis. Damit sollte zumindest plausibel sein, dass

$$\lim_{h \rightarrow 0} \cos(h) = 1$$

ist.

Nun können wir die Variante des Satzes 10.5 für Funktionen anwenden. Da sowohl $\cos(h)$ als auch 1 für $h \rightarrow 0$ gegen 1 gehen, muss also auch gelten:

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

Nun haben wir einen der beiden Grenzwerte bestimmt, die in der Ableitung von der Sinusfunktion auftauchen. Jetzt geht es um den anderen der beiden Grenzwerte: Wir haben

$$\lim_{h \rightarrow 0} \frac{\cos(h) - 1}{h}$$

zu bestimmen. Dafür wenden wir einen der Tricks an, die wir bereits häufiger gesehen haben, und erweitern Zähler und Nenner mit $1 + \cos(h)$. (Es sei angemerkt, dass dieser Term für kleine Werte von h nicht Null sein kann.) Wir erhalten:

$$\lim_{h \rightarrow 0} \frac{\cos(h) - 1}{h} = \lim_{h \rightarrow 0} \frac{(\cos(h) - 1)(\cos(h) + 1)}{h(\cos(h) + 1)}.$$

Nun wenden wir im Zähler die dritte binomische Formel an:

$$\lim_{h \rightarrow 0} \frac{(\cos(h) - 1)(\cos(h) + 1)}{h(\cos(h) + 1)} = \lim_{h \rightarrow 0} \frac{(\cos(h))^2 - 1}{h(\cos(h) + 1)}.$$

Wir erinnern uns an die Formel

$$(\sin(h))^2 + (\cos(h))^2 = 1,$$

die wir vorher bewiesen haben. Umgestellt liefert diese: $(\cos(h))^2 - 1 = -(\sin(h))^2$. Wir setzen das wieder ein und formen etwas um:

$$\lim_{h \rightarrow 0} \frac{(\cos(h))^2 - 1}{h(\cos(h) + 1)} = \lim_{h \rightarrow 0} \frac{-(\sin(h))^2}{h(\cos(h) + 1)} = \lim_{h \rightarrow 0} \left(\frac{\sin(h)}{h} \cdot \frac{-\sin(h)}{(\cos(h) + 1)} \right).$$

Ähnlich wie bei Kosinus kann man auch

$$\lim_{h \rightarrow 0} \sin(h) = 0$$

schließen.

Nach dem Satz 10.7 können wir also die einzelnen Grenzwerte durch Grundrechenarten verknüpfen, um den Grenzwert unseres Ausdrucks zu bestimmen. Wir konnten das nicht von Anfang an machen, da der Grenzwert des Nenners in Ursprünglicher Form 0 ist, und das gerade die Ausnahme im Satz 10.7 war. Da aber der Grenzwert der Funktion $\frac{\sin(h)}{h}$ bestimmt ist und existiert, wurde dieses Problem behoben und so erhalten wir:

$$\lim_{h \rightarrow 0} \left(\frac{\sin(h)}{h} \cdot \frac{-\sin(h)}{(\cos(h) + 1)} \right) = 1 \cdot \frac{-0}{1 + 1} = 0.$$

Insgesamt erhalten wir also

$$\lim_{h \rightarrow 0} \frac{\cos(h) - 1}{h} = 0.$$

Nun setzen wir alles zusammen, um die Ableitung der Sinusfunktion zu bestimmen. Wir haben bereits gesehen, dass

$$\sin'(x_0) = \lim_{h \rightarrow 0} \left(\sin(x_0) \cdot \frac{\cos(h) - 1}{h} + \cos(x_0) \cdot \frac{\sin(h)}{h} \right)$$

gilt und da $x_0 \in \mathbb{R}$ wie bereits erwähnt fest ist, erhalten wir wieder mit Satz 10.7:

$$\sin'(x_0) = \sin(x_0) \cdot \left(\lim_{h \rightarrow 0} \frac{\cos(h) - 1}{h} \right) + \cos(x_0) \cdot \left(\lim_{h \rightarrow 0} \frac{\sin(h)}{h} \right).$$

Jetzt setzen wir die Werte für diese beiden Grenzwerte ein:

$$\sin'(x_0) = \sin(x_0) \cdot 0 + \cos(x_0) \cdot 1 = \cos(x_0).$$

Die Ableitungsfunktion der Sinusfunktion ist also gegeben durch die Kosinusfunktion. Das ist in Übereinstimmung mit den aus der Schule bekannten Ableitungsregeln. Mit den bereits geführten Vorüberlegungen ist es nun nicht schwer zu zeigen, dass die Ableitungsfunktion der Kosinusfunktion gerade die Funktion

$$\begin{aligned} \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto -\sin(x) \end{aligned}$$

ist.

Nachdem wir nun die Ableitungsfunktionen von Sinus und Kosinus kennen, können wir uns der Frage zuwenden, wie man Sinus und Kosinus von einem vorgegebenen Winkel ausrechnen kann. In manchen Fällen ist es recht einfach und bedarf keiner großen Theorie, wie wir etwa bei $\cos(45^\circ)$ oder bei $\cos(0)$ oder bei $\cos(3\pi)$ gesehen haben. Diese Werte sind jedoch recht speziell. Wir wollen uns im Folgenden Gedanken um eine allgemeinere Winkel machen, und dafür Näherungen durch Polynome angeben. Dabei benutzen wir die uns bekannten Taylor-Polynome. Für Sinus- und Kosinusfunktionen wird es nicht weiter schwer sein, allgemeine Formeln für Taylor-Polynome vom beliebigen Grad anzugeben. Der Grund dafür ist, dass die k -te Ableitungsfunktionen der Sinusfunktion „sich immer wiederholen“. Das wollen wir im Folgenden präziser machen.

Unser Ziel ist es also, Taylor-Polynome der Sinusfunktion (und später auch der Kosinusfunktion) an der Stelle $x_0 = 0$ zu bestimmen. Diese Stelle wurde gewählt, da hier der Funktionswert von Sinus sowie auch von allen Ableitungsfunktionen von Sinus leicht zu bestimmen ist. Nach Definition der Taylor-Polynome brauchen wir also gerade den Wert der k -ten Ableitungsfunktion der Sinusfunktion an der Stelle 0. Dazu schauen wir uns zunächst die ersten vier Ableitungen der Sinusfunktion an. Wir erinnern uns, dass die 0-te Ableitung nach unserer Konvention gerade die Funktion selbst bezeichnet. Wir verwenden die gerade bestimmten Regeln zum Ableiten der trigonometrischen Funktionen.

$$\begin{aligned}\sin^{(0)}(x) &= \sin(x), \\ \sin^{(1)}(x) &= (\sin(x))' = \cos(x), \\ \sin^{(2)}(x) &= (\cos(x))' = -\sin(x), \\ \sin^{(3)}(x) &= (-\sin(x))' = -\cos(x), \\ \sin^{(4)}(x) &= (-\cos(x))' = -(-\sin(x)) = \sin(x).\end{aligned}$$

Das ist gerade die Besonderheit der Sinusfunktion, die vorhin angesprochen wurde. Denn ab jetzt wissen wir bereits vorher, was bei Ableitungen passieren wird: Da die vierte Ableitungsfunktion dieselbe ist wie die nullte, wird die fünfte Ableitung mit der ersten übereinstimmen, die sechste mit der zweiten, die siebte mit der dritten, und so weiter. Die Ableitungen wiederholen sich in einem „Viererzyklus“. (Das kann man auch noch etwas formaler aufschreiben, etwa mit Hilfe der Kongruenzen modulo 4, die wir in Mathematischen Grundlagen 1 kennengelernt haben. Wir verzichten an dieser Stelle darauf.) Nun müssen wir $x_0 = 0$ einsetzen, um die Koeffizienten des Taylor-Polynoms bestimmen zu können. Wir haben uns bereits überlegt, dass

$$\cos(0) = 1 \text{ und } \sin(0) = 0$$

gilt. Wir wollen als Beispiel zunächst das 7-te Taylor-Polynom der Sinusfunktion an der Stelle $x_0 = 0$ bestimmen. Dafür brauchen wir die Werte der k -ten

Ableitungsfunktion der Sinusfunktion, für $0 \leq k \leq 7$, an der Stelle 0. Wir nutzen die oben ausgerechnete Ableitungsfunktionen sowie die Bemerkung bezüglich des „Viererzyklus“ aus und erhalten:

$$\begin{aligned}\sin^{(4)}(0) &= \sin^{(0)}(0) = \sin(0) = 0, \\ \sin^{(5)}(0) &= \sin^{(1)}(0) = \cos(0) = 1, \\ \sin^{(6)}(0) &= \sin^{(2)}(0) = -\sin(0) = 0, \\ \sin^{(7)}(0) &= \sin^{(3)}(0) = -\cos(0) = -1.\end{aligned}$$

Dabei merkt man bereits, dass k -te Ableitung der Sinusfunktion für jedes gerade k an der Stelle 0 den Wert 0 haben wird, denn die geraden Ableitungen der Sinusfunktion sind entweder $\sin(x)$ oder $-\sin(x)$, und beide ergeben 0, wenn wir 0 einsetzen. (Für ein präzises Argument müssten wir einen Induktionsbeweis führen.) Die ungeraden Ableitungen (also die k -ten Ableitungen mit k ungerade) liefern hingegen beim Auswerten an der Stelle 0 abwechselnd 1 und -1 . Wir setzen diese Werte nun in die Definition von $T_0^7 \sin(x)$ ein:

$$\begin{aligned}T_0^7 \sin(x) &= \sum_{k=0}^7 \frac{\sin^{(k)}(0)}{k!} (x-0)^k \\ &= \frac{\sin^{(0)}(0)}{0!} x^0 + \frac{\sin^{(1)}(0)}{1!} x^1 + \frac{\sin^{(2)}(0)}{2!} x^2 + \frac{\sin^{(3)}(0)}{3!} x^3 + \frac{\sin^{(4)}(0)}{4!} x^4 \\ &\quad + \frac{\sin^{(5)}(0)}{5!} x^5 + \frac{\sin^{(6)}(0)}{6!} x^6 + \frac{\sin^{(7)}(0)}{7!} x^7 \\ &= \frac{0}{0!} x^0 + \frac{1}{1!} x^1 + \frac{0}{2!} x^2 + \frac{(-1)}{3!} x^3 + \frac{0}{4!} x^4 + \frac{1}{5!} x^5 + \frac{0}{6!} x^6 + \frac{(-1)}{7!} x^7 \\ &= x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \frac{1}{7!} x^7.\end{aligned}$$

Wir merken, wie das allgemeine Prinzip funktioniert: Die geraden Potenzen von x werden, da die entsprechenden Ableitungen der Sinusfunktion an der Stelle 0 den Wert 0 haben, nicht auftauchen. Die ungeraden Potenzen von x haben abwechselnd Plus oder Minus als Vorzeichen, und haben vom Vorzeichen abgesehen den Koeffizienten $\frac{1}{k!}$ bei der Potenz x^k . Insbesondere ergibt es Sinn, nur ungerade Taylor-Polynome der Sinusfunktion zu betrachten. Diese können wir nun auch etwas formaler angeben. Dabei erinnern wir uns, dass jede ungerade natürliche Zahl in der Form $2l + 1$ geschrieben werden kann für eine natürliche Zahl $l \in \mathbb{N}_0$. Wir erhalten für das Taylor-Polynom vom Grad $2l + 1$ der Sinusfunktion an der Stelle $x_0 = 0$:

$$T_0^{2l+1} \sin(x) = \sum_{k=0}^l \frac{(-1)^k}{(2k+1)!} x^{2k+1}.$$

Dadurch haben wir also Polynome angegeben, die die Sinusfunktion approximieren. Im Fall der Sinusfunktion erhält man so sehr gute Approximationen, wenn man den Grad des Polynoms groß genug wählt.

Als nächstes wollen wir die Taylor-Polynome der Kosinusfunktion bestimmen. Das wird sehr ähnlich sein zu dem Fall der Sinusfunktion. Auch hier werden wir einen „Viererzyklus“ erhalten, der fast genauso wie bei Sinusfunktion aussieht, allerdings um einen Index verschoben. Wir berechnen die k -ten Ableitungsfunktionen der Kosinusfunktion für $0 \leq k \leq 4$, um uns davon zu überzeugen:

$$\begin{aligned}\cos^{(0)}(x) &= \cos(x), \\ \cos^{(1)}(x) &= (\cos(x))' = -\sin(x), \\ \cos^{(2)}(x) &= (-\sin(x))' = -\cos(x), \\ \cos^{(3)}(x) &= (-\cos(x))' = -(-\sin(x)) = \sin(x), \\ \cos^{(4)}(x) &= (\sin(x))' = \cos(x).\end{aligned}$$

Wie angekündigt, erhalten wir auch hier einen „Viererzyklus“. Die Werte der k -ten Ableitungsfunktion der Kosinusfunktion an der Stelle 0 sind für $0 \leq k \leq 7$ gegeben durch:

$$\begin{aligned}\cos^{(4)}(0) &= \cos^{(0)}(0) = \cos(0) = 1, \\ \cos^{(5)}(0) &= \cos^{(1)}(0) = -\sin(0) = 0, \\ \cos^{(6)}(0) &= \cos^{(2)}(0) = -\cos(0) = -1, \\ \cos^{(7)}(0) &= \cos^{(3)}(0) = \sin(0) = 0.\end{aligned}$$

(Wir gehen hier wieder bis $k = 7$, um das Muster nochmal hervorzuheben.)

Hier sind also alle ungeraden Ableitungen, ausgewertet an der Stelle 0, gleich 0. Die Taylor-Polynome von Kosinus haben also nur gerade Potenzen von x , und die allgemeine Formel für das $2l$ -te Taylor-Polynom von Kosinus an der Stelle $x_0 = 0$ lautet also:

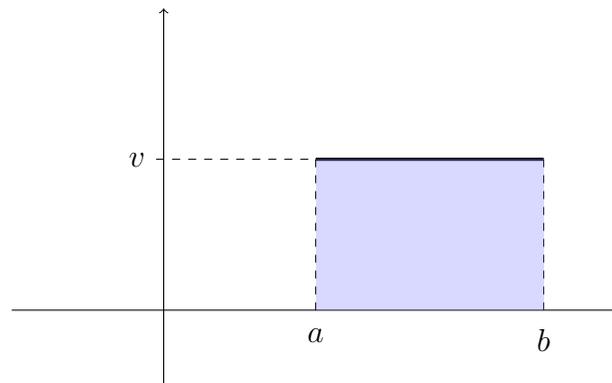
$$T_0^{2l} \cos(x) = \sum_{k=0}^l \frac{(-1)^k}{(2k)!} x^{2k}.$$

Auch für die Kosinusfunktion erhalten wir so gute Approximationen durch Polynome.

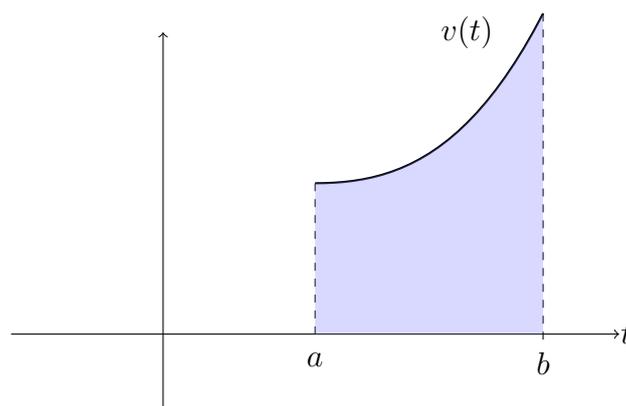
15 Integralrechnung

Nun kommen wir zu einer kurzen Wiederholung einiger Inhalte der Integralrechnung. Während wir uns mit der Differentialrechnung recht ausführlich auseinandergesetzt haben, sollen hier bei der Integralrechnung nur einige Inhalte, die häufig bereits aus der Schule bekannt sind, zusammengestellt werden, und einige typische Beispiele aus der Integralrechnung diskutiert werden.

Während wir den Begriff der Ableitung deutlich präziser gefasst haben als in der Schule üblich, werden wir den Integralbegriff recht vage halten. Die erste Motivation für den Begriff eines Integrals ist die Bestimmung der Fläche unter einem Funktionsgraphen. Das ist beispielsweise in vielen physikalischen Zusammenhängen nötig. Fährt man etwa von einem Zeitpunkt a bis zu einem Zeitpunkt b mit einer konstanten Geschwindigkeit v , so legt man in dieser Zeit die Strecke $v \cdot (b - a)$ zurück. Zeichnet man den Graphen der Geschwindigkeit in Abhängigkeit von der Zeit auf, so ist das gerade die Fläche des Rechtecks, der unter dem Graphen der konstanten Funktion entsteht:



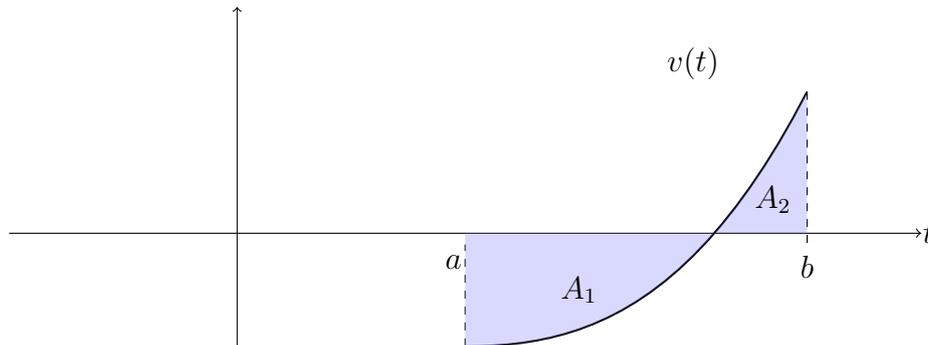
Da nun die Geschwindigkeit in Abhängigkeit von der Zeit bei einer realer Fahrt meist nicht konstant ist, muss man sich überlegen, was eine geeignete Verallgemeinerung der obigen, einfachen Formel zur Streckenberechnung ist. Es stellt sich heraus, dass die zurückgelegte Strecke im Wesentlichen durch die Fläche unter dem Graphen der Geschwindigkeitsfunktion ist:



(Es sei angemerkt, dass es viele weitere Beispiele gibt, gerade aus der Physik, und wir uns auf die Geschwindigkeitsfunktion beschränken, weil hier die physikalischen Zusammenhänge besonders einfach zu erklären sind.)

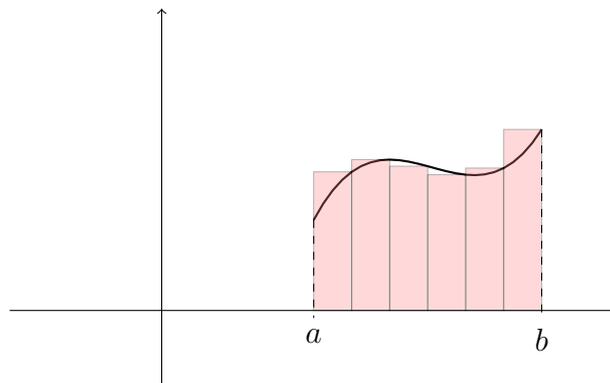
Diese Fläche wird vom Integral berechnet. Es gibt dabei allerdings eine Feinheit, die hier erwähnt werden soll. Diese Feinheit erläutern wir ebenfalls am Beispiel der Geschwindigkeit. Die Geschwindigkeit kann auch negative

Werte annehmen, was physikalisch bedeutet, dass man zurückfährt (oder rückwärts fährt). Hat man in der betrachteten Zeit sowohl positive als auch negative Geschwindigkeiten, fährt man also hin und her, so wird das Integral nicht die gesamte zurückgelegte Strecke, sondern nur die „effektive“ Positionsänderung bestimmen; ist die Geschwindigkeit etwa durch den Graphen



gegeben, so wird das Integral mit den Grenzen von a bis b die Differenz der markierten Flächen $A_2 - A_1$ liefern.

Nun wollen wir eine Idee davon geben, wie das Integral definiert ist. Genau wie bei der Ableitung ist auch beim Integral eine Grenzwertbildung notwendig, allerdings ist diese etwas komplizierter als im Fall der Ableitung. Dabei versucht man, die Fläche unter der Kurve durch Rechtecke anzunähern. Diese Rechtecke sollen immer feiner gewählt werden und die Fläche immer genauer bestimmen. Das kann in etwa wie folgt veranschaulicht werden:

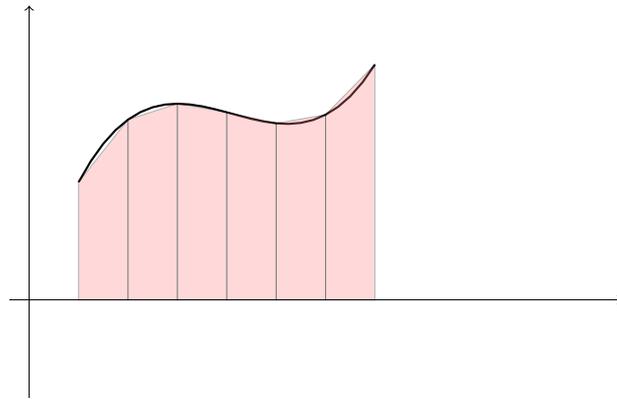


Man sieht, dass eine solche Approximation sowohl Fläche beinhalten wird, die nicht unter dem Graphen ist, als auch manche Teile der Fläche unter dem Graphen unberücksichtigt lässt. Wählt man aber die Rechtecke immer schmaler, so wird der so erhaltene Fehler immer kleiner. Durch geeignete Grenzwertbildung erhält man die Definition des Integrals. Das Integral der Funktion $f: I \rightarrow \mathbb{R}$ zwischen $a, b \in I$ soll mit

$$\int_a^b f(t) dt$$

notiert werden. Dabei ist dt nur ein Platzhalter, der markiert, nach welcher Variable integriert wird. Das ist insbesondere dann wichtig, wenn die Funktion f zusätzlich von Parametern abhängt.

Für präzise Berechnung des Integrals ist diese Definition nicht wirklich hilfreich. Hauptsächlich benutzt man hier stattdessen den Hauptsatz der Differential- und Integralrechnung, an dessen Aussagen wir bald erinnern werden. Hingegen liefert die Definition auch ein erstes Verfahren zum Approximieren des Integrals. Üblicherweise verfeinert man dieses Verfahren in der Praxis, indem man nicht Rechtecke, sondern beispielsweise Trapeze nimmt, um die Fläche unter dem Graphen anzunähern:



In dem Beispiel lässt sich bereits erahnen, dass die Annäherung durch Trapeze bei gleicher Anzahl von Zwischenschritten präziser ist. Gleichzeitig lässt sich die Fläche von einem Trapez, also einem Viereck, in dem zwei Seiten zueinander parallel sind, noch recht einfach ausrechnen. Wir wollen jedoch nicht weiter auf die Näherungen für Integrale eingehen, und wollen nun den Hauptsatz der Differential- und Integralrechnung wiederholen.

In dem Hauptsatz wird die Voraussetzung „stetig“ gebraucht. Wir wollen diesen Begriff nicht präzise definieren. Im Wesentlichen sagt er, dass Grenzwerte der stetigen Funktion durch Einsetzen ausgerechnet werden können, also so, wie im Satz 10.8 bei ganzrationalen Funktionen der Fall war. Jedoch sind fast alle Funktionen, die wir je betrachtet haben, stetig auf ihrem Definitionsbereich, insofern geben wir die Voraussetzung nur an, um die Aussage korrekt zu machen.

Satz 15.1 (Hauptsatz der Differential- und Integralrechnung). *Sei $I \subseteq \mathbb{R}$ ein offenes Intervall. Sei $f: I \rightarrow \mathbb{R}$ eine stetige Funktion und seien a, b Punkte im Intervall I mit $a \leq b$. Sei ferner $F: I \rightarrow \mathbb{R}$ eine differenzierbare Funktion, deren Ableitungsfunktion gerade f ist, also $F' = f$. Dann gilt:*

$$\int_a^b f(t) dt = F(b) - F(a).$$

In diesem Fall wird F eine Stammfunktion von f genannt.

Will man den Hauptsatz anwenden, so braucht man zu einer (stetigen) Funktion $f: I \rightarrow \mathbb{R}$ eine andere, stetig differenzierbare Funktion F , deren Ableitung gerade die ursprüngliche Funktion f ist. Es ist wichtig zu bemerken, dass ein und dieselbe Funktion f unterschiedliche Stammfunktionen haben kann. Allerdings können sich unterschiedliche Stammfunktionen einer Funktion f nur um Konstanten unterscheiden.

Wir betrachten nun ein einfaches Beispiel zur Anwendung des Hauptsatzes.

Beispiel. Wir wollen den Wert des Integrals

$$\int_0^1 t^5 dt$$

bestimmen. Dafür suchen wir eine Stammfunktion der Funktion

$$\begin{aligned} \mathbb{R} &\rightarrow \mathbb{R} \\ t &\mapsto t^5. \end{aligned}$$

Dafür erinnern wir uns, dass man beim Ableiten einer Potenzfunktion $t \mapsto t^n$ erhält:

$$(t^n)' = n \cdot t^{n-1}.$$

Inbesondere gilt:

$$(t^6)' = 6 \cdot t^5.$$

Da wir nun nicht $6t^5$, sondern t^5 als Ableitung erhalten wollen, teilen wir die Funktion durch 6:

$$\left(\frac{t^6}{6}\right)' = \frac{6 \cdot t^5}{6} = t^5.$$

Also ist die Funktion

$$\begin{aligned} \mathbb{R} &\rightarrow \mathbb{R} \\ t &\mapsto \frac{t^6}{6} \end{aligned}$$

eine Stammfunktion unserer ursprünglichen Funktion. Aber auch

$$\begin{aligned} \mathbb{R} &\rightarrow \mathbb{R} \\ t &\mapsto \frac{t^6}{6} + \pi \end{aligned}$$

ist eine Stammfunktion der Funktion $t \mapsto t^5$, denn beim Ableiten der Konstante π erhalten wir 0. Also lässt sich auch diese zweite Stammfunktion verwenden, um mit Hilfe des Hauptsatzes der Differential- und Integralrechnung das fragliche Integral auszurechnen. Wir verwenden, wie meist auch in

der Schule, die abkürzende Notation $F|_a^b$ für $F(b) - F(a)$:

$$\begin{aligned} \int_0^1 t^5 dt &= \left(\frac{t^6}{6} + \pi \right) \Big|_0^1 \\ &= \left(\frac{1^6}{6} + \pi \right) - \left(\frac{0^6}{6} + \pi \right) \\ &= \frac{1}{6} + \pi - \pi = \frac{1}{6}. \end{aligned}$$

Das gesuchte Integral hat also den Wert $\frac{1}{6}$.

Im Folgenden betrachten wir einige weitere Beispiele, anhand derer wir einige Integrationstechniken wiederholen wollen. Zunächst geht es um die partielle Integration. Wir erinnern zunächst an diese Integrationsregel:

Proposition 15.2 (Partielle Integration). *Will man das Produkt zweier Funktionen integrieren, von denen eine bereits als Ableitung einer anderen Funktion erkannt wurde, so kann unter Umständen die partielle Integration hilfreich sein. Seien dafür $I \subseteq \mathbb{R}$ offenes Intervall und $u, v: I \rightarrow \mathbb{R}$ stetig differenzierbare Funktionen. Dann gilt:*

$$\int_a^b u(x)v'(x)dx = (u(b)v(b) - u(a)v(a)) - \int_a^b u'(x)v(x)dx.$$

Die partielle Integration ist das Gegenstück der Produktregel beim Ableiten. Man sieht, dass hier auch ähnliche Terme wie in der Produktregel vorkommen. Wir erläutern die Anwendungsmöglichkeiten der partiellen Integration exemplarisch an einem Beispiel.

Beispiel. Wir wollen das Integral

$$\int_1^2 \ln(t) dt$$

berechnen. Da wir hier zunächst kein Produkt haben, verwenden wir den Trick, den wir sonst bereits häufiger benutzt haben, und schreiben statt $\ln(t)$ nämlich $1 \cdot \ln(t)$, was natürlich das Integral nicht ändert. Nun ist es leicht zu sehen, dass 1 die Ableitung einer Funktion ist, beispielsweise von $t \mapsto t$. Wir setzen also $v(t) = t$ und $u(t) = \ln(t)$ in der partiellen Integration. Dann

erhalten wir:

$$\begin{aligned}
 \int_1^2 \ln(t) dt &= \int_1^2 1 \cdot \ln(t) dt \\
 &= (t \ln(t)) \Big|_1^2 - \int_1^2 t \cdot \frac{1}{t} dt \\
 &= (t \ln(t)) \Big|_1^2 - \int_1^2 1 dt \\
 &= (t \ln(t)) \Big|_1^2 - t \Big|_1^2 \\
 &= (t \ln(t) - t) \Big|_1^2.
 \end{aligned}$$

Dabei haben wir benutzt, dass die Ableitungsfunktion der Logarithmusfunktion gerade $t \mapsto \frac{1}{t}$ ist, und außerdem erneut, dass $t \mapsto t$ eine Stammfunktion der konstanten Funktion $t \mapsto 1$ ist. An dieser Stelle sei angemerkt, dass

$$\begin{aligned}
 F: (0, \infty) &\rightarrow \mathbb{R} \\
 t &\mapsto t \ln(t) - t
 \end{aligned}$$

eine Stammfunktion des Logarithmus ist, was durch Ableiten bewiesen werden kann. Um das Integral auszurechnen, müssen wir noch die Grenzen einsetzen:

$$\begin{aligned}
 \int_1^2 \ln(t) dt &= (t \ln(t) - t) \Big|_1^2 \\
 &= (2 \cdot \ln(2) - 2) - (1 \cdot \ln(1) - 1).
 \end{aligned}$$

Während $\ln(2)$ nicht weiter vereinfacht werden kann, haben wir $\ln(1)$ bereits ausgerechnet, und es gilt $\ln(1) = 0$. Also erhalten wir:

$$\begin{aligned}
 \int_1^2 \ln(t) dt &= (2 \ln(2) - 2) - (1 \ln(1) - 1) \\
 &= 2 \ln(2) - 1.
 \end{aligned}$$

Als nächstes wollen wir zwei Beispiele zur Substitutionsregel betrachten. Zunächst wiederholen wir die Substitutionsregel:

Proposition 15.3 (Substitutionsregel). *Die Substitutionsregel erlaubt es, die Verkettung zweier Funktionen mit einem geeigneten Zusatzfaktor zu integrieren. Seien dafür $I, J \subseteq \mathbb{R}$ offene Intervalle und $g: J \rightarrow I$ und $f: I \rightarrow \mathbb{R}$ differenzierbare Funktionen. Seien ferner a, b Punkte in J mit $a \leq b$. Dann*

gilt:

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(t)dt.$$

Die Substitutionsregel ist das Gegenstück der Kettenregel beim Ableiten. Man sollte beachten, dass sich hier die Integrationsgrenzen ändern; um das nochmal zu betonen, wurde die Integrationsvariable umbenannt. Es gibt viele unterschiedliche Anwendungsbeispiele für die Substitutionsregel. Dabei gibt es zwei grundsätzlich verschiedene Einsatzmöglichkeiten. In der ersten, unmittelbarereren, erkennt man in dem zu berechnendem Integral die linke Seite der Substitutionsregel wieder, und kann diese dann durch die rechte Seite ersetzen. Etwas weniger intuitiv ist die umgekehrte Richtung, bei der man eine „innere“ Funktion „hineinschmuggelt“, um die Berechnung des Integrals zu vereinfachen. Wir wollen zu beiden Möglichkeiten jeweils ein Beispiel angeben.

Beispiel. Wir wollen das Integral

$$\int_0^{\frac{\sqrt{\pi}}{2}} \cos(t^2)t dt$$

berechnen. Die Funktion $t \mapsto \cos(t^2)$ kann als Verkettung der Funktionen $t \mapsto t^2$ und $u \mapsto \cos(u)$ gesehen werden. Um die Substitutionsregel anzuwenden, muss also die Ableitung der inneren Funktion als Faktor im Integrand (also in dem zu integrierenden Term) vorkommen. Die Ableitung von $t \mapsto t^2$ ist die Funktion $t \mapsto 2t$. Uns fehlt also der Faktor 2, um die Substitutionsregel anwenden zu können. Wir ergänzen also den Integrand um den Faktor $1 = \frac{1}{2} \cdot 2$ und benutzen, dass feste Faktoren aus dem Integral „herausgezogen“ werden können:

$$\begin{aligned} \int_0^{\frac{\sqrt{\pi}}{2}} \cos(t^2)t dt &= \int_0^{\frac{\sqrt{\pi}}{2}} \frac{1}{2} \cdot 2 \cdot \cos(t^2)t dt \\ &= \frac{1}{2} \cdot \int_0^{\frac{\sqrt{\pi}}{2}} \cos(t^2) \cdot 2t dt. \end{aligned}$$

Dieses Integral entspricht nun genau der linken Seite der Substitutionsregel. Wir wenden diese an. Dafür müssen wir zunächst die neuen Integralgrenzen ausrechnen:

$$\begin{aligned} \left(\frac{\sqrt{\pi}}{2}\right)^2 &= \frac{\pi}{4} \\ 0^2 &= 0 \end{aligned}$$

Also gilt:

$$\begin{aligned}\int_0^{\frac{\sqrt{\pi}}{2}} \cos(t^2)t dt &= \frac{1}{2} \cdot \int_0^{\frac{\sqrt{\pi}}{2}} \cos(t^2) \cdot 2t dt \\ &= \frac{1}{2} \cdot \int_0^{\frac{\pi}{4}} \cos(u) du.\end{aligned}$$

Wir hatten uns bereits überlegt, dass $(\sin(u))' = \cos(u)$ ist, also ist die Sinusfunktion eine Stammfunktion der Kosinusfunktion und wir können nun den Hauptsatz der Differential- und Integralrechnung anwenden:

$$\begin{aligned}\int_0^{\frac{\sqrt{\pi}}{2}} \cos(t^2)t dt &= \frac{1}{2} \cdot \int_0^{\frac{\pi}{4}} \cos(u) du \\ &= \frac{1}{2} (\sin(u)) \Big|_0^{\frac{\pi}{4}} \\ &= \frac{1}{2} \left(\sin\left(\frac{\pi}{4}\right) - \sin(0) \right).\end{aligned}$$

Wir hatten bereits die beiden fraglichen Sinuswerte im Wesentlichen ausgerechnet:

$$\sin\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}} \text{ und } \sin(0) = 0.$$

(Tatsächlich haben wir explizit nur $\cos(45^\circ)$ bestimmt, aber die Berechnung für $\sin(45^\circ)$ verläuft fast exakt genauso.) Wir setzen diese Werte ein und erhalten:

$$\int_0^{\frac{\sqrt{\pi}}{2}} \cos(t^2)t dt = \frac{1}{2} \left(\sin\left(\frac{\pi}{4}\right) - \sin(0) \right) = \frac{1}{2\sqrt{2}}.$$

Nun kommen wir zu einem Beispiel der komplizierteren Anwendungsart für die Substitutionsregel.

Beispiel. Wir wollen das Integral

$$\int_0^1 \sqrt{1-x^2} dx$$

berechnen. Als Faustregel kann man sagen, dass die Integration von Wurzeltermen häufig etwas schwieriger ist. In diesem Fall lohnt sich die sogenannte *trigonometrische Substitution*. In diesem Fall setzen wir für x , wie angekündigt, eine andere Funktion ein, und benutzen die Substitutionsregel „von rechts nach links“. Dabei entsteht ein vermeintlich schwierigerer Ausdruck, der dann jedoch einfacher zu integrieren ist. In diesem Fall setzen

wir $x = \sin(t)$ ein. Nach der Substitutionsregel erhalten wir als einen zusätzlichen Faktor die Ableitung der eingesetzten Funktion, in diesem Fall $\sin'(t) = \cos(t)$. Ferner müssen wir die neuen Integrationsgrenzen finden. Dabei brauchen wir also reelle Zahlen $a, b \in \mathbb{R}$ mit $\sin(a) = 0$ und $\sin(b) = 1$. Diese Zahlen a, b werden im Allgemeinen und auch in diesem Fall nicht eindeutig sein; es ist aber sinnvoll, die einfachsten solchen Werte zu wählen. Wir wissen bereits, dass $\sin(0) = 0$ gilt. Ferner lässt sich leicht aus der Definition herleiten, dass $\sin\left(\frac{\pi}{2}\right) = 1$ ist. Diese Integralgrenzen können wir nun an dieser Stelle verwenden und erhalten mit der Substitutionsregel:

$$\int_0^1 \sqrt{1-x^2} dx = \int_0^{\frac{\pi}{2}} \sqrt{1-(\sin(t))^2} \cdot \cos(t) dt$$

Nun sieht das zunächst nicht einfacher aus als zuvor. Wir erinnern uns jedoch an die wichtige Identität

$$(\sin(t))^2 + (\cos(t))^2 = 1.$$

Durch Umstellen dieser Identität erhalten wir

$$1 - (\sin(t))^2 = (\cos(t))^2.$$

Das können wir nun im obigen Integral einsetzen. Dabei beachten wir, dass für $0 \leq t \leq \frac{\pi}{2}$ die Kosinuswerte nicht-negativ sind, sodass

$$\sqrt{(\cos(t))^2} = |\cos(t)| = \cos(t)$$

gilt. Also vereinfacht sich das Integral nun zu

$$\int_0^1 \sqrt{1-x^2} dx = \int_0^{\frac{\pi}{2}} (\cos(t))^2 dt.$$

Den Wert dieses Integrals kann man nun wiederum mit Hilfe partieller Integration ermitteln. Wir wollen an dieser Stelle auf die Details verzichten und notieren nur das Ergebnis:

$$\int_0^1 \sqrt{1-x^2} dx = \int_0^{\frac{\pi}{2}} (\cos(t))^2 dt = \frac{\pi}{4}.$$

Teil III

Numerische Verfahren

Zum Schluss wollen wir noch einen ganz kurzen Blick auf die numerischen Verfahren der Mathematik wagen. Dabei handelt es sich um Fragestellungen, die das Theoretische mit den praktischen Fragestellungen verknüpfen.

Es geht insbesondere darum, Näherungslösungen für mathematische Probleme zu finden, die unmöglich exakt gelöst werden können, oder bei denen die exakten Lösungen zu schwierig oder zu aufwendig wären, beispielsweise auch, wenn man nur eine gewisse Präzision der Lösung braucht. Eines der Probleme der exakten Lösungen ist die Unmöglichkeit (oder jedenfalls große Schwierigkeit) der Speicherung unendlicher nicht-periodischer Dezimalzahlen (also irrationaler reeller Zahlen, wie beispielsweise $\sqrt{2}$, π oder e) in einem Computer. Also rechnet man (fast) zwangsweise mit Näherungen, und muss sich sowohl mit der Frage beschäftigen, wie man gute Näherungen erhält, als auch damit, wie gut die existierenden Näherungen sind. Die Numerik, die sich mit solchen Fragestellungen beschäftigt, ist ein riesiges Gebiet, in das wir nur ganz wenig Einblick an zwei exemplarischen Fragestellungen erhalten werden.

Heutzutage hat die Numerik, dank der Existenz von Computern, eine besondere Relevanz. Allerdings sind die Fragestellungen natürlich nicht neu, und auch vor Computern gab es beispielsweise Verfahren, mit denen man Hunderte Nachkommastellen von π ermittelt hat. Allerdings müssen die numerischen Verfahren an die Technik angepasst sein, und die Relevanz unterschiedlicher Verfahren hat sich stark gewandelt und wandelt sich immer noch, etwa dank der großer Verfügbarkeit vom preiswerten Speicherplatz für die Rechner.

Dabei gibt es bei jedem Verfahren, das auf ein bestimmtes Problem oder eine Klasse von Problemen zugeschnitten wird, immer wieder ähnliche Anforderungen. Natürlich möchte man ein numerisches Verfahren zur Lösung des jeweiligen Problems haben, das schnell ist, nur geringe Ungenauigkeiten aufweist, in allen Fällen anwendbar ist und möglichst wenig Speicherplatz verbraucht. Das funktioniert so natürlich nicht, aber es zeigt einige Kriterien auf, nach denen ein Verfahren beurteilt werden kann.

Insbesondere sollte man sich vor Augen führen, dass jedes numerische Verfahren fehlerbehaftet ist. Wir wollen an dieser Stelle auf drei wichtige Fehlerquellen hinweisen.

- Die Eingabedaten eines Verfahrens, die beispielsweise aus statistischen oder physikalischen Messungen stammen, oder aus anderen Rechnungen, sind meist bereits fehlerbehaftet. Allerdings „reagieren“ die Verfahren auf solche Fehler in Eingabedaten unterschiedlich. Man spricht auch von der „Kondition“ eines Problems. Bei einer schlecht konditionierten Fragestellung variieren die Ergebnisse stark bereits bei kleinen Fehlern in den Eingabedaten. Wir wollen ein typisches Beispiel dafür anführen.

Beispiel. Die Frage nach der inversen Matrix zu einer vorgegebenen Matrix ist schlecht konditioniert. Beispielsweise ist die Matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

nicht invertierbar. Betrachtet man stattdessen für eine kleine reelle Zahl ε die Matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & \varepsilon \end{pmatrix},$$

so ist diese Matrix stets invertierbar. Außerdem werden die Einträge der inversen Matrix zum Teil sehr groß sein und ebenfalls stark variieren, wenn man ε nur leicht verändert. Hat man etwa $\varepsilon = 10^{-7}$, so erhält man als inverse Matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 10^7 \end{pmatrix},$$

man hat also insbesondere einen Eintrag, der 10 Millionen ist. Das ist einer der Gründe, weswegen von der Berechnung der inversen Matrix abgeraten wird. Arbeitet man mit „geschickteren“ Versionen des Gauß-Verfahrens, so tritt dieses Problem nicht auf.

Als nächstes wollen wir auf die systematische Fehler im Verfahren hinweisen. Damit ist nicht gemeint, dass man das Verfahren falsch anwendet (auch wenn das in der Realität aber ebenfalls eine Fehlerquelle ist, gegen die wir mathematisch allerdings wenig ausrichten können). Vielmehr sind einige Verfahren so konzipiert, dass sie fast nie das präzise Ergebnis liefern werden. Will man beispielsweise den Wert $\sin(x)$ ausrechnen und verwendet dafür das fünfte Taylor-Polynom

$$T_0^5 \sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!},$$

so lässt man sich bei fast allen Werten von x darauf ein, dass das Ergebnis unpräzise ist, auch wenn diese Näherung ganz gute Ergebnisse liefert. In solchen Fällen ist es wichtig zu wissen, wie groß der durch das Verfahren erzeugte Fehler ist.

Schließlich gibt es bei so gut wie jeder numerisch durchgeführten Rechnung Rundungsfehler. Daher ist es sinnvoll, das Verfahren so zu gestalten, dass die Rundungsfehler nur im geringen Maße auftreten, und die auftretenden Rundungsfehler auch abzuschätzen und zu berücksichtigen.

Nach dieser Einleitung widmen wir uns zwei Beispielen von numerischen Verfahren. Die Weiterentwicklungen dieser Verfahren werden in der Praxis eingesetzt, während die Grundversion leicht zu verstehen ist, was ausschlaggebend für die Auswahl dieser Beispiele war.

16 Vektoriteration

Vektoriteration (engl. *power method*, auch im deutschen manchmal Potenzmethode) ist ein Verfahren zur Approximation der Eigenwerte und Eigenvektoren einer Matrix. Solche Eigenwertprobleme tauchen in diversen realen Sachverhalten auf, etwa bei Bildkompression, beim Google-PageRank-Algorithmus (jedenfalls in seiner öffentlich bekannten, älteren Form) oder bei der Suche nach stabilen Zuständen in Vorgängen mit gewissen Veränderungsmustern (diese werden unter dem Stichwort „Markow-Ketten“ behandelt).

Es ist also eine $n \times n$ -Matrix A vorgegeben, und wir suchen die Vektoren $v \in \mathbb{R} \setminus \{0\}$ und reelle Zahlen $\lambda \in \mathbb{R}$, sodass

$$Av = \lambda v$$

gilt. Wir haben das Problem mit einem naiven Ansatz, allerdings präzise, für 2×2 - und 3×3 -Matrizen behandelt. Bereits diese Fälle wurden schnell kompliziert. In realen Anwendungen ist n sehr groß, und unsere naive Methode nicht anwendbar.

Es existiert eine kompliziertere Methode, um auf präzise Eigenwerte der Matrix A zu kommen. Dabei ermittelt man das sogenannte *charakteristische Polynom* der Matrix A , das die Eigenschaft hat, genau die Eigenwerte der Matrix als Nullstellen zu haben. Das charakteristische Polynom einer Matrix ist recht aufwendig zu bestimmen. Wir hatten bereits auch bei unserer naiven Methode festgestellt, dass bei der Eigenwertsuche bei 2×2 -Matrizen manchmal quadratische Gleichungen gelöst werden müssen. Für 3×3 -Matrizen muss man im Allgemeinen die Nullstellen eines Polynoms dritten Grades finden, und noch allgemeiner bei $n \times n$ -Matrizen kommt man zu dem charakteristischen Polynom vom Grad n . Allerdings ist die Lösung der polynomiellen Gleichungen für Polynome vom Grad $n \geq 5$ im Allgemeinen nicht mit einer Lösungsformel wie der $p - q$ -Formel möglich, und stellt ein schwieriges Problem dar. Erschwerend kommt hinzu, dass die Frage nach Nullstellen vom Polynom mit vorgegebenen Koeffizienten schlecht konditioniert ist: Bereits kleine Abweichungen in den Koeffizienten können große Änderungen der Nullstellen erzeugen; insbesondere kann es passieren, dass man bereits durch kleine Änderung der Koeffizienten beispielsweise aus einem Polynom ohne reellen Nullstellen ein Polynom mit reellen Nullstellen erhält und umgekehrt. Diese Überlegungen bewirken, dass das charakteristische Polynom zwar vom großen theoretischen Nutzen ist, aber nicht zur numerischen Ermittlung der Eigenwerte eingesetzt wird.

Stattdessen nutzt man das Verfahren der Vektoriteration.

Wie der Name schon sagt, handelt es sich dabei um ein iteratives Verfahren. Man beginnt mit einem (fast) beliebigen Vektor $x_0 \in \mathbb{R}^n \setminus \{0\}$. In jedem Schritt multipliziert man den vorherigen Vektor mit der Matrix A , deren Eigenwerte wir ermitteln wollen, und ermitteln so die Vektoren

$$x_{k+1} = A \cdot x_k.$$

Insbesondere hat man also $x_1 = Ax_0$, dann $x_2 = Ax_1 = A^2x_0$, als nächstes $x_3 = Ax_2 = A^3x_0$, und so weiter. Allgemeiner erhalten wir $x_k = A^k \cdot x_0$, was den Namen „Potenzmethode“ erklärt. Nun stellt sich heraus, dass unter gewissen Voraussetzungen an den Vektor x_0 und die Matrix A diese Vektoren etwas mit einem Eigenvektor von A zu tun haben. Wir würden gerne sagen, dass die Vektoren x_k einen Eigenvektor von A annähern. Das funktioniert jedoch im Allgemeinen nicht, und ein wesentliches Problem dabei ist die Tatsache, dass die Komponenten des Vektors x_k mit wachsendem k im Allgemeinen unbeschränkt wachsen und keine reellen Zahlen approximieren, sondern beliebig groß werden. Dagegen hat man verschiedene Auswege, die im Wesentlichen auf ähnliche Ergebnisse Hinauslaufen. Die Idee dabei ist, die Einträge der Vektoren x_k zu skalieren, sodass die Einträge nicht mehr unbeschränkt wachsen. Das kann man auf verschiedene Arten erreichen; wir stellen im Folgenden zwei Möglichkeiten dafür vor.

Die erste Möglichkeit besteht darin, aus den bisherigen Iterationen x_k die neuen Vektoren y_k zu erhalten, indem man jeweils durch die Norm der Vektoren x_k teilt, also setzt man

$$y_k = \frac{x_k}{\|x_k\|}.$$

Dadurch entsteht in jedem Schritt ein Vektor, dessen Komponenten alle im Bereich zwischen -1 und 1 liegen. Diese Vektoren haben nun (in den meisten Fällen) die gewünschte Eigenschaft, einen Eigenvektor von A anzunähern. Etwas präziser: Die Folge der so entstandenen Vektoren $(y_k)_{k \in \mathbb{N}}$ konvergiert (unter milden Voraussetzungen an die Matrix A und den Startvektor x_0) gegen einen Eigenvektor der Matrix A , der zum betragsgrößten Eigenwert gehört. Dabei sind einige Erläuterungen angebracht. Zunächst haben wir den Begriff der Konvergenz soweit nur für reelle Zahlen definiert; jetzt sprechen wir von der Konvergenz von Vektoren. Dabei ist die Konvergenz komponentenweise zu verstehen: Die Folge der ersten Komponenten der $(y_k)_{k \in \mathbb{N}}$ ist eine konvergente Folge von reellen Zahlen, die gegen die erste Komponente des erwähnten Eigenvektors konvergiert, die Folge der zweiten Komponenten der $(y_k)_{k \in \mathbb{N}}$ konvergiert gegen die zweite Komponente des gesuchten Eigenvektors, und so weiter.

Als zweites soll darauf eingegangen werden, gegen welchen Eigenvektor von A die Folge (y_k) konvergiert. Wir haben bereits erwähnt, dass eine $n \times n$ -Matrix bis zu n reelle Eigenwerte haben kann und wir schreiben für diese Eigenwerte $\lambda_1, \dots, \lambda_n$. Wir wollen nun annehmen, dass die so eingeordneten Eigenwerte der Größe ihrer Beträge nach geordnet sind, also dass

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$$

gilt. Das ist leicht durch Ummummerierung der Eigenwerte zu erreichen. Die obige Aussage besagt also, dass in dieser Nummerierung die Folge (y_k) gegen

einen Eigenvektor mit Eigenwert λ_1 konvergiert. Wir werden später eine Beweisskizze für diese Aussagen sowie ein Beispiel dazu betrachten. Zunächst wollen wir jedoch noch kurz auf die zweite Skalierungsmethode eingehen.

Anstelle durch die Norm zu teilen, was das Berechnen von Wurzeln notwendig macht, können wir den Vektor x_k durch den Betrag der betragsgrößten Komponente des Vektors x_k teilen. Ist a_k die betragsgrößte Komponente von x_k , so setzen wir

$$z_k = \frac{x_k}{|a_k|}.$$

Dadurch erreichen wir, dass der betragsgrößte Eintrag in z_k genau ± 1 ist. Insbesondere liegen auch hier alle Komponenten des Vektors z_k zwischen -1 und 1 . Auch für die Folge (z_k) erhalten wir eine ähnliche Konvergenzaussage.

Wie die meisten numerischen Verfahren, funktioniert auch dieses Verfahren nur unter gewissen Voraussetzungen; auch unterscheidet sich die Konvergenzgeschwindigkeiten - also die Anzahl der Iterationen, die man braucht, um eine Näherung gewisser Präzision zu erreichen - je nach Anfangsvoraussetzungen. Wir notieren einige dieser Voraussetzungen; später geben wir eine Beweisidee dafür, warum das Verfahren funktioniert. Da wird man sehen können, warum diese Voraussetzungen das Problem stark vereinfachen.

Zunächst brauchen wir gewisse Voraussetzungen an die Matrix A . Das Verfahren funktioniert insbesondere dann gut, wenn die Matrix A diagonalisierbar ist. Zur Erinnerung: Das heißt, dass es eine Basis von \mathbb{R}^n gibt, die aus Eigenvektoren von A besteht. Es mag etwas seltsam anmuten, dass man, um zu prüfen, ob das Verfahren zur Bestimmung der Eigenvektoren funktionieren wird, bereits wissen muss, ob es eine Basis von \mathbb{R}^n aus Eigenvektoren gibt. Allerdings gibt es einige einfache hinreichende Kriterien, um das zu überprüfen. Beispielsweise sind symmetrische Matrizen, also solche, für die $A^T = A$ gilt (wobei A^T die transponierte Matrix zu A bezeichnet), stets diagonalisierbar. Dieses Kriterium ist leicht zu überprüfen; außerdem kommen symmetrische Matrizen häufig in den Anwendungen vor. Ferner funktioniert das Verfahren in vielen Fällen auch für nicht-diagonalisierbare Matrizen, allerdings nicht immer, wie wir später sehen werden.

Eine weitere Voraussetzung betrifft die Eigenwerte der Matrix A . Wir hatten bereits gesagt, dass wir diese nach der Größe der Beträge ordnen:

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$$

Nun setzen wir zusätzlich voraus, dass $|\lambda_1|$ strikt größer als $|\lambda_2|$ ist. Tatsächlich funktioniert das Verfahren umso besser, je größer der Unterschied zwischen den betragsgrößten Eigenwerten ist, wie wir sowohl in den Beispielen als auch in der Beweisskizze zur Funktionsweise des Verfahrens sehen werden.

Als letztes braucht man noch eine gewisse Voraussetzung an den Startvektor x_0 , die wir später erläutern werden. Diese wird nicht wirklich im Voraus zu testen sein, allerdings ist diese Voraussetzung nicht wirklich entsprechend.

Nun, da einige grundlegende Fragen zum Verfahren und seiner Anwendbarkeit vorerst geklärt sind, wollen wir das Verfahren an einem Beispiel veranschaulichen. Für die Rechnungen wurde hierbei stets SageMath verwendet.

Beispiel. Wir suchen die Eigenwerte der Matrix

$$A = \begin{pmatrix} 245 & -254 & -252 & -46 & -224 \\ 161 & -168 & -174 & -32 & -148 \\ -39 & 40 & 45 & 7 & 38 \\ 27 & -28 & -32 & -6 & -26 \\ 110 & -113 & -110 & -21 & -101 \end{pmatrix}.$$

Man merkt sofort, dass bereits bei dieser 5×5 -Matrix unsere früheren Methoden sehr mühsam anzuwenden wären. Wir wenden hier die Vektoriteration mit dem Startvektor

$$x_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

an.

Die ersten vier Iterationen liefern

$$x_1 = \begin{pmatrix} -531 \\ -361 \\ 91 \\ -65 \\ -235 \end{pmatrix}, x_2 = \begin{pmatrix} -5703 \\ -3817 \\ 979 \\ -641 \\ -2527 \end{pmatrix}, x_3 = \begin{pmatrix} -78891 \\ -52765 \\ 13279 \\ -8885 \\ -35011 \end{pmatrix}, x_4 = \begin{pmatrix} -1021119 \\ -681529 \\ 171091 \\ -113969 \\ -453559 \end{pmatrix}.$$

(Dabei wird die Iterationsvorschrift $x_{k+1} = Ax_k$ angewandt.)

An diesen Zahlen merkt man bereits, dass die Einträge der Vektoren x_k sehr groß werden, wie auch früher angekündigt. Nun führen wir die Skalierung durch, zuerst mit der Norm der Vektoren. Für die Zwecke dieses Beispiels wurde zunächst eine recht geringe Genauigkeit gewählt, die bereits genügt, um das Verfahren zu illustrieren. Wir verwenden der Einfachheit hier trotzdem das Gleichheitszeichen.

$$y_1 = \begin{pmatrix} -0,77 \\ -0,52 \\ 0,13 \\ -0,094 \\ -0,34 \end{pmatrix}, y_2 = \begin{pmatrix} -0,77 \\ -0,52 \\ 0,13 \\ -0,087 \\ -0,34 \end{pmatrix}, y_3 = \begin{pmatrix} -0,77 \\ -0,51 \\ 0,13 \\ -0,087 \\ -0,34 \end{pmatrix}, y_4 = \begin{pmatrix} -0,77 \\ -0,51 \\ 0,13 \\ -0,08 \\ -0,34 \end{pmatrix}.$$

Hier merkt man bereits, dass die Einträge sich „stabilisieren“. Ferner merken wir nochmal an, dass die Komponenten der Vektoren y_i stets zwischen -1 und 1 liegen.

Führt man die Skalierung an der betragsgrößten Komponente durch, so erhält man Vektoren z_i , die gegen einen anderen Vektor konvergieren (hier wieder in recht geringen Genauigkeit):

$$z_1 = \begin{pmatrix} -1,0 \\ -0,68 \\ 0,17 \\ -0,12 \\ -0,44 \end{pmatrix}, z_2 = \begin{pmatrix} -1,0 \\ -0,67 \\ 0,17 \\ -0,11 \\ -0,44 \end{pmatrix}, z_3 = \begin{pmatrix} -1,0 \\ -0,67 \\ 0,17 \\ -0,11 \\ -0,44 \end{pmatrix}, z_4 = \begin{pmatrix} -1,0 \\ -0,67 \\ 0,17 \\ -0,11 \\ -0,44 \end{pmatrix}.$$

Man merkt auch hier, dass die Komponenten der Vektoren stets zwischen -1 und 1 liegen.

Bereits bei der recht geringen Rechengenauigkeit merken wir, dass wir eine Approximation von einem Eigenvektor erhalten haben. Das ist beim Vektor z_4 etwas einfacher zu erklären. Wendet man nämlich auf z_4 die Matrix A an, so erhält man (im Rahmen der Rechengenauigkeit):

$$A \cdot z_4 = \begin{pmatrix} -13 \\ -8,8 \\ 2,2 \\ -1,5 \\ -5,9 \end{pmatrix}$$

Vergleicht man die erste Komponente von z_4 mit der ersten Komponente von Az_4 , so stellt man fest, dass - sollte z_4 ein Eigenvektor sein - der Eigenwert 13 sein müsste. Also vergleichen wir Az_4 mit $13z_4$ (wieder im Rahmen der gewählten Rechengenauigkeit):

$$13 \cdot z_4 = \begin{pmatrix} -13 \\ -8,7 \\ 2,2 \\ -1,5 \\ -5,8 \end{pmatrix}$$

Das macht die Vermutung plausibel, dass z_4 eine Approximation an ein Eigenvektor der Matrix A zum Eigenwert 13 ist. Obwohl der Vektor y_4 ganz anders aussieht, liefert auch dieser eine Approximation zu einem Eigenvektor mit Eigenwert 13 . Dazu vergleichen wir, wieder approximativ:

$$A \cdot y_4 = \begin{pmatrix} -10 \\ -6,8 \\ 1,7 \\ -1,1 \\ -4,5 \end{pmatrix}$$

mit

$$13 \cdot y_4 = \begin{pmatrix} -10 \\ -6,7 \\ 1,7 \\ -1,1 \\ -4,5 \end{pmatrix}.$$

Man stellt außerdem fest, dass die Folgen (y_k) und (z_k) zwar gegen unterschiedliche Eigenvektoren der Matrix A konvergieren, allerdings nicht wesentlich unterschiedliche, denn die beiden sind Vielfachen (im Rahmen der Rechengenauigkeit) voneinander:

$$1,3 \cdot y_4 = \begin{pmatrix} -1,0 \\ -0,67 \\ 0,17 \\ -0,11 \\ -0,45 \end{pmatrix} \text{ vs. } z_4 = \begin{pmatrix} -1,0 \\ -0,67 \\ 0,17 \\ -0,11 \\ -0,44 \end{pmatrix}.$$

Führt man die Rechnung mit etwas höherer Genauigkeit bis zum 8-ten Iterationsschritt durch, so erhält man

$$z_8 = \begin{pmatrix} -1.0000000 \\ -0.66670458 \\ 0.16670492 \\ -0.11113638 \\ -0.44443180 \end{pmatrix}$$

Hier kann man bereits einen Eigenvektor mit rationalen Einträgen vermuten, der approximiert wird. (Natürlich müssen die Eigenvektoren im Allgemeinen nicht rational sein; dieses Beispiel wurde allerdings so gewählt, dass man den Eigenvektor tatsächlich präzise angeben kann, und damit die Darstellung zu vereinfachen.) Tatsächlich stellt man fest, dass der Vektor

$$v = \begin{pmatrix} -1 \\ -2/3 \\ 1/6 \\ -1/9 \\ -4/9 \end{pmatrix}$$

ein Eigenvektor der Matrix A zum Eigenwert 13 ist (und das wirklich präzise). Durch einen Vorfaktor erhält man - in diesem zur Vorführungs-zwecken gewählten Beispiel - sogar einen Eigenvektor von A zum Eigenwert 13, dessen alle Komponenten ganze Zahlen sind:

$$18v = \begin{pmatrix} -18 \\ -12 \\ 3 \\ -2 \\ -8 \end{pmatrix}.$$

Nun haben wir die Vektoriteration erfolgreich eingesetzt, um einen Eigenvektor und einen Eigenwert der Matrix in dem Beispiel zu finden. Da der Eigenwert 13, wie sich herausstellt, tatsächlich der betragsgrößte Eigenwert dieser Matrix ist, werden wir durch erneute Anwendung der Vektoriteration, etwa mit einem anderen Startvektor, „mit hoher Wahrscheinlichkeit“ (und das kann man präzisieren) wieder eine Folge erhalten, die gegen ein Vielfaches des bereits gefundenen Eigenvektors konvergiert. Wir brauchen also eine Modifikation des Verfahrens, um weitere Eigenvektoren und Eigenwerte zu finden.

Die erste Modifikation wird uns erlauben, den Eigenvektor zu finden, der die betragsmäßig größte Differenz zu der vorgegebenen Zahl hat. Ist eine Zahl $c \in \mathbb{R}$ vorgegeben, so können wir die Matrix $A - c \cdot E_n$ betrachten. Diese Matrix sieht fast genauso aus wie die Matrix A , allerdings wurde von den Einträgen auf der Diagonale jeweils c subtrahiert. Ist v nun ein Eigenvektor der Matrix A zum Eigenwert λ , so erhalten wir:

$$(A - cE_n) \cdot v = Av - cE_nv = \lambda v - cv = (\lambda - c)v.$$

Dabei haben wir das Distributivgesetz ausgenutzt, die Definition eines Eigenwerts sowie die Tatsache, dass die Multiplikation mit der Einheitsmatrix einen Vektor nicht ändert. Also ist $v \neq 0$ auch ein Eigenvektor der Matrix $A - cE_n$, uns zwar zum Eigenwert $\lambda - c$. Umgekehrt zeigt eine ähnliche Rechnung, dass jeder Eigenvektor der Matrix $A - cE_n$ auch ein Eigenvektor der Matrix A ist, und die Eigenwerte sich wieder um c unterscheiden. Bestimmt man daher den betragsgrößten Eigenwert von $A - cE_n$ durch das Vektoriterationsverfahren, so erhält man durch Addition von c den Eigenwert von A , der den größten Abstand von c hat.

Das zeigen wir wieder an einem Beispiel.

Beispiel. Wir arbeiten weiter mit der Matrix

$$A = \begin{pmatrix} 245 & -254 & -252 & -46 & -224 \\ 161 & -168 & -174 & -32 & -148 \\ -39 & 40 & 45 & 7 & 38 \\ 27 & -28 & -32 & -6 & -26 \\ 110 & -113 & -110 & -21 & -101 \end{pmatrix}.$$

Wir wollen nun einen weiteren Eigenwert dieser Matrix finden, also setzen wir in unserer Vorüberlegung $c = 13$ ein; der davon am weitesten entfernte Eigenwert wird nicht 13 sein, sofern überhaupt ein anderer Eigenwert existiert. Wir berechnen also

$$B = A - 13E_5 = \begin{pmatrix} 232 & -254 & -252 & -46 & -224 \\ 161 & -181 & -174 & -32 & -148 \\ -39 & 40 & 32 & 7 & 38 \\ 27 & -28 & -32 & -19 & -26 \\ 110 & -113 & -110 & -21 & -114 \end{pmatrix}.$$

Auch hier wieder merkt man, dass die Einträge der Vektoren sehr schnell wachsen:

$$x_1 = \begin{pmatrix} -544 \\ -374 \\ 78 \\ -78 \\ -248 \end{pmatrix}, x_2 = \begin{pmatrix} 8272 \\ 5738 \\ -1218 \\ 1218 \\ 3752 \end{pmatrix}, x_3 = \begin{pmatrix} -127888 \\ -89126 \\ 19038 \\ -19038 \\ -57800 \end{pmatrix}, x_4 = \begin{pmatrix} 1993360 \\ 1392842 \\ -297858 \\ 297858 \\ 898376 \end{pmatrix}.$$

Wir führen zunächst wieder die Skalierung mit dem betragsgrößten Eintrag mit geringer Genauigkeit durch:

$$z_1 = \begin{pmatrix} -1,0 \\ -0,69 \\ 0,14 \\ -0,14 \\ -0,46 \end{pmatrix}, z_2 = \begin{pmatrix} 1,0 \\ 0,69 \\ -0,15 \\ 0,15 \\ 0,45 \end{pmatrix}, z_3 = \begin{pmatrix} -1,0 \\ -0,70 \\ 0,15 \\ -0,15 \\ -0,45 \end{pmatrix}, z_4 = \begin{pmatrix} 1,0 \\ 0,70 \\ -0,15 \\ 0,15 \\ 0,45 \end{pmatrix}.$$

Hier merkt man allerdings, dass diese Werte noch nicht sehr nah an einem Eigenvektor sind. Führt man hier mehr Iterationen mit größerer Genauigkeit durch, so kommt man etwa in der 60-ten Iterationen zu

$$z_{60} = \begin{pmatrix} 1,0000000 \\ 0,71372610 \\ -0,14313695 \\ 0,14313695 \\ 0,42941084 \end{pmatrix},$$

und dieser Vektor ist zwar eine vernünftige, doch eine nicht allzu gute Approximation an einen Eigenvektor von A bzw. B . Tatsächlich wird hier ein Eigenvektor der Matrix A zum Eigenwert -3 (bzw. ein Eigenvektor von B zum Eigenwert -16) angenähert, doch beim Anwenden der Matrix A erhalten wir (im Rahmen unserer Rechengenauigkeit):

$$A \cdot z_{60} = \begin{pmatrix} -2,9882479 \\ -2,1333438 \\ 0,42745221 \\ -0,42745222 \\ -1,2823565 \end{pmatrix}.$$

In diesem Fall kann man feststellen, dass es viel länger dauert als im ersten Beispiel, um eine gute Näherung an den Eigenvektor zu bekommen. Das liegt daran, dass die Beträge der Eigenwerte der Matrix B „zu nah beieinander“ liegen. Man kann nämlich nachrechnen, dass die Matrix B gerade die Eigenwerte $0, -7, -12, -15, -16$ hat, und $|-16|$ zwar größer als $|-15|$ ist, doch die beiden Werte recht ähnlich sind. Hingegen hat die Matrix A Eigenwerte, die gerade um 13 verschoben sind, wie wir bereits erläutert haben, also $13, 6, 1, -2, -3$, und die beiden betragsgrößten Eigenwerte 13 und 6 liegen schon recht weit auseinander.

Bevor wir die Funktionsweise des Verfahrens untersuchen, geben wir nochmal eine (seltene) Ausnahme an, bei der die Vektoriteration nicht funktioniert.

Beispiel. Wir suchen die Eigenwerte der Matrix

$$C = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Wie wir bereits erläutert haben, gilt für die k -ten Vektor x_k in der Vektoriteration $x_k = C^k \cdot x_0$. Da es in diesem Fall informativer ist, berechnen wir also die Potenzen C^k der Matrix C für $2 \leq k \leq 8$.

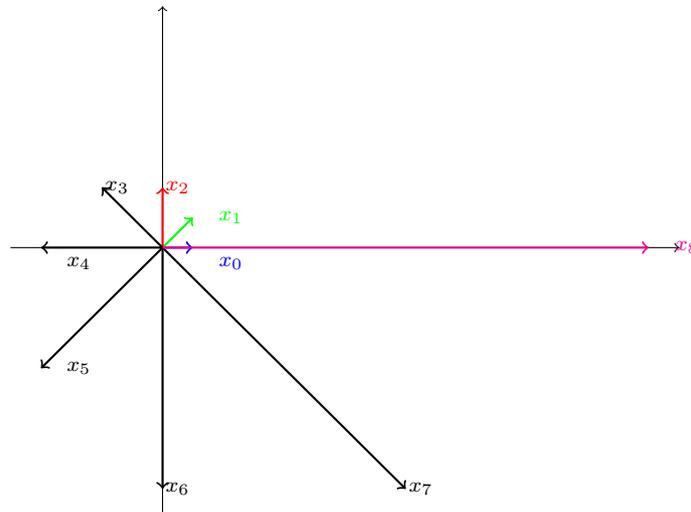
$$\begin{aligned} C^2 &= \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}, & C^3 &= \begin{pmatrix} -2 & -2 \\ 2 & -2 \end{pmatrix}, & C^4 &= \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix} \\ C^5 &= \begin{pmatrix} -4 & 4 \\ -4 & -4 \end{pmatrix}, & C^6 &= \begin{pmatrix} 0 & 8 \\ -8 & 0 \end{pmatrix}, & C^7 &= \begin{pmatrix} 8 & 8 \\ -8 & 8 \end{pmatrix} \\ C^8 &= \begin{pmatrix} 16 & 0 \\ 0 & 16 \end{pmatrix}. \end{aligned}$$

Egal, mit welchem Startvektor x_0 wir anfangen, wir bekommen im 8-ten Iterationsschritt also

$$x_8 = C^8 x_0 = \begin{pmatrix} 16 & 0 \\ 0 & 16 \end{pmatrix} x_0 = 16x_0.$$

Wenden wir das Verfahren weitere 8 mal an, so erhalten wir wieder ein Vielfaches von x_0 , und dann wieder nach weiteren 8 Schritten. Da das für jede Richtung des Vektors x_0 funktioniert, merkt man schon, dass, selbst wenn wir die Vektoren skalieren, die sich daraus ergebende Folge nicht konvergieren wird. Etwas genauer: Wir hatten uns in einem früheren Beispiel bereits überlegt, dass die Matrix C eine Drehstreckung in der Ebene beschreibt, bei der jeder Vektor mit dem Faktor $\sqrt{2}$ skaliert und um 45° gedreht wird.

Wir veranschaulichen die ersten 8 Iterationen in einem Bild.



Die Skalierung wird dann zwar die Längen, aber nicht die Richtungen der Vektoren ändern. Die Vektoren y_k werden also sich in jedem nächsten Schritt um einen Winkel von 45° unterscheiden. Dabei bemerken wir, dass sich die Richtungen alle 8 Schritte wiederholen, da man nach 8 Drehungen um 45° eine Volldrehung vollzogen hat. Man sieht also, dass hier die Folgen aus dem Vektoriterationsverfahren nicht konvergieren werden.

Das passt allerdings ausgezeichnet mit der früher ermittelten Tatsache zusammen, dass die Matrix C gar keine Eigenwerte besitzt.

Wir haben bereits gesehen, dass es allerdings auch Matrizen gibt, die zwar Eigenwerte besitzen, aber nicht diagonalisierbar sind. In diesen Fällen muss genauer geprüft werden, ob die Vektoriterationen anwendbar sind.

Nun wollen wir untersuchen, warum das Verfahren überhaupt funktioniert. Dazu geben wir die Beweisidee zum folgenden Satz.

Satz 16.1. *Sei A eine diagonalisierbare $n \times n$ -Matrix. Sei v_1, \dots, v_n eine Basis von \mathbb{R}^n aus Eigenvektoren von A zu Eigenwerten $\lambda_1, \dots, \lambda_n$. Diese seien betragsmäßig nicht-aufsteigend angeordnet, d.h. es gelte*

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|.$$

Zusätzlich nehmen wir an, dass $|\lambda_1|$ strikt größer als $|\lambda_2|$ ist. Sei x_0 ein Vektor in $\mathbb{R}^n \setminus \{0\}$, und sei

$$x_0 = \alpha_1 v_1 + \dots + \alpha_n v_n$$

die Darstellung von x_0 als Linearkombination der Basisvektoren v_1, \dots, v_n . Als letztes nehmen wir an, dass $\alpha_1 \neq 0$ gilt.

Dann konvergieren die Folgen $(y_k)_{k \in \mathbb{N}}$ und $(z_k)_{k \in \mathbb{N}}$ des Vektoriterationsverfahrens gegen ein Vielfaches des Eigenvektors v_1 der Matrix A zum betragsgrößten Eigenwert λ_1 .

Beweisidee. Wir beschäftigen uns nur mit dem Fall der Skalierung mit der Norm. Wir können durch Skalierung des Vektors v_1 zusätzlich erreichen, dass $\|v_1\| = 1$ ist, was wir von jetzt an annehmen werden.

Man wendet nun die Matrix A auf den Vektor x_0 an, um den Vektor x_1 zu erhalten. Wir können die Koordinaten des Vektors x_1 bezüglich der Basis v_1, \dots, v_n in Termen der Koordinaten $\alpha_1, \dots, \alpha_n$ des Vektors x_0 bezüglich derselben Basis ermitteln. Dabei nutzen wir die Linearität der Multiplikation mit einer Matrix aus, sowie die Tatsache, dass v_i Eigenvektoren von A zu Eigenwerten λ_i sind, und erhalten:

$$\begin{aligned} x_1 &= Ax_0 = A(\alpha_1 v_1 + \dots + \alpha_n v_n) \\ &= \alpha_1 A v_1 + \dots + \alpha_n A v_n \\ &= \alpha_1 \lambda_1 v_1 + \dots + \alpha_n \lambda_n v_n. \end{aligned}$$

Die Koordinate, die zum Vektor v_i gehört, wird also mit λ_i multipliziert. Mittels vollständiger Induktion kann man leicht beweisen, dass daraus eine ähnliche Formel für die k -te Iteration x_k erhalten werden kann:

$$x_k = \alpha_1 \lambda_1^k v_1 + \dots + \alpha_n \lambda_n^k v_n.$$

Nun verwenden wir etwas, was wir nicht beweisen werden und auch nicht gänzlich präzisieren werden. Man kann nämlich zeigen, dass für die Größenordnung der Norm von x_k nur die Koordinate von v_1 entscheidend ist, also dass

$$\|x_k\| \approx \alpha_1 \lambda_1^k$$

gilt. Hat man das und setzt $\alpha_1 \lambda_1^k$ für den Skalierungsfaktor, so erhalten wir

$$\begin{aligned} y_k = \frac{x_k}{\|x_k\|} &\approx \frac{\alpha_1 \lambda_1^k v_1 + \alpha_2 \lambda_2^k v_2 \dots + \alpha_n \lambda_n^k v_n}{\alpha_1 \lambda_1^k} \\ &= v_1 + \frac{\alpha_2}{\alpha_1} \left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \dots + \frac{\alpha_m}{\alpha_1} \left(\frac{\lambda_n}{\lambda_1}\right)^k v_n. \end{aligned}$$

Nun sieht man sofort, dass man die Voraussetzung $\alpha_1 \neq 0$ gebraucht hat, da es sonst nicht möglich wäre, durch diese Zahl zu teilen. (Dann würde auch $\|x_k\| \approx \alpha_1 \lambda_1^k$ nicht mehr korrekt sein.)

Nun benutzen wir die Tatsache, dass λ_1 der betragsgrößte Eigenwert ist, der vom Betrag her strikt größer als die anderen Eigenwerte ist. Daraus folgt, dass all die Brüche $\frac{\lambda_2}{\lambda_1}, \dots, \frac{\lambda_n}{\lambda_1}$ vom Betrag her strikt kleiner als 1 sind. Potenziert man eine Zahl, deren Betrag strikt kleiner als 1 ist, so wird der Betrag des Ergebnisses immer kleiner, und tatsächlich kann man sich überlegen, dass die Folge der Potenzen einer Zahl, deren Betrag strikt kleiner als 1 ist, gegen 0 konvergiert. (Das ist der Grund, wieso wir strikte Ungleichung in den Voraussetzungen gebraucht haben. Wäre einer dieser Quotienten genau 1, so würde die Folge der Potenzen davon 1 bleiben und das nachfolgende Argument nicht funktionieren.) Für große k sind die Vorfaktoren aller Basisvektoren - mit Ausnahme des Vorfaktors von v_1 , der stets 1 ist - in y_k immer näher an 0. Die Vektoren y_k sind also immer näher an v_1 ; etwas präziser gefasst, zeigt dieses Argument die Konvergenz der Folge (y_k) gegen v_1 . (Hier ist kein Vorfaktor nötig, da wir vorher v_1 skaliert haben; sonst müssten wir hier noch einen Faktor einfügen.) Das beendet die Beweisskizze für den obigen Satz. \square

Bislang haben wir gesehen, wie man mit der Vektoriterationsmethode den Eigenvektor zum betragsgrößten Eigenwert findet. Außerdem haben wir durch eine Verschiebung (engl. *shift*) eine Methode, um den am weitesten von dem vorgegebenen Wert $c \in \mathbb{R}$ entfernten Eigenwert einer Matrix bestimmen zu können. Allerdings möchte man häufig genau das Umgekehrte erreichen: Wir haben beispielsweise eine Vermutung über die Größenordnung

eines Eigenwerts, und möchten den Eigenwert der Matrix finden, der einer bestimmten Zahl $c \in \mathbb{R}$ am nächsten ist.

Wir wollen im Folgenden eine Modifikation der Vektoriteration, die inverse Vektoriteration, die es ermöglicht, den betragskleinsten Eigenwert einer Matrix zu finden. In Kombination mit der Verschiebung kann man (in den meisten Fällen) dann den Eigenwert der Matrix finden, der einer vorgegebenen Zahl $c \in \mathbb{R}$ am nächsten ist.

Dafür erinnern wir uns an die aus dem Übungsbetrieb bekannte Tatsache, dass für eine *invertierbare* Matrix A gilt: Eine reelle Zahl $\lambda \in \mathbb{R}$ ist genau dann ein Eigenwert von A , wenn $\frac{1}{\lambda}$ ein Eigenwert von der inversen Matrix A^{-1} zu A ist. Insbesondere kann eine invertierbare Matrix nicht den Eigenwert 0 haben. Für numerische Überlegungen ist die Einschränkung, dass die Matrix A invertierbar sein soll, nicht allzu groß. Hat man die Eigenwerte von A wieder der Größe des Betrags nach geordnet, also

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|,$$

so sind die Eigenwerte von A^{-1} gerade umgekehrt angeordnet, da der Kehrwert einer positiven Zahl umso größer ist, je kleiner die Zahl ist. Die Eigenwerte von A^{-1} erfüllen also:

$$\left| \frac{1}{\lambda_n} \right| \geq \left| \frac{1}{\lambda_{n-1}} \right| \geq \dots \geq \left| \frac{1}{\lambda_1} \right|,$$

und die Anwendung der Vektoriteration auf die Matrix A^{-1} liefert den Eigenwert $\frac{1}{\lambda_n}$, aus dem dann λ_n , der betragskleinste Eigenwert von A , bestimmt werden kann.

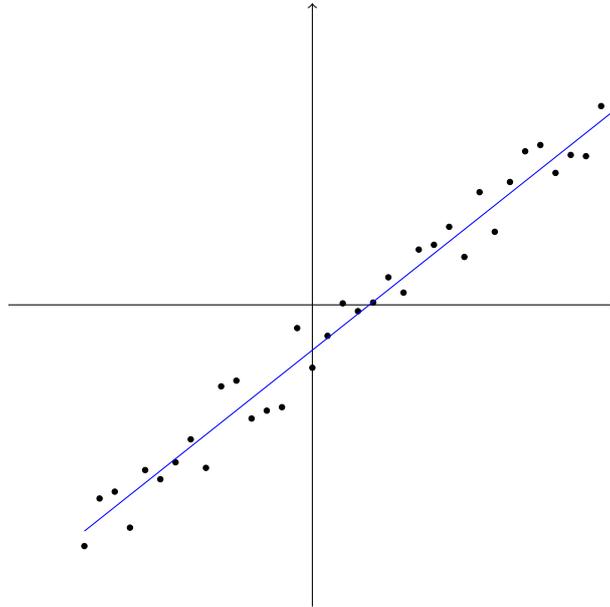
Wir wollen noch auf eine numerische Schwierigkeit der inversen Vektoriteration eingehen. Wie bereits erläutert, ist es numerisch nicht ratsam, die inverse Matrix zu einer Matrix zu berechnen. Allerdings kann man das bei der Vektoriteration umgehen. Anstatt die Matrix A^{-1} auszurechnen und diese in dem Iterationsschritt $x_{k+1} = A^{-1}x_k$ zu verwenden, bemerkt man, dass $x_{k+1} = A^{-1}x_k$ äquivalent ist zu $Ax_{k+1} = x_k$. Hat man also x_k vorgegeben und will x_{k+1} bestimmen, braucht man nur ein lineares Gleichungssystem zu lösen. Dafür gibt es numerisch sinnvolle Verfahren, beispielsweise die LR-Zerlegung und ihre Abwandlungen. Insbesondere muss man für jeden Iterationsschritt ein lineares Gleichungssystem mit derselben linken Seite lösen, was die Aufgabe stark erleichtert.

Damit haben wir also insgesamt auch ein Verfahren, um den Eigenwert einer Matrix A zu bestimmen, der am nächsten an einer vorgegebenen reellen Zahl c liegt.

17 Lineare Regression

Als letztes wollen wir ein weiteres numerisches Verfahren vorstellen, dessen Grundidee auf C.F. Gauß zurückgeht. Das grundlegende Problem dabei

besteht darin, zu einem Satz von Daten, die typischerweise experimentell bestimmt wurden, etwa aus einem physikalischen Experiment, eine bestmögliche Beschreibung durch eine affin-lineare Funktion zu finden. Eine solche Funktion bzw. ihr Graph wird dann auch „Ausgleichsgerade“ genannt. Das kann in etwa wie folgt aussehen:



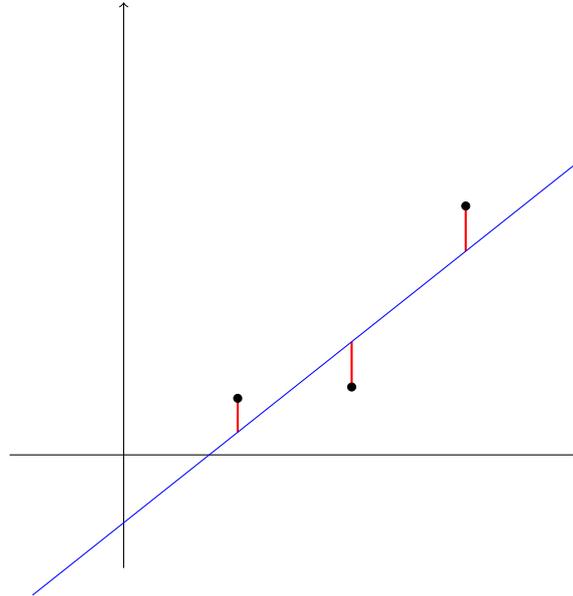
Eine Voraussetzung dafür, dass eine Ausgleichsgerade sinnvoll ist, ist ein affin-linearer Zusammenhang, der den Daten zugrundeliegt; in anderen Fällen ist eine Ausgleichsgerade nur bedingt sinnvoll.

Es stellt sich die Frage, wie gemessen werden sollte, wie gut eine Ausgleichsgerade an die Daten passt. Dafür gibt es unterschiedliche Möglichkeiten. Sicher ist es meist nicht der Fall, dass alle Datenpunkte auf einer Geraden bereits liegen. Man kann zwar versuchen, die Gerade möglichst durch zwei Datenpunkte zu legen (denn zwei Punkte bestimmen die Gerade bereits eindeutig), aber das hieße, dass man die anderen Messungen im Wesentlichen ignoriert. Auf C.F. Gauß geht ein Maß für die Abweichung der Ausgleichsgeraden von den Meßdaten zurück, das auch heute noch vielfach verwendet wird. Dabei handelt es sich um die *Methode der kleinsten Quadrate*.

Man hat also eine Gerade $y = mx + b$ gewählt und will feststellen, wie gut diese zu den Daten $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ passt. Dabei kann man die Vorschrift für die Gerade als eine Art Vorhersage ansehen: Man versucht zu schätzen, wie der Messwert zu dem Argument x sein würde. Bei der Methode der kleinsten Quadrate vergleicht man nun diese Vorhersage für die Argumente x_i , mit den tatsächlich gemessenen Messwerten. Dafür bestimmt man die Differenz $(mx_i + b) - y_i$ des Vorhersagewerts und des tatsächlichen Wertes.

Das sind fast die Längen der im Bild rot markierten Strecken, wobei die Zahlen $(mx_i + b) - y_i$ auch negativ sein können, wenn der geschätzte Wert

kleiner ist als der tatsächliche Wert.



Nun möchte man alle diese Abweichungen auf einmal berücksichtigen. Der erste Gedanke, der einem da vielleicht kommt, besteht darin, diese Abweichungen aufsummieren. Allerdings kann es passieren, da manche Abweichungen positiv und manche negativ sind, dass man als Summe 0 erhält, obwohl die einzelnen Abweichungen gar nicht so klein sind. Das ist natürlich nicht wünschenswert. Bei der Methode der kleinsten Quadrate summiert man nun nicht die Abweichungen auf, sondern deren Quadrate, und versucht diese Summe zu minimieren, also den Term

$$\sum_{i=1}^n ((mx_i+b)-y_i)^2 = ((mx_1+b)-y_1)^2 + ((mx_2+b)-y_2)^2 + \dots + ((mx_n+b)-y_n)^2.$$

(Das erklärt auch den Namen der Methode.)

Nun wollen wir untersuchen, wie wir die Steigung und den y -Achsenabschnitt der Geraden finden, die diesen Term minimieren.

Dafür erinnern wir uns an die Suche nach Minima und Maxima, die wir aus der Schule kennen. Dort lernt man, dass eine notwendige Bedingung für ein lokales Extremum einer differenzierbaren Funktion f an der Stelle x_0 durch $f'(x_0) = 0$ gegeben ist. Nun hängt unsere Funktion allerdings nicht, wie gewohnt, von einer, sondern von zwei Variablen ab. Für differenzierbare Funktionen von zwei Variablen ist das notwendige Kriterium für ein lokales Minimum allerdings ganz ähnlich: Es muss die Ableitung nach jeder Variablen 0 in dem betrachteten Punkt sein. Eine hinreichende Bedingung für ein Minimum anzugeben (also das Analogon für $f''(x_0) > 0$ bzw. für das Vorzeichenwechsel der ersten Ableitung) ist bereits schwieriger. Wir wollen hier das Kriterium nicht weiter angeben. Es sei erwähnt, dass die Nullstelle der Ableitung, die wir gleich erhalten, in den Situationen der Methode der kleinsten Quadrate stets ein Minimum liefert.

Wir bestimmen nun zunächst die Ableitung unserer Funktion nach der Variable m :

$$\sum_{i=1}^n 2 \cdot ((mx_i + b) - y_i) \cdot x_i,$$

dabei ist x_i jeweils die „innere Ableitung“ bei der Anwendung der Kettenregel. Wir erhalten also die erste Gleichung

$$\sum_{i=1}^n 2 \cdot ((mx_i + b) - y_i) \cdot x_i = 0.$$

Die Ableitung der Funktion nach der Variable b ist hingegen

$$\sum_{i=1}^n 2 \cdot ((mx_i + b) - y_i),$$

und wir erhalten die Gleichung

$$\sum_{i=1}^n 2 \cdot ((mx_i + b) - y_i) = 0.$$

Um das potentielle Minimum zu finden, müssen wir also das Gleichungssystem

$$\begin{cases} \sum_{i=1}^n 2 \cdot ((mx_i + b) - y_i) \cdot x_i & = 0, \\ \sum_{i=1}^n 2 \cdot ((mx_i + b) - y_i) & = 0. \end{cases}$$

Dieses Gleichungssystem formen wir nun etwas um. Dabei sortieren wir so, dass alle Summanden, die die Variablen m bzw. b enthalten, zusammen stehen, und bringen alle Summanden ohne m bzw. b auf die rechte Seite. Schließlich teilen wir noch beide Gleichungen durch 2.

$$\begin{aligned} & \begin{cases} \sum_{i=1}^n 2 \cdot ((mx_i + b) - y_i) \cdot x_i & = 0, \\ \sum_{i=1}^n 2 \cdot ((mx_i + b) - y_i) & = 0, \end{cases} \\ \Leftrightarrow & \begin{cases} \sum_{i=1}^n (mx_i^2 + bx_i - y_i x_i) & = 0, \\ \sum_{i=1}^n (mx_i + b - y_i) & = 0, \end{cases} \\ \Leftrightarrow & \begin{cases} \sum_{i=1}^n mx_i^2 + \sum_{i=1}^n bx_i - \sum_{i=1}^n y_i x_i & = 0, \\ \sum_{i=1}^n mx_i + \sum_{i=1}^n b - \sum_{i=1}^n y_i & = 0, \end{cases} \end{aligned}$$

Wir bemerken, dass die Summe $\sum_{i=1}^n b$ aus n Summanden, von denen jeder b ist, gerade $b \cdot n$ ist. Wir vertauschen außerdem die Reihenfolge der Gleichungen.

$$\begin{cases} b \cdot n + m \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n y_i, \\ b \cdot \sum_{i=1}^n x_i + m \cdot \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i x_i. \end{cases}$$

Nun sehen wir, dass dies in den Variablen b und m ein lineares inhomogenes Gleichungssystem ist. Die erweiterten Koeffizientenmatrix ist gegeben durch

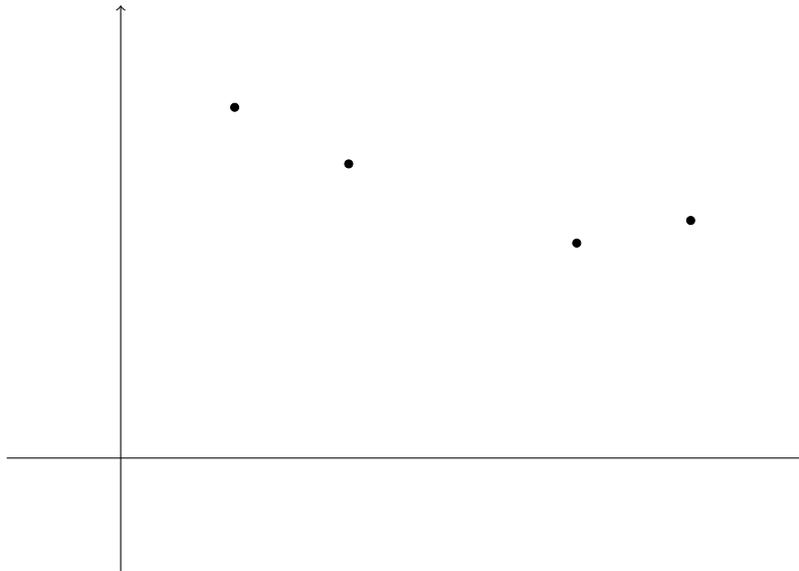
$$\left(\begin{array}{cc|c} n & \sum_{i=1}^n x_i & \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i y_i \end{array} \right)$$

Die Zahlen b und m , die sich als Lösungen dieses linearen Gleichungssystems ergeben, sind die Kandidaten für die Beschreibung der „besten“ Ausgleichsgeraden. Man kann nachrechnen, dass diese tatsächlich die Summe der quadratischen Abweichungen minimieren.

Wir geben an dieser Stelle noch ein Zahlenbeispiel zur Funktionsweise des Verfahrens.

Beispiel. Wir suchen eine Ausgleichsgerade für die Punkte $(1, 3, 1)$, $(2, 2, 6)$, $(4, 1, 9)$, $(5, 2, 1)$.

Die Punkte liegen in etwa wie folgt:



Hier ahnt man schon, dass die gesuchte Gerade eine betragsmäßig relativ kleine, negative Steigung haben muss. Wir setzen die Werte in das soeben

aufgestellte Gleichungssystem ein.

$$\begin{aligned}
 n &= 4, \\
 \sum_{i=0}^n x_i &= 1 + 2 + 4 + 5 &= 12, \\
 \sum_{i=0}^n x_i^2 &= 1^2 + 2^2 + 4^2 + 5^2 &= 46, \\
 \sum_{i=0}^n y_i &= 3,1 + 2,6 + 1,9 + 2,1 &= 9,7, \\
 \sum_{i=0}^n x_i y_i &= 1 \cdot 3,1 + 2 \cdot 2,6 + 4 \cdot 1,9 + 5 \cdot 2,1 &= 26,4.
 \end{aligned}$$

Wir erhalten also ein lineares Gleichungssystem mit der erweiterten Koeffizientenmatrix

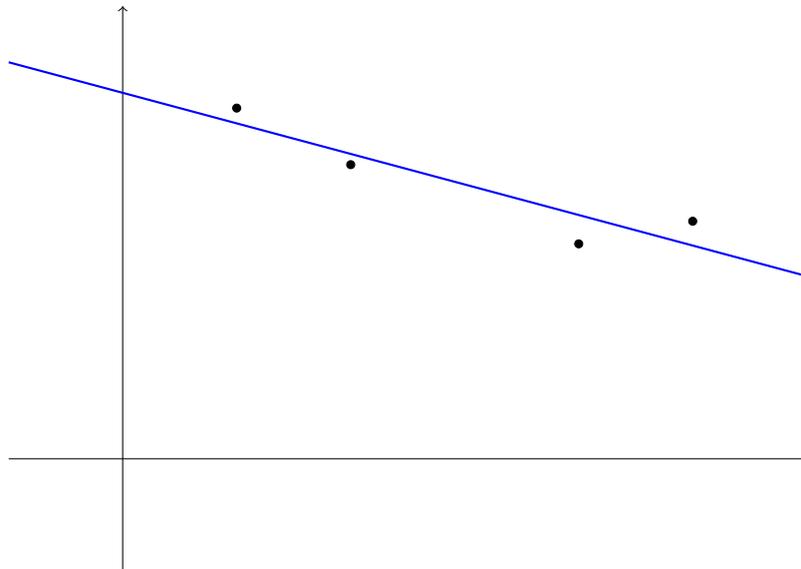
$$\left(\begin{array}{cc|c} 4 & 12 & 9,7 \\ 12 & 46 & 26,4 \end{array} \right),$$

das wir nun lösen müssen.

Das tun wir, wie üblich, mit dem Gaußschen Eliminationsverfahren:

$$\begin{aligned}
 &\left(\begin{array}{cc|c} 4 & 12 & 9,7 \\ 12 & 46 & 26,4 \end{array} \right) \xrightarrow{II-3I} \left(\begin{array}{cc|c} 4 & 12 & 9,7 \\ 0 & 10 & -2,7 \end{array} \right) \\
 &\xrightarrow{II/10, I/4} \left(\begin{array}{cc|c} 1 & 3 & 2,425 \\ 0 & 1 & -0,27 \end{array} \right) \xrightarrow{I-3II} \left(\begin{array}{cc|c} 1 & 0 & 3,235 \\ 0 & 1 & -0,27 \end{array} \right)
 \end{aligned}$$

Also ist die optimale Ausgleichsgerade gegeben durch $m = -0,27$, $b = 3,235$, d.h. durch $y = -0,27x + 3,235$. Das sieht im Bild etwa so aus:



Zuletzt gehen wir noch auf eine andere Art ein, auf das Gleichungssystem für die Parameter der Ausgleichsgeraden zu kommen. Diese Sichtweise, die mehr an die lineare Algebra angelehnt ist als unser erster Zugang mit den Ableitungen, hat den Vorteil, dass sie auf kompliziertere Gebilde, beispielsweise Ausgleichsparabeln, leichter ausgeweitet werden kann. Dabei bemerken wir, dass wir am liebsten m und b als Lösungen des linearen Gleichungssystems

$$\begin{cases} b + mx_1 = y_1 \\ b + mx_2 = y_2 \\ \vdots \\ b + mx_n = y_n \end{cases}$$

erhalten würden, das durch die erweiterte Koeffizientenmatrix

$$\left(\begin{array}{cc|c} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ & \vdots & \\ 1 & x_n & y_n \end{array} \right)$$

gegeben ist. Wir nennen die linke Seite der erweiterten Koeffizientenmatrix abkürzend A und die rechte y . Dieses Gleichungssystem hat natürlich keine Lösung, da die Datenpunkte nicht wirklich auf einer Geraden liegen. Allerdings können wir die Menge aller Vektoren

$$\{Av - y \mid v \in \mathbb{R}^n\}$$

in \mathbb{R}^n betrachten. Da das lineare Gleichungssystem keine Lösung hat, ist 0 nicht in dieser Menge, also kann sie insbesondere kein Unterraum von \mathbb{R}^n sein. Allerdings ist das ein „verschobener“ Unterraum von \mathbb{R}^n , etwa wie eine Gerade oder Ebene in \mathbb{R}^2 oder \mathbb{R}^3 , die nicht durch den Ursprung geht. Um die besten Parameter zu finden, müssen wir das v finden, für das $Av - y$ am nächsten am Ursprung dran ist. Dafür fallen wir das Lot vom Ursprung auf die „verschobene“ Ebene, die wir betrachten. Das lässt sich mit Hilfe des Skalarproduktes gut in Termen von linearer Algebra umsetzen, und dafür hat man auch effektive numerische Verfahren. Daher ist diese Sichtweise auf die lineare Regression sehr hilfreich.

Literatur

- [1] Peter Hartmann. *Mathematik für Informatiker. Ein praxisbezogenes Lehrbuch. 3., überarbeitete und erweiterte Auflage.* Wiesbaden: Vieweg, 3rd revised and expanded edition edition, 2004.