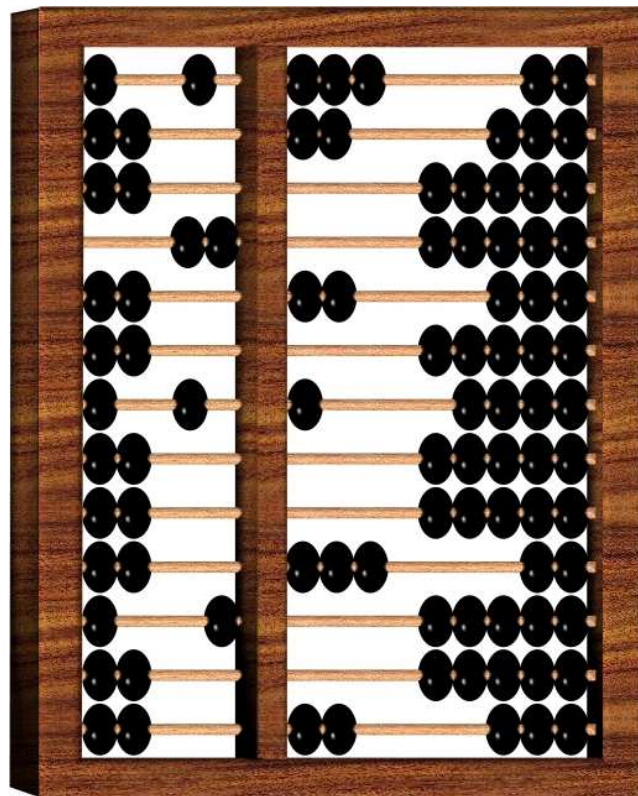


Christian Bär

Elemente der linearen Algebra und analytischen Geometrie

Wintersemester 2008/09



Version vom 28. Juli 2011

Inhaltsverzeichnis

Vorwort	iii
1 Grundlagen	1
1.1 Vorbemerkungen zur Logik	1
1.2 Mengen und Abbildungen	7
2 Matrixrechnung	21
2.1 Lineare Gleichungssysteme	21
2.2 Lineare Unabhängigkeit und Basen	34
2.3 Der Gauß-Algorithmus	46
2.4 Geometrie der Ebene	56
2.5 Die komplexen Zahlen	78
2.6 Graphentheorie - etwas Kombinatorik	102
2.7 Determinanten	109
2.8 Affine Abbildungen und affine Unterräume	133
2.9 Volumina	136
2.10 Eigenwerte und Eigenvektoren	146
2.11 Orientierungen und Vektorprodukt	160
3 Vektorräume	167
3.1 Gruppen	167
3.2 Körper	172
3.3 Vektorräume	174
Abbildungsnachweise	177
Index	179

Vorwort

Dies ist das Vorlesungsskript zur Vorlesung „Elemente der linearen Algebra und analytischen Geometrie“, die ich im Wintersemester 2008/09 an der Universität Potsdam gehalten habe. Die Hörer waren Studierende im ersten Semester des Lehramts für die Primar- und die Sekundarstufe I. Aus diesem Grund und weil die Vorlesung nur einsemestrig angelegt war, habe ich bei der Stoffwahl konkrete Themen, wie lineare Gleichungssysteme oder die Geometrie der Ebene, zunächst in den Vordergrund gestellt. Die abstrakten Begriffsbildungen wie Gruppen, Körper, Vektorräume kamen erst später. Dies hat auch den Vorteil, dass man dann bereits über mehr Beispiele, z.B. Matrixgruppen verfügt.

Da die Studierenden des ersten Semesters zunächst an korrektes mathematisches Formulieren und logisches Argumentieren herangeführt werden müssen, wurde dem eigentlichen Stoff der Vorlesung ein Kapitel über elementarste Aussagenlogik und Mengenlehre vorangestellt.

Mein herzlicher Dank geht an Martin Naumann, der die Erstversion dieses Skriptes in hervorragender Qualität verfasst hat.

Potsdam, Januar 2011,

Christian Bär

1 Grundlagen

Was sich überhaupt sagen lässt, lässt sich klar sagen; und wovon man nicht reden kann, darüber muss man schweigen.

*(Ludwig Wittgenstein
Tractatus logico-philosophicus)*

Bevor wir uns mit linearer Algebra oder analytischer Geometrie befassen, behandeln wir zunächst einige grundsätzliche Fragen, die korrektes Formulieren und Argumentieren in der Mathematik und darüber hinaus betreffen.

1.1 Vorbemerkungen zur Logik

In der Mathematik machen wir Aussagen, die entweder wahr oder falsch sind. So ist etwa die Aussage „ 8 ist eine gerade Zahl“ wahr wohingegen „Jede Quadratzahl ist gerade“ falsch ist. Durch Aussagenverbindungen können aus Aussagen A und B neue Aussagen gebildet werden:

$A \& B$	A und B
A oder B	A oder B (kein ausschließendes oder)
$A \Rightarrow B$	A impliziert B
$A \Leftrightarrow B$	A und B sind äquivalent, d.h. $A \Rightarrow B$ und $B \Rightarrow A$

Ein Doppelpunkt vor oder hinter dem Äquivalenzpfeil drückt aus, dass eine Aussage durch eine andere definiert wird, z.B. $A : \Leftrightarrow$ „ 8 ist eine gerade Zahl“. Ferner können Aussagen verneint werden. Die Verneinung einer Aussage A ist wahr, wenn A selbst falsch ist, und umgekehrt. Ist z.B. $A : \Leftrightarrow$ „ 8 ist eine gerade Zahl“, so ist ihre Verneinung $\neg A \Leftrightarrow$ „ 8 ist keine gerade Zahl“. Für die Aussage $B : \Leftrightarrow$ „Jede Quadratzahl ist gerade“ haben wir $\neg B \Leftrightarrow$ „Nicht jede Quadratzahl ist gerade“. Für letzteres könnten wir auch $\neg B \Leftrightarrow$ „Es gibt wenigstens eine ungerade Quadratzahl“ schreiben.

Häufig ist es nützlich, eine ganze Serie von Aussagen in strukturierter Form zusammenzufassen. So könnten wir z.B. die Aussage

$$A(n) : \Leftrightarrow \text{„}n \text{ ist eine gerade Zahl“}$$

1 Grundlagen

machen, wobei n irgendeine positive ganze Zahl sein darf. Offensichtlich ist dann $A(1)$ falsch, $A(2)$ wahr, $A(3)$ falsch usw.

Machen wir doch mal die Aussage $A(n)$:

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

Für welche positiven ganzen Zahlen n ist diese Aussage wahr? Tatsächlich für alle. Wie kann man das einsehen?

Hier kommt das wichtige Beweisprinzip der *vollständigen Induktion* zum Einsatz. Wir gehen dabei folgendermaßen vor: Zunächst überzeugen wir uns, dass $A(1)$ richtig ist (der Induktionsanfang) und dann prüfen wir nach, dass für jedes n die Aussage $A(n)$ die Aussage $A(n+1)$ impliziert (der Induktionsschritt). Wenn wir das erledigt haben, sind wir fertig, denn dann wissen wir, dass $A(n)$ für jedes n richtig ist. Der Induktionsschritt mit $n = 1$ sagt uns, dass $A(2)$ aus $A(1)$ folgt, also ist auch $A(2)$ richtig. Der Induktionsschritt mit $n = 2$ sagt uns, dass $A(3)$ aus $A(2)$ folgt, also ist auch $A(3)$ richtig usw.

Führen wir dies im obigen Beispiel der Aussage $A(n) \Leftrightarrow$

$$1 + 2 + \dots + n = \frac{n(n+1)}{2}$$

durch.

Induktionsanfang: Wir müssen $A(1)$ als wahr erkennen. In der Tat,

$$1 = \frac{1(1+1)}{2}$$

ist wahr.

Induktionsschritt: Nun müssen wir zeigen, dass für beliebiges n die Aussage $A(n)$ (die Induktionsvoraussetzung) die Aussage $A(n+1)$ (die Induktionsbehauptung) impliziert, in Symbolen $A(n) \Rightarrow A(n+1)$:

$$\begin{aligned} 1 + 2 + \dots + n &= \frac{n(n+1)}{2} \\ &\Downarrow \\ 1 + 2 + \dots + n + (n+1) &= \frac{n(n+1)}{2} + (n+1) \\ &\Downarrow \\ 1 + 2 + \dots + n + (n+1) &= \frac{n(n+1) + 2(n+1)}{2} \\ &\Downarrow \\ 1 + 2 + \dots + n + (n+1) &= \frac{(n+1)(n+2)}{2} \\ &\Downarrow \\ 1 + 2 + \dots + n + (n+1) &= \frac{(n+1)((n+1)+1)}{2} \end{aligned}$$

Die erste Zeile ist die Aussage $A(n)$, die letzte ist $A(n+1)$. Damit ist die vollständige Induktion vollzogen und somit die Aussage $A(n)$ für alle n bewiesen.

Ist die Induktionsbehauptung eine Gleichung, so beginnt man im Induktionsschritt häufig mit einer Seite der Induktionsbehauptung und führt diese mit Hilfe der Induktionsvoraussetzung (IV) zum Ergebnis, z.B.:

$$\begin{aligned} 1 + 2 + \cdots + n + (n + 1) &= (1 + 2 + \cdots + n) + (n + 1) \\ &\stackrel{\text{IV}}{=} \frac{n(n + 1)}{2} + (n + 1) \\ &= \frac{(n + 1)((n + 1) + 1)}{2} \end{aligned}$$

In jedem Fall muss die Induktionsvoraussetzung benutzt werden, sonst ist es keine Induktion!

Nun ist nicht jede Aussage von positiven ganzen Zahlen abhängig und damit mit vollständiger Induktion nachprüfbar. Auch erweist sich in einigen Fällen ein direkter Beweis als umständlich oder gar unmöglich. So könnte man z.B. für die Aussage, dass unendlich viele Primzahlen existieren, versuchen sie dadurch zu beweisen, dass man unendlich viele Primzahlen angibt. Das ist allerdings nicht so einfach. Daher verwendet man hier ein Beweisprinzip, dem die Beobachtung zu Grunde liegt, dass $A \Rightarrow B$ logisch äquivalent zu $\neg B \Rightarrow \neg A$ ist. Gilt B nicht, so kann auch A nicht gelten. Mit dem darauf beruhenden indirekten Beweisprinzip (Prinzip des Widerspruchsbeweises), werden wir sehen, dass es tatsächlich keine größte Primzahl gibt, d.h. unendlich viele Primzahlen existieren. Dazu nehmen wir genau die Negierung der Behauptung an und folgern daraus eine falsche Aussage, einen so genannten Widerspruch, den wir stets abkürzend mit ζ kennzeichnen.

Für unser Primzahl-Beispiel gehen wir von folgender Festlegung aus.

Definition 1.1.1. Eine ganze Zahl heißt **Primzahl**, falls sie größer als 1 und nur durch sich selbst und 1 teilbar ist.

Beispiele für Primzahlen sind 2, 3, 5, 7, etc. Dagegen ist 9 keine Primzahl, da $9 = 3 \cdot 3$.

Satz 1.1.2

Es gibt unendlich viele verschiedene Primzahlen.

Beweis.

Wir gehen von der Widerspruchsannahme aus, dass es nur endlich viele Primzahlen gäbe. Seien

1 Grundlagen

also

$$p_1 = 2, p_2 = 3, p_3 = 5, \dots, p_N$$

diese Primzahlen. Nun setzen wir

$$k := p_1 \cdot p_2 \cdot \dots \cdot p_N + 1.$$

Bei Division der Zahl k durch irgendeine der Primzahlen bleibt stets der Rest 1. Daher ist die positive ganze Zahl k durch keine der Primzahlen teilbar. Andererseits kann man jedoch jede positive ganze Zahl in eindeutiger Weise als Produkt von Primzahlen darstellen und damit insbesondere durch bestimmte Primzahlen teilen. An dieser Stelle schlägt der ζ ein, d.h. die Annahme nur endlich vieler Primzahlen war falsch. \square

Bemerkung 1.1.3. Dieser Widerspruchsbeweis wird Euklid zugeschrieben und ist somit über 2000 Jahre alt. Euklid lebte ca. von 365 bis 300 v. Chr. und war einer der bedeutendsten Mathematiker der Antike. Sein berühmtestes Werk, die vielbändigen *Elemente*, fasst den damaligen Kenntnisstand in Geometrie und Arithmetik zusammen. Wir zitieren die Wikipedia: *Über Euklid erzählt man sich viele Anekdoten: Ein Schüler fragte, als er den ersten Satz gelernt hatte: „Was kann ich verdienen, wenn ich diese Dinge lerne?“ Da rief Euklid seinen Sklaven und sagte: „Gib ihm drei Obolen, denn der arme Mann muss Geld verdienen mit dem, was er lernt.“*



Abb. 1

Wir sind nun vertraut mit den Prinzipien des direkten Beweises, des Induktionsbeweises sowie des Widerspruchsbeweises. Hier eine wichtige **Warnung**: Es mag für den ein oder anderen überraschend erscheinen, aber falsche Aussagen können wahre implizieren! So folgt etwa aus der falschen Aussage $-1 = 1$ die Aussage $(-1)^2 = 1^2$ und daraus wiederum $1 = 1$, eine wahre Aussage. Man kann also Aussagen nicht dadurch beweisen, dass man wahre Aussagen aus ihnen herleitet¹. Eine derart fehlerhafte Argumentation kann man leider an Schulen, von manchem Lehrer abgesehen, häufig antreffen.

Wir wollen diesen Abschnitt mit einigen Bemerkungen zu den Aussagenverbindungen kompletieren. Seien dazu A und B Aussagen und bezeichnen W und F die Wahrheitswerte wahr und falsch.

Dann besitzen die Aussagen $A \& B$ und A oder B folgende Wertetafeln:

A	B	$A \& B$	A	B	A oder B
F	F	F	F	F	F
F	W	F	F	W	W
W	F	F	W	F	W
W	W	W	W	W	W

¹Ein Widerspruchsbeweis tut etwas anderes: Er beweist eine Aussage, indem er aus ihrer Negierung eine falsche Aussage herleitet. Das ist ok.

Für die Aussagen $A :\Leftrightarrow „8 \text{ ist gerade}“$ (wahr) und $B :\Leftrightarrow „3 \text{ ist gerade}“$ (falsch) ergibt sich zum Beispiel $A \& B \Leftrightarrow „8 \text{ und } 3 \text{ sind gerade}“$, d.h. eine falsche Aussage. Andererseits gilt $A \text{ oder } B \Leftrightarrow „8 \text{ oder } 3 \text{ ist gerade}“$. Auch die Aussage „2 oder 4 ist gerade“ ist wahr. Dieses Beispiel zeigt, dass es sich bei dem mathematischen *oder* nicht um ein exklusives oder handelt. In anderen Worten: die Aussage $A \text{ oder } B$ ist wahr, falls nur eine der beiden Teilaussagen oder auch beide wahr sind.

Die Negation von $A \& B$ ist äquivalent zu $\neg A \text{ oder } \neg B$, da für die Aussage $\neg(A \& B)$ bereits genügt, dass nur eine der Teilaussagen nicht gilt. Umgekehrt ist $\neg(A \text{ oder } B)$ äquivalent zu $\neg A \& \neg B$.

Zuletzt hatten wir die die Aussage $A(n) \Leftrightarrow 1 + \dots + n = \frac{n(n+1)}{2}$ für alle positiven ganzen Zahlen als richtig erkannt. Für derartige Quantifikationen stellt der mathematische Zeichenapparat die ausdrucksstarken Symbole \forall (**All-Quantor**) und \exists (**Existenz-Quantor**) bereit. Für eine Serie von Aussagen $A(n)$ ist die neue Aussage

$$\forall n : A(n)$$

genau dann wahr, wenn $A(n)$ für alle n wahr ist. So ist z.B. die Aussage $\forall n$ positiv und ganz: $1 + \dots + n = \frac{n(n+1)}{2}$ wahr. Die Aussage

$$\exists n : A(n)$$

dagegen ist genau dann wahr, wenn die Aussage $A(n)$ für wenigstens ein n wahr ist. So ist z.B. für $B(n) :\Leftrightarrow „n^2 \text{ ist gerade}“$ die Aussage „ $\forall n$ ganz und positiv: $B(n)$ “ falsch, weil nicht alle Quadratzahlen gerade sind (z.B. $1^2 = 1$ nicht), aber die Aussage „ $\forall n$ ganz und positiv: $B(n)$ “ ist wahr, weil es gerade Quadratzahlen gibt (z.B. $2^2 = 4$).

Wie negiert man nun quantorenbehaftete Aussagen?

Dass eine Serie von Aussagen $A(n)$ nicht für alle n wahr ist, bedeutet dasselbe wie, dass sie wenigstens für ein n falsch ist, d.h. $\neg(\forall n : A(n)) \Leftrightarrow \exists n : \neg A(n)$. Ähnlich gilt $\neg(\exists n : A(n)) \Leftrightarrow \forall n : \neg A(n)$.

Bemerkung 1.1.4. Anders als im umgangssprachlichen Gebrauch üblich, bedeutet in der Mathematik „es gibt ein...“ (was durch den Quantor \exists zum Ausdruck gebracht wird) stets „es gibt *mindestens* ein...“. Möchte man dagegen ausdrücken, dass es ein und nur ein (also genau ein) n gibt, für das $A(n)$ wahr ist, so notiert man $\exists! n : A(n)$.

Beispiel 1.1.5. Wir verdeutlichen den Einsatz von \exists und $\exists!$ an unterschiedlichen Quantifizierungen der Aussageform $n^2 = 25$:

$$\begin{array}{ll} \exists! n \text{ positiv und ganz: } n^2 = 25 & \text{(wahr)} \\ \exists n \text{ positiv und ganz: } n^2 = 25 & \text{(wahr)} \\ \exists! n \text{ ganz: } n^2 = 25 & \text{(falsch)} \\ \exists n \text{ ganz: } n^2 = 25 & \text{(wahr)} \end{array}$$

1 Grundlagen

Selbstverständlich begegnen uns Aussagenverbindungen und Quantifizierungen auch in Textform. Zum Beispiel lässt sich die Aussage „Alle viereckigen Kreise sind blau.“ wie folgt kodieren:

$$\forall \text{ Kreise } K \text{ viereckig} : K \text{ ist blau.} \quad (1.1)$$

Gleichwohl diese Aussage vollständig quantifiziert ist, erscheint eine Entscheidung für wahr oder falsch auf den ersten Blick schwierig. Gehen wir einmal zur Negierung der Aussage über:

$$\exists \text{ Kreise } K \text{ viereckig} : \neg(K \text{ ist blau}),$$

also umgangssprachlich „Es gibt viereckige Kreise, die nicht blau sind.“ Diese Aussage ist falsch, da es überhaupt keine viereckigen Kreise gibt. Somit ist deren Negation, also die ursprüngliche Aussage (1.1), wahr.

Oft treten Quantifizierungen geschachtelt auf. Betrachten wir für $n > m$ einmal alle Möglichkeiten der Quantifizierung, also eine Verschachtelung mit zwei Quantoren:

$$\begin{aligned} \forall n \text{ ganz} : \forall m \text{ ganz} : n > m & \quad (\text{falsch}) \\ \exists n \text{ ganz} : \forall m \text{ ganz} : n > m & \quad (\text{falsch}) \\ \forall n \text{ ganz} : \exists m \text{ ganz} : n > m & \quad (\text{wahr}) \\ \exists n \text{ ganz} : \exists m \text{ ganz} : n > m & \quad (\text{wahr}) \end{aligned}$$

Besonders in Beweisen ergibt sich häufig die Notwendigkeit, auch verschachtelte Aussagen zu negieren:

$$\begin{aligned} \neg(\forall n : \forall m : A(n, m)) & \Leftrightarrow \exists n : \neg(\forall m : A(n, m)) \\ & \Leftrightarrow \exists n : \exists m : \neg A(n, m). \end{aligned}$$

Wir bemerken: Die Negation verschachtelter Quantifikationen erfolgt durch das Vertauschen aller auftretenden Quantoren und die Verneinung der Aussageform. Wir betrachten als Beispiel die Aussage: „In so manchem Land haben alle Städte weniger als 1 Million Einwohner“ mit dem Ziel, das eben erarbeitete Kalkül anzuwenden und schreiben:

$$\exists \text{ Land } L : \forall \text{ Stadt } S \text{ in } L : \#(\text{Einwohner in } S) < 10^6.$$

Die Negation kann nun quasi algorithmisch erfolgen:

$$\forall \text{ Land } L : \exists \text{ Stadt } S \text{ in } L : \#(\text{Einwohner in } S) \geq 10^6.$$

Dies bedeutet aber wiederum umgangssprachlich: „In jedem Land gibt es (mindestens) eine Stadt mit mindestens 1 Million Einwohnern“.

1.2 Mengen und Abbildungen

Nach dem Begründer der Mengenlehre Georg Cantor (1845-1918) ist eine **Menge** eine Zusammenfassung bestimmter, wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens zu einem Ganzen.

Genau genommen ist diese Definition nicht wirklich präzise, da keinerlei Vereinbarung oder Definition der *wohlunterschiedenen Objekte* vorliegt. Dieser Umstand soll uns jedoch nicht beirren, da wir die Mengennotation vor allem als Hilfsmittel sehen, die Objekte unseres Interesses prägnant auszudrücken.

Beispiel 1.2.1. Folgende Mengen stellen die klassischen Zahlbereiche der Schule dar:

$\mathbb{N} = \{1, 2, 3, \dots\}$	natürlichen Zahlen
$\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$	natürlichen Zahlen inkl. der Null
$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$	ganze Zahlen
$\mathbb{Q}_+ = \left\{ \frac{p}{q} \mid p, q \in \mathbb{N} \right\}$	gebrochene Zahlen
$\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z} \ \& \ q \neq 0 \right\}$	rationale Zahlen
$\mathbb{R} = \{\text{reelle Zahlen}\}$	reelle Zahlen

Wir wollen unsere Gedanken nun auch auf andere Mengen und Mengenkonstruktionen richten. Eine wichtige Eigenschaft einer Menge A ist die Anzahl ihrer Elemente $\#A$, etwa $\#\{1, 2, 3\} = 3$, $\#\mathbb{N} = \infty$ und $\#\{\text{Primzahlen}\} = \infty$. Die wohl einfachste Menge ist die leere Menge $\{\} = \emptyset$. Für sie gilt $\#\emptyset = 0$.

Mengen können wiederum Elemente anderer Mengen sein. So ist z.B. $\{\emptyset\} \neq \emptyset$, denn $\{\emptyset\}$ hat genau ein Element, nämlich \emptyset . Es ist also $\#\{\emptyset\} = 1$. Ferner ist $\#\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} = 3$ mit den Elementen \emptyset , $\{\emptyset\}$ und $\{\emptyset, \{\emptyset\}\}$, unabhängig davon, dass $\{\emptyset, \{\emptyset\}\}$ hierbei selbst wieder eine zweielementige Menge ist.

Die Elemente einer Menge haben keine Reihenfolge, etwa ist $\{2, 3\} = \{3, 2\}$. Zudem tritt jedes Element einer Menge nur einmal auf, z.B. ist $\{1, 1\} = \{1\}$ und $\{2, \frac{4}{2}, \frac{8}{4}\} = \{2\}$.

Definition 1.2.2. Sind A und B Mengen, so heißt A **Teilmenge** von B , in Symbolen

$$A \subset B :\Leftrightarrow \forall a \in A : a \in B.$$

Analog ist A **Obermenge** von B , symbolisch $A \supset B :\Leftrightarrow B \subset A$.

Ferner gilt die Mengengleichheit $A = B \Leftrightarrow A \subset B \ \& \ B \subset A$.

Für $A = \{\diamond, \heartsuit, \spadesuit, \clubsuit\}$ ist $\diamond \in A$ und damit $\{\diamond\} \subset A$. Ebenso kann für $\heartsuit, \clubsuit \in A$ auch $\{\heartsuit, \clubsuit\} \subset A$ notiert werden.

1 Grundlagen

Bemerkung 1.2.3. Aus der Definition von *Teilmenge* folgt sofort, dass die leere Menge Teilmenge jeder Menge ist.

Definition 1.2.4. Sei A eine Menge. Die **Potenzmenge**

$$\mathcal{P}(A) := \{T \mid T \subset A\}$$

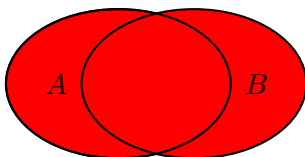
ist die Menge aller Teilmengen von A . In anderen Worten: $T \subset A \Leftrightarrow T \in \mathcal{P}(A)$.

Nehmen wir einmal $A = \{\heartsuit, \clubsuit\}$ her. Dann ist $\mathcal{P}(A) = \{\emptyset, \{\heartsuit\}, \{\clubsuit\}, A\}$. Ein anderes Beispiel wäre $A = \{1, 2, 3\}$ und $\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, A\}$.

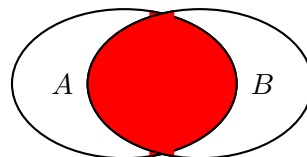
Bemerkung 1.2.5. Für endliche Mengen A gilt stets $\#\mathcal{P}(A) = 2^n$, wobei $n = \#A$. Dies kann man durch vollständige Induktion beweisen, was wir dem engagierten Leser zur Übung überlassen.

Wir haben die Verbindungen $\&$, oder, \Rightarrow und \Leftrightarrow kennen gelernt, um Aussagen logisch zu verbinden. Zudem sind aus der Schule die Rechenoperationen $+$, $-$, \cdot und $:$ für Zahlen bekannt. Auch für Mengen A und B existieren Operationen, die Mengenoperationen:

Vereinigung $A \cup B = \{x \mid x \in A \text{ oder } x \in B\}$



Durchschnitt $A \cap B = \{x \mid x \in A \& x \in B\}$



Differenz $A \setminus B = \{x \mid x \in A \& x \notin B\}$

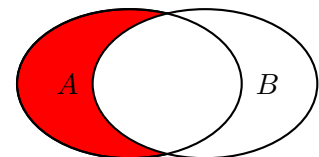


Abb. 2

Eine vierte wichtige Mengenoperation ist das kartesische Produkt deren Aufbau wir uns zunächst einmal an einem Beispiel klar machen: Bei der Analyse von Lehrer-Schüler-Interaktion im Unterricht unterscheidet der Pädagoge Flanders nach

Sozialformen			Grad der vorgesehenen Schülerinitiative	
s_1	Unterricht im Klassenverband	und	g_1	Instruiert werden
s_2	Teilgruppenunterricht		g_2	Zur Entdeckung gelenkt werden
s_3	Einzelunterricht		g_3	Impulse erhalten

Aus den so erzeugten Mengen $S = \{s_1, s_2, s_3\}$ und $G = \{g_1, g_2, g_3\}$ bildet Flanders geordnete Paare (s_i, g_i) , geordnet, da stets im ersten Eingang die Sozialform und als zweite Komponente der Grad der Schülerinitiative notiert werden soll. Damit ergeben sich z.B. die Interpretationen $(s_1, g_1) \hat{=}$ Frontalunterricht, $(s_1, g_2) \hat{=}$ Fragend-entwickelnder Unterricht oder $(s_3, g_1) \hat{=}$ Individualisierte Instruktion.

Das hier angewendete Konzept der Bildung geordneter n -Tupel (a_1, \dots, a_n) lässt sich für endlich viele Mengen A_1, \dots, A_n wie folgt mathematisch fassen:

$$A_1 \times \dots \times A_n := \{(a_1, \dots, a_n) \mid a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n\}$$

heißt **kartesisches Produkt** der Mengen A_1, \dots, A_n . Wie oben beschrieben, bedeutet geordnet dann z.B. $(1, 0) \neq (0, 1)$, d.h. es kommt auf die Reihenfolge der Eingänge an. Betrachten wir z.B. $A_1 = \{0, 1\}$ und $A_2 = \{1, 2, 3\}$. Dann ist $A_1 \times A_2 = \{(0, 1), (0, 2), (0, 3), (1, 1), (1, 2), (1, 3)\}$.

Im Spezialfall $A_1 = \dots = A_n$ für irgendeine Menge $A_1 = A$ schreibt man auch

$$A^n := \underbrace{A \times \dots \times A}_{n\text{-mal}}$$

etwa im Fall $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. Je nach Konvention notiert man statt Tupeln (a_1, \dots, a_n) auch

Spalten $\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, d.h. geht zu

$$A^n = \left\{ \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mid a_1 \in A, \dots, a_n \in A \right\}$$

über.

Bemerkung 1.2.6. Für endliche Mengen A_1, \dots, A_n kann man sich überlegen, dass $\#(A_1 \times \dots \times A_n) = \#A_1 \cdot \dots \cdot \#A_n$ gilt. Diesen Sachverhalt können wir als kleine Rechtfertigung für den Ausdruck kartesisches *Produkt* auffassen.

Das soll zunächst alles zu Mengen und Mengenkonstruktionen gewesen sein. Der zweite Teil dieses Abschnitts widmet sich nun dem Abbildungsbegriff. In der Mathematik geht die Einführung eines so grundlegenden und gebietsübergreifenden Begriffes meist mit etlichen Definitionen einher, die dann für das gesamte weitere Mathematikstudium und auch Lehrer-Dasein von Bedeutung sind, so auch hier.

Definition 1.2.7. Sind A und B Mengen, so ist eine **Abbildung** $f : A \rightarrow B$ eine Vorschrift, die jedem Element $a \in A$ genau ein Element $f(a) \in B$ zuordnet. Hierbei heißt dann A **Definitionsbereich** und B **Wertebereich** von f .

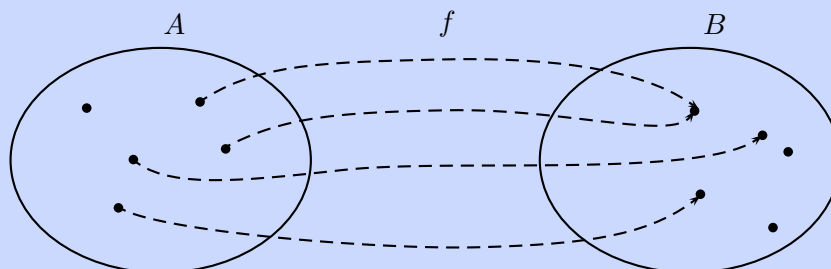


Abb. 3

Beispiel 1.2.8. Beginnen wir mit einigen Beispielen. Man beachte, dass bei der Funktionsdefinition sehr oft $x \mapsto f(x)$ statt $f(x) =$ „Funktionsterm“ notiert wird. Hierbei bedeutet $x \mapsto f(x)$ nichts weiter als: „ x wird abgebildet auf $f(x)$ “.

- 1.) $f : \mathbb{Z} \rightarrow \mathbb{Z}, \quad x \mapsto x^2$
- 2.) $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \quad (x, y) \mapsto x + y$
- 3.) $f : \mathbb{R} \rightarrow \mathbb{N}_0, \quad x \mapsto \begin{cases} 1 & , \text{ falls } x \in \mathbb{Q} \\ 0 & , \text{ falls } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$
- 4.) $\text{id}_A : A \rightarrow A, \quad x \mapsto x$ für eine beliebige Menge A .

Die 4. Abbildung heißt **identische Abbildung** oder Identität.

Bereits im 1. Beispiel wird deutlich, dass nicht auf den gesamten Wertebereich \mathbb{Z} abgebildet wird. Für die Menge der tatsächlich von f getroffenen Elemente gibt es einen Terminus ebenso wie für die Menge der „Rückverfolgungen“ von Bildpunkten im Wertebereich zu den Originalen im Definitionsbereich.

Definition 1.2.9. Sei $f : A \rightarrow B$ eine Abbildung und $A' \subset A$ eine Teilmenge, dann heißt

$$f(A') := \{f(a') \mid a' \in A'\} \subset B$$

das **Bild** von A' unter f .

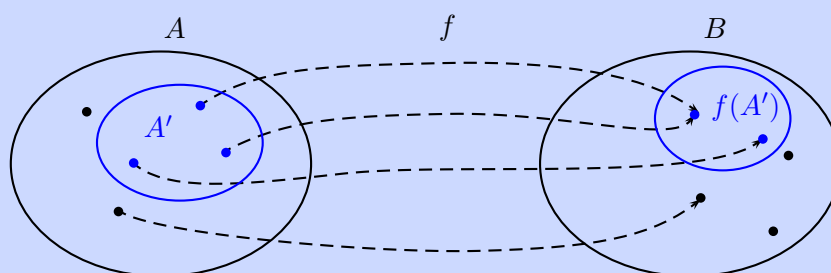


Abb. 4

Definition 1.2.10. Sei $f : A \rightarrow B$ eine Abbildung und $B' \subset B$ eine Teilmenge, dann heißt

$$f^{-1}(B') := \{a \in A \mid f(a) \in B'\} \subset A$$

das **Urbild** von B' unter f .

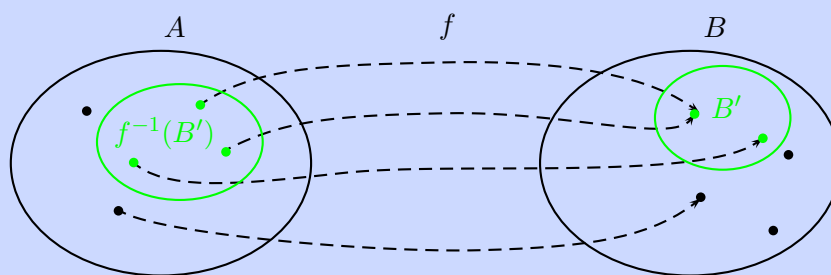


Abb. 5

Bestimmen wir einmal Bild und Urbild an Hand des Beispiels:

$f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x, y) = x + y$.

Sei $A' := \{(0, y) \in \mathbb{R} \times \mathbb{R} \mid y \in \mathbb{R}\}$. Dann ist das Bild von A' unter f gegeben durch

$$\begin{aligned} f(A') &= \{f(0, y) \mid y \in \mathbb{R}\} \\ &= \{0 + y \mid y \in \mathbb{R}\} \\ &= \mathbb{R}. \end{aligned}$$

Sei nun $B' = \{0\}$. Wir bestimmen das Urbild:

$$f^{-1}(B') = \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid f(x, y) \in B'\}$$

1 Grundlagen

$$\begin{aligned} &= \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x + y \in \{0\}\} \\ &= \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x + y = 0\} \\ &= \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid y = -x\} \\ &= \{(x, -x) \mid x \in \mathbb{R}\}. \end{aligned}$$

Beachtlich hierbei ist einerseits $\#B' = 1$ und andererseits $\#f^{-1}(B') = \infty$.

Sehen wir uns nun an, wie wir Abbildungen unter bestimmten Voraussetzungen kombinieren können.

Definition 1.2.11. Seien $g : A \rightarrow B$ und $f : B \rightarrow C$ zwei Abbildungen. Dann heißt

$$f \circ g : A \rightarrow C \quad \text{mit} \quad (f \circ g)(a) = f(g(a))$$

Komposition oder **Verkettung** von f mit g .

Etwa ist für $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$ und $g : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^3$ die Verkettung von f mit g gegeben durch

$$(f \circ g)(x) = f(g(x)) = f(x^3) = (x^3)^2 = x^6.$$

Bemerkung 1.2.12. Ein Wort der Warnung: Anders als vielleicht im letzten Beispiel suggeriert, ist die Verkettung im Allgemeinen nicht kommutativ, d.h. es ist $f \circ g \neq g \circ f$.

Beispiel 1.2.13. Für $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x + 1$, und $g : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$, ist

$$(f \circ g)(x) = f(g(x)) = f(x^2) = x^2 + 1$$

während

$$(g \circ f)(x) = g(f(x)) = g(x + 1) = (x + 1)^2 = x^2 + 2x + 1$$

ist. Für alle $x \neq 0$ ist somit $(f \circ g)(x) \neq (g \circ f)(x)$, also $f \circ g \neq g \circ f$.

Bemerkung 1.2.14. Liegen drei Abbildungen $h : A \rightarrow B$, $g : B \rightarrow C$ und $f : C \rightarrow D$ vor, so können wir zunächst $g \circ h$ bilden und schließlich f dahinterschalten. Dabei kommt dann $f \circ (g \circ h)$ heraus. Andererseits können wir auch erst $(f \circ g)$ bilden und schließlich h von rechts „herankringeln“. Dies liefert $(f \circ g) \circ h$. Diese Unterscheidung macht uns glücklicherweise keine Sorgen, denn in der Tat gilt:

$$(f \circ g) \circ h = f \circ (g \circ h).$$

Durch viermaliges Anwenden der Definition der Verkettung finden wir für alle $a \in A$:

$$((f \circ g) \circ h)(a) = (f \circ g)(h(a)) = f(g(h(a))) = f((g \circ h)(a)) = (f \circ (g \circ h))(a).$$

Die Verkettung von Abbildungen ist also assoziativ und wir brauchen uns keinerlei Gedanken um die Klammerung zu machen. Daher können wir einfach

$$f \circ g \circ h := (f \circ g) \circ h = f \circ (g \circ h)$$

schreiben.

Wir setzen den Reigen der Definitionen mit drei zentralen Eigenschaften von Abbildungen fort.

Definition 1.2.15. Eine Abbildung $f : A \rightarrow B$ heißt **surjektiv** $:\Leftrightarrow$

$$\forall b \in B \exists a \in A : f(a) = b.$$

Eine alternative und überaus kompakte Notation für die Surjektivität lautet: $f(A) = B$.

Jedes Element des Wertebereichs B wird also von mindestens einem Element des Definitionsbereichs getroffen.

Definition 1.2.16. Eine Abbildung $f : A \rightarrow B$ heißt **injektiv** $:\Leftrightarrow$

$$\forall a_1, a_2 \in A : f(a_1) = f(a_2) \Rightarrow a_1 = a_2.$$

Wie wir natürlich sofort erkennen, ist $\forall a_1, a_2 \in A : a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)$ dazu äquivalent.

Auf jedes Element in B wird also höchstens einmal abgebildet wird. Jedes Element der Zielmenge ist somit höchstens Bild von einem Element des Definitionsbereiches.

Definition 1.2.17. Eine Abbildung $f : A \rightarrow B$ heißt **bijektiv** $:\Leftrightarrow f$ ist injektiv und surjektiv.

1 Grundlagen

Die definierenden Eigenschaften surjektiv und injektiv bringen die Begriffe mindestens und höchstens mit sich. Zusammengenommen ist f somit bijektiv \Leftrightarrow

$$\forall b \in B \exists ! a \in A : f(a) = b.$$

Eine bijektive Abbildung f ist stets umkehrbar, d.h. sie besitzt eine Umkehrabbildung f^{-1} , welche auch als inverse Abbildung bezeichnet wird und für die dann $f \circ f^{-1} = \text{id}_B$ und $f^{-1} \circ f = \text{id}_A$ gilt. Die Eigenschaften surjektiv und injektiv liefern dagegen nur die Existenz einer rechtsinversen bzw. linksinversen Abbildung. Formulieren wir dies in folgendem

Satz 1.2.18

Sei $f : A \rightarrow B$ eine Abbildung mit $A \neq \emptyset$. Dann gilt:

1. f ist surjektiv $\Leftrightarrow \exists g : B \rightarrow A : f \circ g = \text{id}_B$.
2. f ist injektiv $\Leftrightarrow \exists g : B \rightarrow A : g \circ f = \text{id}_A$.
3. f ist bijektiv $\Leftrightarrow \exists g : B \rightarrow A : f \circ g = \text{id}_B \ \& \ g \circ f = \text{id}_A$.

Beweis.

Zu 1.)

Für eine Äquivalenz müssen wir die beiden Beweisrichtungen „ \Leftarrow “ und „ \Rightarrow “ überprüfen.

„ \Leftarrow “:

Sei $g : B \rightarrow A$, so dass $f \circ g = \text{id}_B$. Wir zeigen, dass f dann surjektiv ist. Die Surjektivität ist eine All-Aussage (vgl. Definition). Sei daher also $b \in B$ beliebig. Nun setzen wir $a := g(b) \in A$, wenden auf dieses Element die Abbildung f an und nutzen schließlich die Voraussetzung $f \circ g = \text{id}_B$. Es gilt:

$$f(a) = f(g(b)) = (f \circ g)(b) = \text{id}_B(b) = b.$$

„ \Rightarrow “:

Sei nun f surjektiv. Für beliebiges $b \in B$ konstruieren wir g wie folgt: Da f surjektiv ist, existiert mindestens ein (von b abhängiges) Element $a_b \in A$ mit $f(a_b) = b$. Dann setzen wir einfach $g(b) := a_b$, d.h. die Abbildung wählt zu jedem vorgegebenen b ein Urbild a_b aus. Dann gilt für alle $b \in B$:

$$(f \circ g)(b) = f(g(b)) = f(a_b) = b = \text{id}_B(b),$$

also

$$f \circ g = \text{id}_B.$$

Zu 2.)

„ \Leftarrow “:

Sei $g : B \rightarrow A$, so dass $g \circ f = \text{id}_A$. Wir zeigen, dass f injektiv ist. Seien dazu $a_1, a_2 \in A$ mit $f(a_1) = f(a_2)$ vorgegeben. Wir wenden die Abbildung g auf beide Seiten an und erhalten

$$\begin{aligned} g(f(a_1)) &= g(f(a_2)) \\ \Rightarrow (g \circ f)(a_1) &= (g \circ f)(a_2) \\ \Rightarrow \text{id}_A(a_1) &= \text{id}_A(a_2) \\ \Rightarrow a_1 &= a_2. \end{aligned}$$

„ \Rightarrow “:

Sei nun f injektiv. An dieser Stelle geht ein, dass A nichtleer ist. Demnach existiert ein Element $a_0 \in A$, welches wir als „Sonstiges-Element“ nutzen und somit $g : B \rightarrow A$ wie folgt definieren:

$$g(b) := \begin{cases} \text{eindeutiges } a \in A \text{ mit } f(a) = b & , \text{ falls } b \in f(A) \\ a_0 & , \text{ sonst.} \end{cases}$$

Dann gilt für alle $a \in A$:

$$(g \circ f)(a) = g(f(a)) = a = \text{id}_A(a).$$

Zu 3.)

„ \Leftarrow “:

Sei $g : B \rightarrow A$, so dass d.h. $f \circ g = \text{id}_B$ und $g \circ f = \text{id}_A$ gilt. Dann folgt nach „ \Leftarrow “ aus 1.), dass f surjektiv ist und nach „ \Leftarrow “ aus 2.), dass f injektiv ist. Definitionsgemäß ist f somit bijektiv.

„ \Rightarrow “:

Sei f bijektiv, also surjektiv und injektiv.

Nach „ \Leftarrow “ aus 1.) folgt die Existenz einer Abbildung $g_1 : B \rightarrow A : f \circ g_1 = \text{id}_B$ und

nach „ \Leftarrow “ aus 2.) folgt die Existenz einer Abbildung $g_2 : B \rightarrow A : g_2 \circ f = \text{id}_A$.

A priori ist nicht klar, dass $g_1 = g_2$ gilt. Dies sehen wir aber wie folgt ein:

$$g_1 = \text{id}_A \circ g_1 = (g_2 \circ f) \circ g_1 = g_2 \circ f \circ g_1 = g_2 \circ (f \circ g_1) = g_2 \circ \text{id}_B = g_2.$$

Somit hat $g := g_1 = g_2$ die verlangten Eigenschaften. □

Bemerkung 1.2.19. g wie in 3.) heißt dann die zu f inverse Abbildung, für die wir statt g fortan f^{-1} schreiben.

Bemerkung 1.2.20. Zwei Mengen A und B haben gleich viele Elemente, d.h. sind **gleichmächtig**, wenn es eine bijektive Abbildung $f : A \rightarrow B$ zwischen ihnen gibt. Statt Bijektion sagt man auch eineindeutige Zuordnung oder 1:1-Beziehung.

1 Grundlagen

Zum Nachdenken präsentieren wir: \mathbb{N}_0 und \mathbb{Z} sind gleichmächtig, da folgende Abbildung $f : \mathbb{N}_0 \rightarrow \mathbb{Z}$ bijektiv ist:

a	0	1	2	3	4	5	6	...
$f(a)$	0	-1	1	-2	2	-3	3	...

Obwohl \mathbb{N}_0 eine echte Teilmenge von \mathbb{Z} ist, hat \mathbb{N}_0 nicht weniger Elemente als \mathbb{Z} !

Ein komplexeres Beispiel, welches die „verschiedenen Typen“ des Unendlichkeitsbegriffes illustriert, ist bekannt unter dem Namen **Hilberts Hotel**.²

In einem Hotel mit endlich vielen Zimmern können keine Gäste mehr aufgenommen werden, sobald alle Zimmer belegt sind. Stellen wir uns nun ein Hotel mit unendlich vielen Zimmern vor, durchnummeriert, beginnend bei 1 und mit nur einem möglichen Gast pro Zimmer. Man könnte annehmen, dass dasselbe Problem auch hier auftritt. Die naive Vermutung hierzu wäre: Wenn unendlich viele Gäste im Hotel sind, kann kein weiterer Gast aufgenommen werden. Sehen wir uns einmal an, welche Erlebnisse eine Arbeitswoche in Hilberts Hotel bereithält:



Abb. 6

Montag: Alle Zimmer sind belegt und ein neuer Gast reist an.

Der Hoteldirektor bittet den Gast im Zimmer 1 nach Zimmer 2 umzuziehen, den Gast im Zimmer 2 nach Zimmer 3 umzuziehen, usw. Auf diese Weise ziehen alle „alten“ Gäste in das Zimmer mit der nächsthöheren Zimmernummer um. Dadurch wird Zimmer 1 frei und der neue Gast kann einziehen. Mathematisch steckt folgende injektive und nicht surjektive „Umzieh-Abbildung“ hinter diesem Prinzip:

$$f : \mathbb{N} \rightarrow \mathbb{N}, n \mapsto n + 1.$$

Die Injektivität stellt dabei sicher, dass niemals zwei Gäste in einem Zimmer einquartiert werden. Andererseits ist f nicht surjektiv, denn offenbar wird $1 \in \mathbb{N}$ nicht als Bild von f angenommen, d.h. es gilt

$$f(\mathbb{N}) = \{f(n) \mid n \in \mathbb{N}\} = \mathbb{N} \setminus \{1\} = \{2, 3, 4, \dots\},$$

also $1 \notin f(\mathbb{N})$.

Dienstag: Alle Zimmer sind belegt und ein Bus mit k neuen Gästen trifft ein ($k \in \mathbb{N}$).

Wieder hat der clevere Hoteldirektor eine Lösung parat: Jeder „alte“ Gast zieht in das Zimmer, dessen Nummer um k größer ist als das bisherige. Dadurch werden die Zimmer $1, \dots, k$ frei. Diesmal liegt die injektive aber nicht surjektive Abbildung

$$f : \mathbb{N} \rightarrow \mathbb{N}, n \mapsto n + k,$$

²benannt nach dem berühmten Mathematiker David Hilbert (1862 (Königsberg) - 1943 (Göttingen)), siehe http://de.wikipedia.org/wiki/David_Hilbert

zu Grunde, für die

$$f(\mathbb{N}) = \mathbb{N} \setminus \{1, \dots, k\} = \{k+1, k+2, \dots\}$$

gilt. Daher ist $\{1, \dots, k\} \cap f(\mathbb{N}) = \emptyset$.

Mittwoch: Alle Zimmer sind belegt und ein Bus mit ∞ vielen Gästen trifft ein, deren Sitzplätze mit $1, 2, 3, \dots$ durchnummeriert sind.

Der Plan: Die „alten“ Gäste ziehen in ein Zimmer mit der doppelten Zimmernummer und die neuen Gäste belegen dann die ungeraden Zimmernummern. Hier ist unsere injektive und nicht surjektive „Umzieh-Abbildung“ durch

$$f : \mathbb{N} \rightarrow \mathbb{N}, n \mapsto 2n,$$

gegeben. Das Bild der Abbildung lautet

$$f(\mathbb{N}) = \{2n \mid n \in \mathbb{N}\}.$$

Für die neu eingetroffene Gäste nehmen wir die „Neubelegungs-Abbildung“:

$$g : \mathbb{N} \rightarrow \mathbb{N}, n \mapsto 2n - 1.$$

Hierbei ist g injektiv und es gilt

$$g(\mathbb{N}) = \{2n - 1 \mid n \in \mathbb{N}\}$$

und damit $f(\mathbb{N}) \cap g(\mathbb{N}) = \emptyset$.

Donnerstag: Alle Zimmer sind belegt und ∞ viele Busse mit jeweils ∞ vielen Gästen treffen ein, wobei Busse und Sitzplätze jeweils mit $1, 2, 3, \dots$ durchnummeriert sind.

Donnerstag ist ein sehr stressiger Tag in Hilberts Hotel. Die „alten“ Gäste müssen zunächst wie am Mittwoch in die Zimmer mit den doppelten Zimmernummern umziehen. Wie wir wissen sind diese Zimmernummern dann insbesondere gerade. Diese Erkenntnis und sein Fable für Euklid bringen den Hoteldirektor auf folgende Idee: Er ordnet als erstes jedem Bus eine ungerade Primzahl zu, d.h.

$$p_1 = 3, p_2 = 5, p_3 = 7, p_4 = 11, \dots$$

Genau wie uns, ist auch ihm bewusst, dass es unendlich viele Primzahlen gibt, so dass wir alle Busse mit je einer eigenen ungeraden Primzahl versorgen können. Die Fahrgäste aus Bus 1 belegen nun nach und nach die Zimmer $3^1, 3^2, 3^3, \dots$. Die Fahrgäste aus Bus 2 belegen dann sukzessive die Zimmer $5^1, 5^2, 5^3, \dots$. Damit hat es der Direktor mal wieder geschafft!

Nun legen wir uns diese Angelegenheit noch einmal mathematisch zurecht, behalten dazu unsere „Mittwochs-Umzieh-Abbildung“ $f : \mathbb{N} \rightarrow \mathbb{N}$ mit $f(n) = 2n$ bei und definieren schließlich die „Neubelegungs-Abbildung“

$$g : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}, (m, n) \mapsto p_m^n.$$

1 Grundlagen

Dabei beschreibt m die Busnummer und n die Sitzplatznummer. Überlegen wir also, dass g injektiv ist: Für alle (m, n) und (m', n') folgt aus $g(m, n) = g(m', n')$, dass $p_m^n = p_{m'}^{n'}$, also $p_m = p_{m'}$ & $n = n'$, d.h. $m = m'$ & $n = n'$. Nun ist

$$f(\mathbb{N}) \cap g(\mathbb{N}) \subset \{n \in \mathbb{N} \mid n \text{ gerade}\} \cap \{n \in \mathbb{N} \mid n \text{ ungerade}\} = \emptyset$$

und daher $f(\mathbb{N}) \cap g(\mathbb{N}) = \emptyset$.

Freitag: Alle Zimmer sind frei und ein neuer Bus mit ∞ vielen Gästen trifft ein, wobei die Sitzplätze mit den reellen Zahlen des Intervalls $(0, 1)$ durchnummeriert sind, also unter anderem mit $\frac{1}{5}$, $\frac{\sqrt{2}}{2}$ oder etwa $\frac{\pi}{10}$.

Ausgerechnet heute muss der Hoteldirektor passen und kann zu seinem Bedauern nicht allen angereisten Gästen ein Zimmer anbieten. Wie man nun richtig vermutet, unterscheidet sich die Anzahl ∞ der Sitzplätze im Bus zur Anzahl ∞ der Hotelzimmer. Exakt begründen wir diesen Sachverhalt mit

Satz 1.2.21

Es gibt keine injektive Abbildung $g : (0, 1) \rightarrow \mathbb{N}$.

Beweis.

Wir machen die Widerspruchsannahme, dass es doch eine injektive Abbildung $g : (0, 1) \rightarrow \mathbb{N}$ gibt und nutzen das so genannte Cantor'sche Diagonalverfahren. Obige Annahme gestattet es, die Elemente $x_i \in (0, 1)$ wie folgt aufzulisten bzw. aufzuzählen:

„Sitzplatz im Bus“	„zugeordnete Zimmernummer“
$x_1 = 0, a_{11}a_{12}a_{13} \dots$	$g(x_1)$ minimal
$x_2 = 0, a_{21}a_{22}a_{23} \dots$	$g(x_2)$ zweitkleinstes Element
$x_3 = 0, a_{31}a_{32}a_{33} \dots$	$g(x_3)$ drittkleinstes Element
\vdots	\vdots

Dabei ist $a_{ij} \in \{0, \dots, 9\}$ die j -te Ziffer in der Dezimalentwicklung der i -ten Zahl x_i . Wir konstruieren nun eine neue reelle Zahl $y = 0, b_1b_2b_3 \dots$ mit $b_j \in \{0, \dots, 9\}$ und setzen hierzu

$$b_j := \begin{cases} 3, & \text{falls } a_{jj} \neq 3 \\ 4, & \text{falls } a_{jj} = 3. \end{cases}$$

Dann erhalten wir in der Tat eine reelle Zahl $y \in (0, 1)$ mit der Nachkommaunterscheidung $b_j \neq a_{jj}$, so dass für alle $j \in \mathbb{N}$ gilt: $y \neq x_j$. Also ist $y \in (0, 1)$ eine neue reelle Zahl, die noch nicht in unserer Liste auftaucht:

⚡ zur Annahme, dass alle Zahlen aus $(0, 1)$ bereits in der Liste vorkommen. □

1.2 Mengen und Abbildungen

Nach all diesen Grundlagen steigen wir im folgenden Kapitel nun endlich in den eigentlichen Stoff der linearen Algebra und analytischen Geometrie ein.

2 Matrixrechnung

Wake up, Neo. The Matrix has you. Follow the white rabbit.

(Larry und Andy Wachowski
The Matrix)

2.1 Lineare Gleichungssysteme

Zunächst wärmen wir uns an ein paar Beispielen für **lineare Gleichungssysteme** (LGS) auf.

Beispiel 2.1.1. Gegeben seien zwei Obstkörbe, einer voller Äpfel und einer gefüllt mit Bananen. Wir bestimmen einmal das Gewicht der Körbe auf der Grundlage folgender zwei Wiegeungen bzw. Messungen:

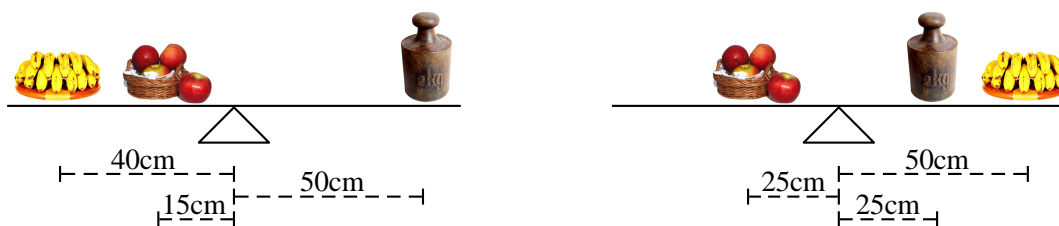


Abb. 7

Sei nun A das Gewicht der Äpfel (in kg) und B das Gewicht der Bananen (in kg). Wir erhalten

$$\begin{aligned} A \cdot 15 + B \cdot 40 &= 2 \cdot 50 \\ A \cdot 25 &= 2 \cdot 25 + B \cdot 50 \end{aligned} \quad (2.1)$$

Dieses System ist äquivalent zu

$$\begin{aligned} 3 \cdot A + 8 \cdot B &= 20 \\ A - 2B &= 2 \end{aligned}$$

Addition der ersten Zeile mit dem -3 -fachen der zweiten Zeile liefert

$$14 \cdot B = 14 \Rightarrow B = 1$$

2 Matrixrechnung

und anschließendes Auflösen der zweiten Gleichung dann

$$A = 2 + 2B = 4.$$

Bisher ist damit gezeigt: Wenn das LGS (2.1) Lösungen besitzt, so müssen es $A = 4$ und $B = 1$ sein. Einsetzen der Lösungen in (2.1) lässt den Umkehrschluss zu, dass für $A = 4$ und $B = 1$ die Gleichungen (2.1) gelten.

Insgesamt ist dann (2.1) erfüllt $\Leftrightarrow A = 4 \ \& \ B = 1$.

Example 2.1.2

Nun zu einem explosiven Beispiel aus der Chemie. Betrachtet werden die Substanzen **Toluol** C_7H_8 , **Salpetersäure** HNO_3 , **Wasser** H_2O sowie **Trinitrotoluol** $C_7H_5O_6N_3$, besser bekannt als der Sprengstoff TNT. Die zugehörige Reaktionsgleichung lautet:



Abb. 8



Bekanntlich bleiben chemische Elemente bei (chemischen) Reaktion erhalten. Aus (2.2) erhalten wir dann folgendes LGS mit vier Gleichungen in vier Unbekannten für:

$$\begin{array}{ll} \text{Kohlenstoff: } C : & 7 \cdot x + 0 \cdot y = 7 \cdot z + 0 \cdot w \\ \text{Wasserstoff: } H : & 8 \cdot x + 1 \cdot y = 5 \cdot z + 2 \cdot w \\ \text{Stickstoff: } N : & 0 \cdot x + 1 \cdot y = 3 \cdot z + 0 \cdot w \\ \text{Sauerstoff: } O : & 0 \cdot x + 3 \cdot y = 6 \cdot z + 1 \cdot w \end{array}$$

Dieses LGS ist äquivalent zu

$$\begin{array}{rcl} x & = & z \\ 8x + y & = & 5z + 2w \\ y & = & 3z \\ 3y & = & 6z + w \end{array}$$

Da die zweite Gleichung aus den Gleichungen eins, drei und vier folgt, ist das System sogar unterbestimmt und damit äquivalent zu

$$\begin{array}{rcl} x & = & z \\ y & = & 3z \\ w & = & 3z \end{array}$$

Die Lösungsmenge ist also

$$\left\{ \left(\begin{array}{c} z \\ 3z \\ z \\ 3z \end{array} \right) \mid z \in \mathbb{R} \right\},$$

2.1 Lineare Gleichungssysteme

d.h. wir erhalten eine unendliche Lösungsmenge, denn jedes $z \in \mathbb{R}$ liefert eine Lösung. Dies ist auch genau, was die Aufgabenstellung erwarten lässt, denn in der Reaktionsgleichung können natürlich die einzelnen Zutaten für das TNT mengenmäßig vervielfacht werden, und wir erhalten dann die entsprechende vielfache Menge an TNT und Wasser. Daher ergibt es nicht nur eine Lösung, sondern unendlich viele.

Example 2.1.3

Als drittes und letztes Beispiel kommen wir nun noch zu einem Exkurs in die Finanzwelt. Und zwar wollen wir Aktien erwerben, wobei sich unser Interesse an folgenden drei DAX-Unternehmen ausrichtet:



Abb. 9

Anzahl	Firma	Kurs (07.11.08, 13:14)
x	(B)MW	$K_B = 21,18$ Euro
y	(D)aimler	$K_D = 24,69$ Euro
z	(M)ünchener Rück	$K_M = 100,78$ Euro

Insgesamt beträgt unsere Anlagesumme 10 000 Euro. Somit haben wir:

$$x \cdot K_B + y \cdot K_D + z \cdot K_M = 10\,000$$

Da die bayerischen Finanztitel in letzter Zeit besonders gelitten haben, gewichten wir zu Gunsten von Daimler wie folgt:

$$x \cdot K_B + 2 \cdot z \cdot K_M = y \cdot K_D$$

Damit kaufen wir nun weniger Aktien von BMW bzw. Münchener Rück. Nun wollen wir aber auch die momentane Finanzkrise bei den Automobilherstellern berücksichtigen, d.h. es wird eine erneute Gewichtung fällig:

$$x \cdot K_B + 5 \cdot y \cdot K_D = z \cdot K_M$$

Wir interessieren uns also für Lösungen $x, y, z \in \mathbb{R}$ und nehmen daher an, es gäbe solche. Unser LGS lautet:

$$\begin{aligned} x \cdot K_B + y \cdot K_D + z \cdot K_M &= 10\,000 \\ x \cdot K_B - y \cdot K_D + 2z \cdot K_M &= 0 \\ x \cdot K_B + 5y \cdot K_D - z \cdot K_M &= 0 \end{aligned}$$

Subtraktion der zweiten von der ersten Gleichung und der dritten von der zweiten Gleichung ergibt:

$$\begin{aligned} 2y \cdot K_D - z \cdot K_M &= 10\,000 \\ -6y \cdot K_D + 3z \cdot K_M &= 0 \end{aligned}$$

2 Matrixrechnung

Wir vereinfachen weiter und bekommen:

$$\begin{aligned}2y \cdot K_D - z \cdot K_M &= 10\,000 \\ -2yK_D + zK_M &= 0\end{aligned}$$

Schließlich führt die Addition dieser beiden Zeilen auf:

$$0 + 0 = 10\,000 \quad \zeta$$

Der Blitz deutet auf das Ende eines Widerspruchsbeweises hin und tatsächlich haben wir eben, wenn auch nicht in vollem Bewusstsein, einen solchen geführt! Denn: Die Nicht-Existenz von Lösungen negierten wir durch die Widerspruchs-Annahme der Existenz von Lösungen $x, y, z \in \mathbb{R}$. Daraus ergab sich obiger Widerspruch. Folglich war die Annahme falsch und das LGS besitzt in der Tat keine Lösungen.

Problematisierung. Die Beispiele zeigen, dass es keine, genau eine und unendlich viele Lösungen geben kann.

Aber sind dies bereits alle Möglichkeiten? Oder gibt es vielleicht auch LGS mit genau zwei Lösungen? Führen mehr als eine Lösung stets zur Existenz unendlich vieler Lösungen? Und unter welchen Bedingungen entstehen welche qualitativen Typen von Lösungsmengen?

Um Antworten auf diese Fragen zu erhalten, studieren wir lineare Gleichungssysteme nun systematisch und beginnen mit folgender formaler

Definition 2.1.4. Ein **lineares Gleichungssystem** (LGS) ist ein Gleichungssystem der Form

$$\begin{aligned}A_{11} \cdot x_1 + A_{12} \cdot x_2 + \dots + A_{1n} \cdot x_n &= b_1 \\ &\vdots \\ A_{m1} \cdot x_1 + A_{m2} \cdot x_2 + \dots + A_{mn} \cdot x_n &= b_m\end{aligned}$$

Hierbei ist n die Anzahl der Unbekannten und m die Anzahl der Gleichungen. Man beachte, dass m und n im Allgemeinen nicht übereinzustimmen brauchen.

Die **Koeffizienten** A_{ij} sowie die $b_i \in \mathbb{R}$ sind vorgegeben. Die **Unbekannten** x_1, \dots, x_n dagegen sind gesucht.

Gilt $b_1 = \dots = b_m = 0$, so heißt das LGS *homogen* und ansonsten *inhomogen*.

So lagen etwa dem Obstkorb-Beispiel ebenso wie dem Beispiel aus der Finanzwelt inhomogene LGS zu Grunde. Das LGS im Falle des Sprengstoff-Beispiels war homogen.

Mehr und mehr wird deutlich, dass es nicht auf die Platzhalter x_1, \dots, x_n ankommt sondern im Wesentlichen nur auf die Koeffizienten A_{ij} , für die wir nun folgende kompakte Matrix-Schreibweise einführen:

Definition 2.1.5. Eine $m \times n$ -**Matrix** reeller Zahlen ist ein rechteckiges Schema der Form

$$\begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix},$$

wobei $A_{ij} \in \mathbb{R}$ für alle $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$ ist.

So ist zum Beispiel $\begin{pmatrix} 0 & 0 \end{pmatrix}$ eine 1×2 -Matrix, $\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$ eine 3×2 -Matrix und $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ eine 3×1 -Matrix. Wir werden nachfolgend die Menge aller reellen $m \times n$ -Matrizen mit $\text{Mat}(m \times n, \mathbb{R})$ bezeichnen. Für uns ist dann also zum Beispiel $\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in \text{Mat}(3 \times 2, \mathbb{R})$.

Definition 2.1.6. Eine $1 \times n$ -Matrix nennen wir **Zeilenvektor** und eine $m \times 1$ -Matrix **Spaltenvektor**.

Nachfolgend fassen wir Tupel bzw. Elemente aus \mathbb{R}^n stets als Spaltenvektoren, d.h. $m \times 1$ -Matrizen auf. Folglich gilt:

$$\mathbb{R}^n = \text{Mat}(n \times 1, \mathbb{R})$$

Bemerkung 2.1.7. Als Konvention für die Einträge von Matrizen A bzw. B aus $\text{Mat}(m \times n, \mathbb{R})$ nutzen wir der Definition entsprechend A_{ij} bzw. B_{ij} also

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix} \quad \text{bzw.} \quad B = \begin{pmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & \ddots & \vdots \\ B_{m1} & \cdots & B_{mn} \end{pmatrix}.$$

Sehen wir uns einmal an, welche Operationen mit Matrizen möglich sind.

Addition von Matrizen. Seien $A, B \in \text{Mat}(m \times n, \mathbb{R})$ mit der Darstellung wie oben. Die Matrizen stimmen also in Höhe und Breite überein. Dann hat $A + B$ die Einträge

$$(A + B)_{ij} = A_{ij} + B_{ij}.$$

2 Matrixrechnung

Man spricht auch davon, dass die Addition komponentenweise erklärt ist, etwa im Fall

$$\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} + \begin{pmatrix} 7 & 10 \\ 8 & 11 \\ 9 & 12 \end{pmatrix} = \begin{pmatrix} 1+7 & 4+10 \\ 2+8 & 5+11 \\ 3+9 & 6+12 \end{pmatrix} = \begin{pmatrix} 8 & 14 \\ 10 & 16 \\ 12 & 18 \end{pmatrix}$$

oder im Spezialfall

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

Für die prägnante Unterscheidung zwischen Matrizen reeller Zahlen und reellen Zahlen selbst bezeichnet man letztere im Sprachgebrauch auch häufig als **Skalare**.

Multiplikation von Matrizen mit Skalaren. Sei $A \in \text{Mat}(m \times n, \mathbb{R})$ mit Einträgen wie oben und $\lambda \in \mathbb{R}$. Dann hat $\lambda \cdot A \in \text{Mat}(m \times n, \mathbb{R})$ die Einträge

$$(\lambda \cdot A)_{ij} = \lambda \cdot A_{ij},$$

z.B.

$$2 \cdot \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 & 2 \cdot 4 \\ 2 \cdot 2 & 2 \cdot 5 \\ 2 \cdot 3 & 2 \cdot 6 \end{pmatrix} = \begin{pmatrix} 2 & 8 \\ 4 & 10 \\ 6 & 12 \end{pmatrix}.$$

Naiv könnte man nun auf die Idee kommen, auch die Multiplikation zweier Matrizen komponentenweise zu definieren. Wie sich herausstellen wird, ist dies wenig sinnvoll, wohl aber die folgende Definition der

Multiplikation zweier Matrizen. Sei $A \in \text{Mat}(m \times n, \mathbb{R})$ und $B \in \text{Mat}(n \times k, \mathbb{R})$. Die Breite der ersten Matrix muss also mit der Höhe der zweiten Matrix übereinstimmen. Insbesondere tritt die Größe n im Resultat nicht mehr auf, d.h. es gilt $A \cdot B \in \text{Mat}(m \times k, \mathbb{R})$.

Bevor wir nun die recht komplizierte Formel für die Matrix-Multiplikation in Augenschein nehmen, betrachten wir zunächst einmal eine Möglichkeit der schematischen Ausführung der Multiplikation $A \cdot B =: C$, das so genannte Falk-Schema:

$$\begin{array}{c|c} & B \\ \hline A & C \end{array} \quad \text{bzw.} \quad \begin{array}{ccc|ccc} & & & B_{11} & B_{12} & B_{13} \\ & & & B_{21} & B_{22} & B_{23} \\ & & & B_{31} & B_{32} & B_{33} \\ \hline A_{11} & A_{12} & A_{13} & C_{11} & C_{12} & C_{13} \\ A_{21} & A_{22} & A_{23} & C_{21} & C_{22} & C_{23} \\ A_{31} & A_{32} & A_{33} & C_{31} & C_{32} & C_{33} \end{array}$$

Zur Berechnung, etwa von C_{23} , geht man nun simultan die 2. Zeile von A und die 3. Spalte von B durch, multipliziert in jedem Schritt die entsprechenden Einträge und addiert schließlich die

Produkte auf:

$$(A \cdot B)_{23} = C_{23} = A_{21} \cdot B_{13} + A_{22} \cdot B_{23} + A_{23} \cdot B_{33}$$

Ganz allgemein und für jede Kombination von Zeilen i und Spalten j tauglich ist die Formel:

$$\begin{aligned} (A \cdot B)_{ij} &= A_{i1} \cdot B_{1j} + A_{i2} \cdot B_{2j} + \dots + A_{in} \cdot B_{nj} \\ &= \sum_{l=1}^n A_{il} \cdot B_{lj} \end{aligned}$$

Zum Beispiel gilt:

$$\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 4 \cdot 0 & 1 \cdot 1 + 4 \cdot 1 \\ 2 \cdot 1 + 5 \cdot 0 & 2 \cdot 1 + 5 \cdot 1 \\ 3 \cdot 1 + 6 \cdot 0 & 3 \cdot 1 + 6 \cdot 1 \end{pmatrix} = \begin{pmatrix} 1 & 5 \\ 2 & 7 \\ 3 & 9 \end{pmatrix} \in \text{Mat}(3 \times 2, \mathbb{R})$$

Bemerkung 2.1.8. Fassen wir die drei Operationen noch einmal in Gestalt von Abbildungen zusammen und notieren nochmals, auf welchen Mengen diese arbeiten:

$$\begin{aligned} (A, B) &\mapsto A + B && \text{auf } \text{Mat}(m \times n, \mathbb{R}) \times \text{Mat}(m \times n, \mathbb{R}) \rightarrow \text{Mat}(m \times n, \mathbb{R}) \\ (\lambda, A) &\mapsto \lambda \cdot A && \text{auf } \mathbb{R} \times \text{Mat}(m \times n, \mathbb{R}) \rightarrow \text{Mat}(m \times n, \mathbb{R}) \\ (A, B) &\mapsto A \cdot B && \text{auf } \text{Mat}(m \times n, \mathbb{R}) \times \text{Mat}(n \times k, \mathbb{R}) \rightarrow \text{Mat}(m \times k, \mathbb{R}) \end{aligned}$$

Ungeklärt ist, weshalb die dritte Operation auf diese komplizierte Art und Weise definiert wurde und welche Rechenregeln hier überhaupt gelten und welche nicht.

Satz 2.1.9

Reelle Matrizen genügen den folgenden Rechenregeln:

1. $\forall A, B, C \in \text{Mat}(m \times n, \mathbb{R}) :$

$$A + B = B + A \text{ und } (A + B) + C = A + (B + C)$$

2. $\forall \lambda \in \mathbb{R} \forall A \in \text{Mat}(m \times n, \mathbb{R}) \forall B \in \text{Mat}(n \times k, \mathbb{R}) :$

$$(\lambda \cdot A) \cdot B = A \cdot (\lambda \cdot B) = \lambda \cdot (A \cdot B)$$

3. $\forall A \in \text{Mat}(m \times n, \mathbb{R}) \forall B \in \text{Mat}(n \times k, \mathbb{R}) \forall C \in \text{Mat}(k \times l, \mathbb{R}) :$

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

2 Matrixrechnung

4. $\forall A, B \in \text{Mat}(m \times n, \mathbb{R}) \forall C \in \text{Mat}(n \times k, \mathbb{R}) :$

$$(A + B) \cdot C = A \cdot C + B \cdot C$$

Beweis.

Wir zeigen alle Teile außer der 2. Aussage. Diese verbleibt als Übungsaufgabe.

Zu 1.)

Seien $A, B \in \text{Mat}(m \times n, \mathbb{R})$. Dann gilt für alle $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$:

$$\begin{aligned}(A + B)_{ij} &= A_{ij} + B_{ij} \\ &= B_{ij} + A_{ij} \\ &= (B + A)_{ij}\end{aligned}$$

Daraus folgt: $A+B = B+A$. Ähnlich sieht man auch die zweite Gleichung ein: Seien $A, B, C \in \text{Mat}(m \times n, \mathbb{R})$. Dann gilt für jedes $i \in \{1, \dots, m\}$ und jedes $j \in \{1, \dots, n\}$:

$$\begin{aligned}((A + B) + C)_{ij} &= (A + B)_{ij} + C_{ij} \\ &= (A_{ij} + B_{ij}) + C_{ij} \\ &= A_{ij} + (B_{ij} + C_{ij}) \\ &= A_{ij} + (B + C)_{ij} \\ &= (A + (B + C))_{ij}\end{aligned}$$

Aus der Gleichheit aller Komponenten folgt $(A + B) + C = A + (B + C)$.

Zu 2.)

Übungsaufgabe

Zu 3.)

Sei $A \in \text{Mat}(m \times n, \mathbb{R})$, sei $B \in \text{Mat}(n \times k, \mathbb{R})$ und sei $C \in \text{Mat}(k \times l, \mathbb{R})$. Dann gilt für alle $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, l\}$:

$$\begin{aligned}((A \cdot B) \cdot C)_{ij} &= \sum_{s=1}^k (A \cdot B)_{is} \cdot C_{sj} \\ &= \sum_{s=1}^k \left(\sum_{r=1}^n A_{ir} \cdot B_{rs} \right) \cdot C_{sj} \\ &= \sum_{s=1}^k \sum_{r=1}^n (A_{ir} \cdot B_{rs} \cdot C_{sj})\end{aligned}$$

$$\begin{aligned}
&= \sum_{r=1}^n \sum_{s=1}^k (A_{ir} \cdot B_{rs} \cdot C_{sj}) \\
&= \sum_{r=1}^n A_{ir} \cdot \sum_{s=1}^k (B_{rs} \cdot C_{sj}) \\
&= \sum_{r=1}^n A_{ir} \cdot (B \cdot C)_{rj} \\
&= (A \cdot (B \cdot C))_{ij}
\end{aligned}$$

Es folgt: $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

Zu 4.)

Seien $A, B \in \text{Mat}(m \times n, \mathbb{R})$ und sei $C \in \text{Mat}(n \times k, \mathbb{R})$. Dann gilt für alle $i \in \{1, \dots, m\}$ und jedes $j \in \{1, \dots, k\}$:

$$\begin{aligned}
((A + B) \cdot C)_{ij} &= \sum_{l=1}^n (A + B)_{il} \cdot C_{lj} \\
&= \sum_{l=1}^n (A_{il} + B_{il}) \cdot C_{lj} \\
&= \sum_{l=1}^n (A_{il} \cdot C_{lj} + B_{il} \cdot C_{lj}) \\
&= \sum_{l=1}^n A_{il} \cdot C_{lj} + \sum_{l=1}^n B_{il} \cdot C_{lj} \\
&= (A \cdot C)_{ij} + (B \cdot C)_{ij} \\
&= (A \cdot C + B \cdot C)_{ij}
\end{aligned}$$

Dies zeigt $(A + B) \cdot C = A \cdot C + B \cdot C$. □

Bemerkung 2.1.10. Diese Rechenregeln versetzen uns in die Lage, mit Matrizen fast so wie mit reellen Zahlen zu rechnen. Von der Kommutativität des Produktes zweier Matrizen kann jedoch im Allgemeinen nicht ausgegangen werden, denn sei $A \in \text{Mat}(m \times n, \mathbb{R})$ und $B \in \text{Mat}(n \times m, \mathbb{R})$. Wie wir wissen, ist dann $A \cdot B \in \text{Mat}(m \times m, \mathbb{R})$ und $B \cdot A \in \text{Mat}(n \times n, \mathbb{R})$. Gilt nun $m \neq n$, so besitzen die Matrizen offenbar verschiedene Abmaße und insbesondere gilt somit $A \cdot B \neq B \cdot A$. Ist andererseits $m = n$, kann sowohl der Fall $A \cdot B = B \cdot A$ als auch der Fall $A \cdot B \neq B \cdot A$ eintreten. Etwa für die Matrizen $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ und $B = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$ erhalten wir einerseits:

$$A \cdot B = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = B \cdot A$$

2 Matrixrechnung

Mit $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ und B wie oben ergibt sich jedoch:

$$A \cdot B = \begin{pmatrix} 1 & 3 \\ 3 & 7 \end{pmatrix} \neq \begin{pmatrix} 4 & 3 \\ 6 & 4 \end{pmatrix} = B \cdot A$$

Das Matrix-Produkt ist übrigens nur in den seltensten Fällen kommutativ und die Reihenfolge der Multiplikation daher wesentlich.

Doch was spricht nun eigentlich gegen die Einführung einer komponentenweisen Multiplikation? Zur Beantwortung gehen wir nachfolgend auf die Verbindung zwischen Matrizen und LGS ein und betrachten hierfür $A \in \text{Mat}(m \times n, \mathbb{R})$ sowie speziell $x \in \text{Mat}(n \times 1, \mathbb{R}) = \mathbb{R}^n$. Es gilt:

$$A \cdot x = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} A_{11} \cdot x_1 + A_{12} \cdot x_2 + \cdots + A_{1n} \cdot x_n \\ \vdots \\ A_{m1} \cdot x_1 + A_{m2} \cdot x_2 + \cdots + A_{mn} \cdot x_n \end{pmatrix}$$

Nun, die rechte Seite kennen wir bereits. Das LGS

$$\begin{cases} A_{11} \cdot x_1 + \cdots + A_{1n} \cdot x_n = b_1 \\ \vdots \\ A_{m1} \cdot x_1 + \cdots + A_{mn} \cdot x_n = b_m \end{cases} \quad (2.3)$$

ist äquivalent zur Matrix-Gleichung

$$A \cdot x = b,$$

wobei $A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix}$ die **Koeffizientenmatrix** des LGS (2.3) ist und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

sowie $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$.

Nur auf Grundlage der komplizierteren Definition der Matrix-Multiplikation wird diese ausdrucksstarke Verbindung möglich. Die Rechenregeln für Matrizen erleichtern uns den Umgang mit LGS erheblich, wie sich bereits im Beweis des nächsten Satz zeigen wird. Zunächst jedoch führen wir noch einen Bezeichner für die Menge aller Lösungen eines LGS ein.

Definition 2.1.11. Sei $A \in \text{Mat}(m \times n, \mathbb{R})$ und $b \in \mathbb{R}^m$. Dann heißt

$$\text{Lös}(A, b) := \{x \in \mathbb{R}^n \mid A \cdot x = b\}$$

die **Lösungsmenge** von (2.3).

Betrachten wir vorerst homogene LGS, also den Spezialfall $b = \mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ in folgendem

Satz 2.1.12

Die Lösungsmenge des homogenen LGS $A \cdot x = \mathbf{0}$, hat folgende Eigenschaften:

1. $\mathbf{0} \in \text{Lös}(A, \mathbf{0})$, d.h. insbesondere gilt $\text{Lös}(A, \mathbf{0}) \neq \emptyset$.
2. Sind $x, y \in \text{Lös}(A, \mathbf{0})$, so ist auch $x + y \in \text{Lös}(A, \mathbf{0})$.
3. Ist $x \in \text{Lös}(A, \mathbf{0})$ und $\lambda \in \mathbb{R}$, so ist auch $\lambda \cdot x \in \text{Lös}(A, \mathbf{0})$.

Beweis.

Zu 1.)

Sicherlich gilt $A \cdot \mathbf{0} = \mathbf{0}$ und daher ist $\mathbf{0} \in \text{Lös}(A, \mathbf{0})$.

Zu 2.)

Seien $x, y \in \text{Lös}(A, \mathbf{0})$, d.h. es gilt $Ax = \mathbf{0}$ und $Ay = \mathbf{0}$. Mit dem Resultat von Blatt 4, Aufgabe 3b erhalten wir dann:

$$\begin{aligned} A \cdot (x + y) &= A \cdot x + A \cdot y \\ &= \mathbf{0} + \mathbf{0} \\ &= \mathbf{0} \end{aligned}$$

Folglich ist $x + y \in \text{Lös}(A, \mathbf{0})$.

Zu 3.)

Sei schließlich $x \in \text{Lös}(A, \mathbf{0})$ und $\lambda \in \mathbb{R}$. Nun bedienen wir uns der zweiten Aussage von Satz 2.1.9 und bekommen:

$$\begin{aligned} A \cdot (\lambda \cdot x) &= \lambda \cdot (A \cdot x) \\ &= \lambda \cdot \mathbf{0} \\ &= \mathbf{0} \end{aligned}$$

Also gilt $\lambda \cdot x \in \text{Lös}(A, \mathbf{0})$. □

Für homogene und nur für homogene LGS treten Lösungsmengen (als Teilmengen von \mathbb{R}^n) mit den drei Eigenschaften des Satzes auf und verdienen eine eigene

2 Matrixrechnung

Definition 2.1.13. Eine Teilmenge $V \subset \mathbb{R}^n$ heißt **Untervektorraum**, falls gilt:

1. $\mathbf{0} \in V$
2. Sind $x, y \in V$, so ist auch $x + y \in V$.
3. Ist $x \in V$ und $\lambda \in \mathbb{R}$, so ist auch $\lambda \cdot x \in V$.

Ergo ist $\text{Lös}(A, \mathbf{0})$ stets ein Untervektorraum von \mathbb{R}^n .

Bemerkung 2.1.14. Das Aktienbeispiel führte auf ein inhomogenes System ohne Lösung, d.h. hier gilt $\text{Lös}(A, b) = \emptyset$ und somit ist wegen $\mathbf{0} \notin \text{Lös}(A, b)$ die Lösungsmenge des zugehörigen inhomogenen Systems kein Untervektorraum des \mathbb{R}^3 .

Wie bereits erwähnt, sind Lösungsmengen inhomogener Systeme *niemals* Untervektorräume des \mathbb{R}^n . Der Satz ist also für inhomogene Systeme ungültig, denn falls wann immer $\mathbf{0} \in \text{Lös}(A, b)$ ist, so folgt direkt $b = A \cdot \mathbf{0} = \mathbf{0}$, d.h. das LGS muss homogen sein.

Wohl aber lassen sich Lösungsmengen inhomogener Systeme auf Lösungsmengen homogener Systeme zurückführen.

Satz 2.1.15

Sei $A \in \text{Mat}(m \times n, \mathbb{R})$ und sei $b \in \mathbb{R}^m$. Weiterhin sei $\text{Lös}(A, b) \neq \emptyset$ und $y \in \text{Lös}(A, b)$. Dann gilt:

$$\text{Lös}(A, b) = \{x + y \mid x \in \text{Lös}(A, \mathbf{0})\}$$

Beweis.

Wie bei Mengengleichheiten üblich, zeigt man diese durch Nachweis der gegenseitigen Inklusionen, also in zwei Schritten. Wir beweisen:

$$\text{Lös}(A, b) \supset \{x + y \mid x \in \text{Lös}(A, \mathbf{0})\} \text{ sowie } \text{Lös}(A, b) \subset \{x + y \mid x \in \text{Lös}(A, \mathbf{0})\}$$

„ \supset “:

Sei $x \in \text{Lös}(A, \mathbf{0})$ und $y \in \text{Lös}(A, b)$. Wir überprüfen, ob $x + y$ das inhomogene System löst. Es gilt

$$\begin{aligned} A \cdot (x + y) &= A \cdot x + A \cdot y \\ &= \mathbf{0} + b \\ &= b \end{aligned}$$

und somit $x + y \in \text{Lös}(A, b)$.

„ \subset “:

Sei nun $z \in \text{Lös}(A, b)$. Wir setzen $x := z - y$, denn dann gilt $z = x + y$. Nun ist zu zeigen, dass tatsächlich $x \in \text{Lös}(A, \mathbf{0})$ gilt. Betrachte dazu:

$$\begin{aligned} A \cdot x &= A \cdot (z - y) \\ &= A \cdot (z + (-1) \cdot y) \\ &= A \cdot z + A \cdot (-1) \cdot y \\ &= A \cdot z + (-1) \cdot A \cdot y \\ &= b + (-1) \cdot b \\ &= \mathbf{0} \end{aligned}$$

Das heißt nun aber gerade $x \in \text{Lös}(A, \mathbf{0})$ und somit gilt $z \in \{x + y \mid x \in \text{Lös}(A, \mathbf{0})\}$. \square

Im folgenden Beispiel schießen einmal „mit Kanonen auf Spatzen“ und wenden die Erkenntnis des Satzes auf ein sehr übersichtliches LGS an:

Beispiel 2.1.16. Für $x_1 + 3x_2 = 7$ haben wir die Koeffizientenmatrix $\begin{pmatrix} 1 & 3 \end{pmatrix}$ sowie $b = (7)$.

Wir erraten die Lösung $y = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ dieses inhomogenen LGS und brauchen nun dem Satz nach nur noch das zugehörige homogene LGS $x_1 + 3x_2 = 0$ zu lösen. Es ist:

$$\begin{aligned} \text{Lös}(A, \mathbf{0}) &= \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2 \mid x_1 + 3x_2 = 0 \right\} \\ &= \left\{ \begin{pmatrix} -3x_2 \\ x_2 \end{pmatrix} \mid x_2 \in \mathbb{R} \right\} \\ &= \left\{ x_2 \cdot \begin{pmatrix} -3 \\ 1 \end{pmatrix} \mid x_2 \in \mathbb{R} \right\} \\ &= \left\{ t \cdot \begin{pmatrix} -3 \\ 1 \end{pmatrix} \mid t \in \mathbb{R} \right\} \end{aligned}$$

Weiter nach dem Satz genügt es dann, y hinzuzuzählen. Wir erhalten als Lösungsmenge:

$$\begin{aligned} \text{Lös}(A, b) &= \left\{ t \cdot \begin{pmatrix} -3 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} \mid t \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} 1 - 3t \\ 2 + t \end{pmatrix} \mid t \in \mathbb{R} \right\} \end{aligned}$$

Nachfolgend richten wir unser Augenmerk besonders auf die Lösungsmengen homogener LGS, d.h. auf Untervektorräume des \mathbb{R}^n und deren sinnvolle bzw. effektive Beschreibung. Der Folgeabschnitt behandelt die hierfür grundlegenden Konzepte.

2.2 Lineare Unabhängigkeit und Basen

Mittels der Anwendung sämtlicher zugelassener Grundrechenarten für Vektoren, d.h. mittels Vektoraddition und skalarer Multiplikation auf Vektoren v_1, \dots, v_m , lässt sich folgende Menge erzeugen.

Definition 2.2.1. Für $v_1, \dots, v_m \in \mathbb{R}^n$ heißt

$$L(v_1, \dots, v_m) := \left\{ \sum_{k=1}^m \lambda_k \cdot v_k \mid \lambda_1, \dots, \lambda_m \in \mathbb{R} \right\}$$

die **lineare Hülle** von v_1, \dots, v_m .

Dabei heißt ein Vektor der Form $\lambda_1 \cdot v_1 + \dots + \lambda_m \cdot v_m$, also ein Element von $L(v_1, \dots, v_m)$, **Linearkombination** von v_1, \dots, v_m . Folglich ist $L(v_1, \dots, v_m)$ die Menge aller Linearkombinationen von v_1, \dots, v_m .

Betrachten wir einige Beispiele für lineare Hüllen in dem Spezialfall von Vektoren im \mathbb{R}^3 . Es seien dazu:

$$v_0 := \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad v_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 := \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad v_4 := \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Beispiel 2.2.2. Für $v = \mathbf{0} \in \mathbb{R}^n$ bekommen wir: $L(v) = \{\lambda \cdot \mathbf{0} \mid \lambda \in \mathbb{R}\} = \{\mathbf{0}\}$. Insbesondere gilt somit:

$$L(v_0) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

Beispiel 2.2.3. Berechnen wir die lineare Hülle von v_1 :

$$\begin{aligned} L(v_1) &= \left\{ \lambda \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} \lambda \\ 0 \\ 0 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \end{aligned}$$

In geometrischer Deutung ist $L(v_1)$ genau diejenige Ursprungsgerade, welche der x_1 -Achse entspricht:

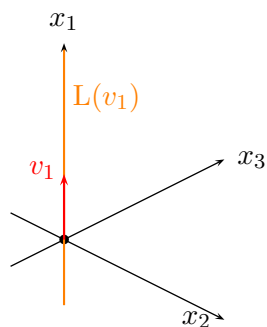


Abb. 10

Beispiel 2.2.4. Die Menge aller Linearkombinationen von v_2 und v_3 ist gegeben durch:

$$\begin{aligned} L(v_2, v_3) &= \{s \cdot v_2 + t \cdot v_3 \mid s, t \in \mathbb{R}\} \\ &= \left\{ \begin{pmatrix} 0 \\ s \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix} \mid s, t \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} 0 \\ s \\ t \end{pmatrix} \mid s, t \in \mathbb{R} \right\} \end{aligned}$$

Veranschaulichen wir diese Teilmenge des \mathbb{R}^3 , so ergibt sich genau die x_2 - x_3 -Ebene:

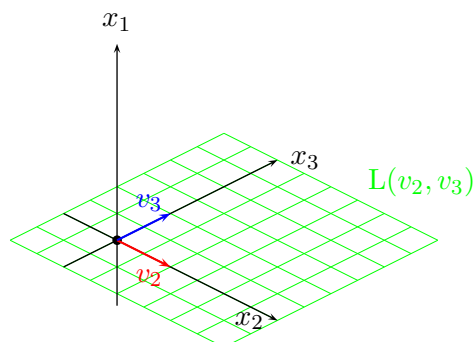


Abb. 11

Beispiel 2.2.5. Nimmt man nun zu v_2 und v_3 noch den Vektor v_4 hinzu, so erhält man:

$$\begin{aligned} L(v_2, v_3, v_4) &= \{\lambda \cdot v_2 + \mu \cdot v_3 + \eta \cdot v_4 \mid \lambda, \mu, \eta \in \mathbb{R}\} \\ &= \left\{ \begin{pmatrix} 0 \\ \lambda \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \mu \end{pmatrix} + \begin{pmatrix} 0 \\ \eta \\ \eta \end{pmatrix} \mid \lambda, \mu, \eta \in \mathbb{R} \right\} \end{aligned}$$

2 Matrixrechnung

$$\begin{aligned}
 &= \left\{ \begin{pmatrix} 0 \\ \lambda + \eta \\ \mu + \eta \end{pmatrix} \mid \lambda, \mu, \eta \in \mathbb{R} \right\} \\
 &= L(v_2, v_3)
 \end{aligned}$$

Die Darstellung $L(v_2, v_3, v_4)$, die durch Hinzunahme von v_4 entsteht, ist offenbar weniger effizient als die Darstellung $L(v_2, v_3)$. Demnach ist v_4 sozusagen überflüssig und die geometrische Deutung offenbart in der Tat:

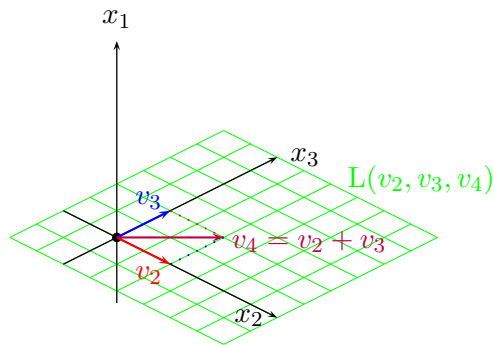


Abb. 12

Beispiel 2.2.6. Bilden wir schließlich noch die lineare Hülle von v_1, v_2 und v_3 :

$$L(v_1, v_2, v_3) = \mathbb{R}^3.$$

Im Beispiel (2.2.5) erhielten wir für die lineare Hülle dreier Vektoren eine Ebene. Für einen leicht abgeänderten Satz von Vektoren, wie im Beispiel (2.2.6), bekamen wir dagegen genau den \mathbb{R}^3 . Unsere Zielstellung ist es nun, die Systematik hinter diesem Phänomen zu verstehen.

Bemerkung 2.2.7. Stets sind v_1, \dots, v_m wieder Elemente ihrer linearen Hülle, d.h. es gilt $v_1, \dots, v_m \in L(v_1, \dots, v_m)$. Um dies einzusehen, schreiben wir einfach

$$v_1 = 1 \cdot v_1 + 0 \cdot v_2 + \dots + 0 \cdot v_m \in L(v_1, \dots, v_m)$$

und analog für v_2, \dots, v_m .

Des Weiteren gilt allgemein:

$$L(v_1, \dots, v_m) \subset L(v_1, \dots, v_m, v_{m+1}, \dots, v_{m'})$$

Denn:

$$\lambda_1 \cdot v_1 + \dots + \lambda_m \cdot v_m = \lambda_1 \cdot v_1 + \dots + \lambda_m \cdot v_m + 0 \cdot v_{m+1} + \dots + 0 \cdot v_{m'}$$

Satz 2.2.8

Für $v_1, \dots, v_m \in \mathbb{R}^n$ ist $L(v_1, \dots, v_m)$ ein Untervektorraum von \mathbb{R}^n .

Beweis.

Prüfen wir die Eigenschaften eines Untervektorraumes Schritt für Schritt nach.

1.)

Wir schreiben und erhalten:

$$\mathbf{0} = 0 \cdot v_1 + 0 \cdot v_2 + \dots + 0 \cdot v_m \in L(v_1, \dots, v_m).$$

2.)

Seien $x, y \in L(v_1, \dots, v_m)$. Wir zeigen: $x + y \in L(v_1, \dots, v_m)$. Für $\lambda_k, \mu_k \in \mathbb{R}$ haben wir

$$\begin{aligned} x &= \lambda_1 \cdot v_1 + \dots + \lambda_m \cdot v_m && \text{sowie} \\ y &= \mu_1 \cdot v_1 + \dots + \mu_m \cdot v_m \end{aligned}$$

und bekommen damit:

$$\begin{aligned} x + y &= \lambda_1 \cdot v_1 + \dots + \lambda_m \cdot v_m + \mu_1 \cdot v_1 + \dots + \mu_m \cdot v_m \\ &= (\lambda_1 + \mu_1) \cdot v_1 + \dots + (\lambda_m + \mu_m) \cdot v_m \\ &\in L(v_1, \dots, v_m) \end{aligned}$$

3.)

Schließlich prüfen wir für $x \in L(v_1, \dots, v_m)$ und $\lambda \in \mathbb{R}$, dass stets $\lambda \cdot x \in L(v_1, \dots, v_m)$ gilt. Wir notieren

$$x = \lambda_1 \cdot v_1 + \dots + \lambda_m \cdot v_m$$

und erhalten:

$$\begin{aligned} \lambda \cdot x &= \lambda \cdot (\lambda_1 \cdot v_1 + \dots + \lambda_m \cdot v_m) \\ &= (\lambda \cdot \lambda_1) \cdot v_1 + \dots + (\lambda \cdot \lambda_m) \cdot v_m \\ &\in L(v_1, \dots, v_m) \end{aligned}$$

□

Untervektorräume treten also auch als lineare Hüllen auf.

Wie angekündigt, wollen wir nun verstehen, weshalb sich für die lineare Hülle dreier Vektoren in den Beispielen (2.2.5) und (2.2.6) einerseits etwas zweidimensionales (d.h. eine Ebene) und andererseits etwas Dreidimensionales (der \mathbb{R}^3) ergab. Den Schlüssel hierfür bildet folgendes Lemma (ein mathematischer Hilfssatz).

2 Matrixrechnung

Lemma 2.2.9

Seien $v_1, \dots, v_m \in \mathbb{R}^n$. Dann sind folgende Aussagen äquivalent:

1. $\exists \lambda_1, \dots, \lambda_m \in \mathbb{R}$, nicht alle gleich 0, so dass: $\lambda_1 v_1 + \dots + \lambda_m v_m = \mathbf{0}$
(Dies nennt man dann eine nichttriviale Linearkombination des Nullvektors.)
2. $\exists l \in \{1, \dots, m\} : v_l \in L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m)$
3. $\exists l \in \{1, \dots, m\} : L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m) = L(v_1, \dots, v_m)$

Beweis.

Für die behauptete Äquivalenz der drei Aussagen genügt der Nachweis folgender Beweisteile:

(1.) \Leftrightarrow (2.) (also die Richtungen 1. \Rightarrow 2. und 2. \Rightarrow 1.) sowie

(2.) \Leftrightarrow (3.) (also die Richtungen 2. \Rightarrow 3. und 3. \Rightarrow 2.)

„(1.) \Rightarrow (2.)“:

Seien $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ und davon wenigstens ein $\lambda_l \neq 0$, so dass $\lambda_1 v_1 + \dots + \lambda_m v_m = \mathbf{0}$. Ist also etwa $\lambda_l \neq 0$, so gilt:

$$\begin{aligned} -\lambda_l v_l &= \lambda_1 v_1 + \dots + \lambda_{l-1} v_{l-1} + \lambda_{l+1} v_{l+1} + \dots + \lambda_m v_m \\ \Rightarrow v_l &= -\frac{\lambda_1}{\lambda_l} v_1 - \dots - \frac{\lambda_{l-1}}{\lambda_l} v_{l-1} - \frac{\lambda_{l+1}}{\lambda_l} v_{l+1} - \dots - \frac{\lambda_m}{\lambda_l} v_m \\ &\in L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m) \end{aligned}$$

„(2.) \Rightarrow (1.)“:

Sei nun umgekehrt $v_l \in L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m)$, d.h. es existieren $\lambda_1, \dots, \lambda_{l-1}, \lambda_{l+1}, \dots, \lambda_m \in \mathbb{R}$, so dass:

$$\begin{aligned} v_l &= \lambda_1 v_1 + \dots + \lambda_{l-1} v_{l-1} + \lambda_{l+1} v_{l+1} + \dots + \lambda_m v_m \\ \Rightarrow 0 &= \lambda_1 v_1 + \dots + \lambda_{l-1} v_{l-1} - v_l + \lambda_{l+1} v_{l+1} + \dots + \lambda_m v_m \end{aligned}$$

Der Koeffizient von v_l ist offenbar gegeben durch $-1 \neq 0$ und damit haben wir eine nichttriviale Linearkombination des Nullvektors.

„(2.) \Rightarrow (3.)“:

Sei $v_l \in L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m)$.

Wir zeigen $L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m) = L(v_1, \dots, v_m)$ und im Prinzip müssten wir uns dafür zweier Inklusionen annehmen. Nach Bemerkung (2.2.7) ist jedoch „ \subset “ bereits klar.

Wir zeigen noch $L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m) \supset L(v_1, \dots, v_m)$, also

„ \supset “:

Sei hierfür $x \in L(v_1, \dots, v_m)$. Dann existieren $\lambda_k \in \mathbb{R}$, so dass:

$$x = \lambda_1 v_1 + \dots + \lambda_m v_m$$

Da nach Voraussetzung $v_l \in L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m)$ ist, können wir schreiben:

$$v_l = \mu_1 v_1 + \dots + \mu_{l-1} v_{l-1} + \mu_{l+1} v_{l+1} + \dots + \mu_m v_m$$

schreiben. Diesen Ausdruck setzen wir in die Darstellung für x ein und sortieren geeignet:

$$\begin{aligned} x &= \lambda_1 v_1 + \dots + \lambda_{l-1} v_{l-1} + \lambda_l v_l + \lambda_{l+1} v_{l+1} + \dots + \lambda_m v_m \\ &= \lambda_1 v_1 + \dots + \lambda_{l-1} v_{l-1} + \lambda_l (\mu_1 v_1 + \dots + \mu_{l-1} v_{l-1} + \mu_{l+1} v_{l+1} + \dots + \mu_m v_m) \\ &\quad + \lambda_{l+1} v_{l+1} + \dots + \lambda_m v_m \\ &= (\lambda_1 + \lambda_l \mu_1) v_1 + \dots + (\lambda_{l-1} + \lambda_l \mu_{l-1}) v_{l-1} \\ &\quad + (\lambda_{l+1} + \lambda_l \mu_{l+1}) v_{l+1} + \dots + (\lambda_m + \lambda_l \mu_m) v_m \\ &\in L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m) \end{aligned}$$

Es folgt $L(v_1, \dots, v_m) \subset L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m)$.

„(3.) \Rightarrow (2.)“:

Offenbar ist $v_l \in L(v_1, \dots, v_m) = L(v_1, \dots, v_{l-1}, v_{l+1}, \dots, v_m)$. □

Definition 2.2.10. Gilt eine (und damit alle) der Aussagen aus dem Lemma, dann heißen die Vektoren v_1, \dots, v_m **linear abhängig**, ansonsten **linear unabhängig**.

Oft besteht die Aufgabe darin, die lineare Unabhängigkeit endlich vieler vorgegebener Vektoren $v_1, \dots, v_m \in \mathbb{R}^n$ nachzuweisen. In diesem Fall muss für alle $\lambda_k \in \mathbb{R}$ die Gültigkeit folgender Implikation nachgeprüft werden:

$$\lambda_1 v_1 + \dots + \lambda_m v_m = \mathbf{0} \Rightarrow \lambda_1 = \dots = \lambda_m = 0$$

Der Nullvektor darf sich also nur auf triviale Art linear kombinieren lassen. Sehen wir uns einige Beispiele vorgegebener linear (un-)abhängiger Vektoren an.

Beispiel 2.2.11. In welchem Fall ist $v \in \mathbb{R}^n$ linear (un-)abhängig?

1. Fall: Sei $v = \mathbf{0}$. Dann gilt:

$$\forall \lambda \in \mathbb{R} \setminus \{0\} : \lambda \cdot v = \mathbf{0}$$

Insbesondere existiert somit ein $\lambda \neq 0$, so dass $\lambda \cdot v = \mathbf{0}$ ist. Der Nullvektor ist also linear abhängig.

2. Fall: Sei $v \neq \mathbf{0}$. Nun existiert kein $\lambda \in \mathbb{R} \setminus \{0\}$, so dass $\lambda \cdot v = \mathbf{0}$, d.h. es gilt:

$$\forall \lambda \in \mathbb{R} \setminus \{0\} : \lambda \cdot v \neq \mathbf{0}$$

Somit ist jeder Vektor $v \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ linear unabhängig.

2 Matrixrechnung

Beispiel 2.2.12. Wann sind $v_1, v_2 \in \mathbb{R}^n$ linear abhängig? Laut Definition gilt unter anderem:

$$\begin{aligned}v_1, v_2 \text{ sind linear abhängig} &\Leftrightarrow v_1 \in L(v_2) \text{ oder } v_2 \in L(v_1) \\ &\Leftrightarrow v_1 \in \{t \cdot v_2 \mid t \in \mathbb{R}\} \text{ oder } v_2 \in \{s \cdot v_1 \mid s \in \mathbb{R}\} \\ &\Leftrightarrow \exists t \in \mathbb{R} : v_1 = t \cdot v_2 \text{ oder } v_2 = t \cdot v_1\end{aligned}$$

Zwei Vektoren sind also genau dann linear abhängig, wenn einer sich als Vielfaches des anderen schreiben lässt.

Als Notation für $X = \{v_1, \dots, v_m\}$ mit $v_1, \dots, v_m \in \mathbb{R}^n$ schreiben wir einerseits $L(X)$ abkürzend für $L(v_1, \dots, v_m)$. Als Konvention vereinbaren wir andererseits $L(\emptyset) = \{0\}$. Zudem soll \emptyset linear unabhängig sein.

Mittels der Konzepte von linearer Unabhängigkeit und Hüllenbildung erschließen sich uns nunmehr die Grundbegriffe Basis und Dimension.

Definition 2.2.13. Sei $V \subset \mathbb{R}^n$ ein Untervektorraum und seien $v_1, \dots, v_m \in V$. Dann heißt $X = \{v_1, \dots, v_m\}$ **Basis** von V , falls zugleich gilt:

1. v_1, \dots, v_m sind linear unabhängig.
2. $L(X) = V$.

Diese Definition mag zunächst abstrakt erscheinen, weshalb wir Anschauung in einigen Beispielen suchen.

Beispiel 2.2.14. Zu $V = \{0\}$ ist \emptyset eine Basis von V .

(Und nicht etwa $\{0\}$, denn wie wir wissen, ist der Nullvektor nicht linear unabhängig!)

Beispiel 2.2.15. Für $V = \left\{ \begin{pmatrix} t \\ 0 \\ 0 \end{pmatrix} \mid t \in \mathbb{R} \right\} \subset \mathbb{R}^3$ ist z.B. $X_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$ eine Basis, ebenso

$X_2 = \left\{ \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \right\}$. Überhaupt liefert jede reelle Zahl ungleich 0 in der ersten Komponente eine Basis von V .

Beispiel 2.2.16. Der Untervektorraum $V = \left\{ \begin{pmatrix} 0 \\ \lambda_1 \\ \lambda_2 \end{pmatrix} \mid \lambda_1, \lambda_2 \in \mathbb{R} \right\} \subset \mathbb{R}^3$ besitzt zum Beispiel

die Basis $X = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$. Hingegen ist die Menge $Y = \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$ keine Basis von V , da die zweite definierende Eigenschaft einer Basis nicht erfüllt ist. Genauer gesagt gilt einerseits

$$L\left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right) = \left\{ \lambda \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \subset V,$$

andererseits ist aber z.B.

$$\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \in V \setminus L\left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right).$$

Beispiel 2.2.17. Der Anschauungsraum $V = \mathbb{R}^3$ besitzt z.B. die Basis

$$X = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Diese Menge X wird auch als „Standard-Basis“ und die Basisvektoren mit e_1, e_2, e_3 bezeichnet.

Bemerkung 2.2.18. Jeder Untervektorraum $V \subset \mathbb{R}^n$ besitzt (mindestens) eine Basis.

Fraglich bleibt, wie man sich im Allgemeinen eine Basis verschafft.

Nachfolgend lernen wir eine Abfolge kennen, die uns in die Lage versetzt, zu einem vorgegebenen Untervektorraum $V \subseteq \mathbb{R}^n$ in endlich vielen Schritten eine Basis zu konstruieren.

Algorithmus zur Bestimmung einer Basis.

0. Ist $\{0\} = V$, so ist \emptyset eine Basis und man ist fertig.
1. Ist $\{0\} \neq V$, so wähle $v_1 \in V \setminus \{0\}$ und bilde $L(v_1)$.
Ist $L(v_1) = V$, so ist $\{v_1\}$ eine Basis und das war's.
2. Ist $L(v_1) \neq V$, so wähle $v_2 \in V \setminus L(v_1)$ und bilde $L(v_1, v_2)$.
Ist $L(v_1, v_2) = V$, so ist $\{v_1, v_2\}$ eine Basis und die Suche ist beendet.
3. Ist $L(v_1, v_2) \neq V$, so wähle $v_3 \in V \setminus L(v_1, v_2)$ und bilde $L(v_1, v_2, v_3)$.
...

2 Matrixrechnung

Das Verfahren endet spätestens nach dem Schritt n . Etwa im Fall \mathbb{R}^3 endet der Algorithmus also nach maximal 3 Schritten. Betrachten wir hierzu ein Beispiel und bestimmen einmal eine konkrete Basis zu einer vorgegebenen Lösungsmenge eines LGS, d.h. eines Untervektorraums $V \subset \mathbb{R}^3$.

Beispiel 2.2.19. Wir ermitteln eine Basis von $V = \{x \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$.

Ist $\{0\} = V$? Nein, denn wähle z.B. $v_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \in V \setminus \{0\}$. Damit ergibt sich:

$$L(v_1) = \left\{ t \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \mid t \in \mathbb{R} \right\}$$

Ist $L(v_1) = V$? Nein, denn wähle z.B. $v_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \in V \setminus \{v_1\}$. Es ist:

$$\begin{aligned} L(v_1, v_2) &= \left\{ t \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + s \cdot \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \mid t, s \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} t - 2s \\ -2t + s \\ t + s \end{pmatrix} \mid t, s \in \mathbb{R} \right\} \end{aligned}$$

Addition der zweiten und dritten Komponente sowie Multiplikation des Resultats mit -1 liefern:

$$\begin{aligned} L(v_1, v_2) &= \left\{ \begin{pmatrix} -(x_2 + x_3) \\ x_2 \\ x_3 \end{pmatrix} \mid x_2, x_3 \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1, x_2, x_3 \in \mathbb{R} \ \& \ x_1 = -(x_2 + x_3) \right\} \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1, x_2, x_3 \in \mathbb{R} \ \& \ x_1 + x_2 + x_3 = 0 \right\} \\ &= V \end{aligned}$$

Somit ist $\{v_1, v_2\} = \left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right\}$ eine Basis von V .

Bemerkung 2.2.20. Wie schon im Beispiel (2.2.15) festgestellt, sind Basen für Untervektorräume $V \subset \mathbb{R}^n$ im Allgemeinen nicht eindeutig. Wohl aber besitzen zwei Basen von V stets gleich viele Elemente, d.h. die so genannte Länge der Basis ist für alle Basen von V identisch.

Beispiel 2.2.21. $V = \mathbb{R}^2$ hat unter anderem die Basen

$$X_1 = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \quad \text{und} \quad X_2 = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}.$$

Definition 2.2.22. Sei $V \subset \mathbb{R}^n$ ein Untervektorraum und X eine Basis von V . Dann heißt die Zahl

$$\dim(V) := \#X$$

die **Dimension** von V .

Diese Definition ist nur auf Grund von Bemerkung (2.2.20) sinnvoll. Besäße ein Untervektorraum Basen verschiedener Länge, so wäre die Dimension nicht eindeutig.

Überlegen wir auch hierzu einige Beispiele.

Beispiel 2.2.23. Für $V = \mathbb{R}^n$ ist $X = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$ eine Basis der Länge n .

Damit besitzen alle Basen von X die Länge n und es gilt:

$$\dim(\mathbb{R}^n) = \#X = n$$

Obige Basis ist die Standard-Basis von $V = \mathbb{R}^n$. Ihre Elemente bezeichnen wir mit e_1, e_2, \dots, e_n .

Beispiel 2.2.24. Für den Nullraum $V = \{0\}$ ist die Dimension 0, da wir \emptyset als Basis von $\{0\}$ festgelegt haben. Somit ist:

$$\dim(\{0\}) = \#\emptyset = 0$$

Beispiel 2.2.25. Im Beispiel (2.2.19) ermittelten für $V = \{x \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$ eine Basis der Länge 2. Folglich gilt auch:

$$\dim(V) = 2$$

2 Matrixrechnung

Beispiel 2.2.26. Betrachten wir den Untervektorraum:

$$V = \{x \in \mathbb{R}^n \mid x_{m+1} = x_{m+2} = \dots = x_n = 0\}$$

Dessen Elemente besitzen die Gestalt

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \\ x_{m+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

mit $\lambda_1, \dots, \lambda_m \in \mathbb{R}$. Offenbar bilden dann die Vektoren $e_1, \dots, e_m \in V$ gegeben durch

$$\begin{pmatrix} 1 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{l} \leftarrow m\text{-te Komponente} \\ \leftarrow (m+1)\text{-te Komponente} \end{array}$$

eine Basis von V . Daher gilt:

$$\dim(V) = m$$

Beispiel 2.2.27. Für den Spezialfall \mathbb{R}^3 ergeben die Untervektorräume

$$V = \{x \in \mathbb{R}^3 \mid x_3 = 0\} \quad \text{bzw.} \quad U = \{x \in \mathbb{R}^3 \mid x_2 = x_3 = 0\}$$

die räumlichen geometrischen Interpretationen der x_1 - x_2 -Ebene (im Fall von V) bzw. der x_1 -Achse (im Fall von U) und lassen somit unmittelbar auf ihre Dimensionen

$$\dim(V) = 2 \quad \text{bzw.} \quad \dim(U) = 1$$

schließen.

Nun zu einem Hilfssatz, der das Dimensionsverhalten zweier Untervektorräume $U, V \subset \mathbb{R}^n$ im Falle der Inklusion $U \subset V$ klärt.

Lemma 2.2.28

Seien $U, V \subset \mathbb{R}^n$ Untervektorräume.

1. Ist $U \subset V$, so gilt:

$$\dim(U) \leq \dim(V)$$

2. Ist $U \subset V$ und ferner $\dim(U) = \dim(V)$, so gilt:

$$U = V.$$

Beweis.

Zu 1.)

Zunächst konstruieren wir wie in Bemerkung (2.2.18) eine Basis $X = \{v_1, \dots, v_m\}$ von U . Das weitere Vorgehen setzt im Prinzip den Algorithmus zur Bestimmung einer Basis fort:

Ist $U = L(X) = V$, so ist X ebenso Basis von V und es gilt $\dim(U) = \dim(V) = m$.

Ist $U = L(X) \neq V$, so wähle $v_{m+1} \in V \setminus \{v_1, \dots, v_m\}$.

Ist $L(v_1, \dots, v_m, v_{m+1}) = V$, so ist $Y = \{v_1, \dots, v_m, v_{m+1}\}$ eine Basis von V .

Ist dagegen $L(Y) \neq V$, so wähle $v_{m+2} \in V \setminus L(Y)$.

...

Somit folgt $\dim(U) \leq \dim(V)$.

Zu 2.)

Wie eben konstruieren wir eine Basis X von U und eine Basis Y von V . Nach Konstruktion und wegen $U \subset V$ gilt dann insbesondere $X \subset Y$. Ferner ist nach Voraussetzung $\#X = \dim(U) = \dim(V) = \#Y$ und damit $X = Y$. Nach Übergang zur linearen Hülle hat die letzte Gleichung schließlich $U = L(X) = L(Y) = V$ zur Folge. \square

Beispiel 2.2.29. Für $U = \{x \in \mathbb{R}^3 \mid x_1 = 0\}$ und $V = \{x \in \mathbb{R}^3 \mid x_2 = 0\}$ gilt $\dim(U) = \dim(V) = 2$, jedoch ist $U \neq V$, da weder $U \subset V$ noch $V \subset U$ zutrifft.

Wir beschließen diesen Abschnitt mit einem nützlichen Hilfssatz, mit dem ein vorgegebener Satz linear unabhängiger Vektoren eines Untervektorraums U zu einer Basis von U ergänzt werden kann.

Lemma 2.2.30 (Basisergänzungssatz)

Ist $U \subset \mathbb{R}^n$ ein Untervektorraum und sind $v_1, \dots, v_m \in U$ linear unabhängig, so gibt es eine Basis X von U mit $v_1, \dots, v_m \in X$.

2 Matrixrechnung

Beweis.

Wieder erfolgt die Konstruktion in Analogie zu Bemerkung (2.2.18).

Ist $L(v_1, \dots, v_m) = U$, so ist $X = \{v_1, \dots, v_m\}$ eine Basis von U .

Ist $L(v_1, \dots, v_m) \neq U$, so wähle $v_{m+1} \in U \setminus L(v_1, \dots, v_m)$.

...

(weiter wie im Algorithmus) □

Nachdem sich unsere Aufmerksamkeit nun einige Zeit fast ausschließlich auf Untervektorräume, deren Basen und den Dimensionsbegriff gerichtet hat, erinnern wir uns an dessen Ausgangspunkt: die Lösungsmengen linearer Gleichungssysteme. Im folgenden Abschnitt konzentrieren wir uns darauf, LGS in eine optimale Form zu transformieren.

2.3 Der Gauß-Algorithmus

Dem Gauß'schen Lösungsalgorithmus¹ für LGS liegen Überlegungen hinsichtlich Transformationen linearer Gleichungssysteme zu Grunde, bei deren Anwendung die Lösungsmenge nicht verändert wird:

- Multiplikation einer Gleichung mit $\lambda \in \mathbb{R} \setminus \{0\}$:

$$\begin{aligned} A_{i1}x_1 + \dots + A_{in}x_n &= b_i \\ \Leftrightarrow^{\lambda \neq 0} \lambda(A_{i1}x_1 + \dots + A_{in}x_n) &= \lambda b_i \\ \Leftrightarrow (\lambda A_{i1})x_1 + \dots + (\lambda A_{in})x_n &= \lambda b_i \end{aligned}$$

- Vertauschen zweier Gleichungen
- Addition des Vielfachen einer Gleichung zu einer anderen Gleichung:

$$\Leftrightarrow \begin{cases} A_{i1}x_1 + \dots + A_{in}x_n = b_i \\ A_{j1}x_1 + \dots + A_{jn}x_n = b_j \end{cases} \Leftrightarrow \begin{cases} A_{i1}x_1 + \dots + A_{in}x_n = b_i \\ (A_{j1} + \lambda A_{i1})x_1 + \dots + (A_{jn} + \lambda A_{in})x_n = b_j + \lambda b_i \end{cases}$$

Für ein LGS

$$\begin{cases} A_{11}x_1 + \dots + A_{1n}x_n = b_1 \\ \vdots \\ A_{m1}x_1 + \dots + A_{mn}x_n = b_n \end{cases}$$

haben wir die kompakte Schreibweise

$$Ax = b$$

¹benannt nach dem berühmten Mathematiker Carl Friedrich Gauß (1777 (Braunschweig) - 1855 (Göttingen)), siehe http://de.wikipedia.org/wiki/Carl_Friedrich_Gauß

kennen gelernt, wobei wir A als Koeffizientenmatrix bezeichnen. Nun lassen sich auch A und b in einer Matrix gruppieren. Das Resultat ist die so genannte **erweiterte Koeffizientenmatrix** von $Ax = b$, nämlich:

$$(A, b) := \begin{pmatrix} A_{11} & \cdots & A_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ A_{m1} & \cdots & A_{mn} & b_m \end{pmatrix}$$

Definition 2.3.1. Folgende Transformationen einer Matrix bezeichnet man als **elementare Zeilenumformungen**:

- Multiplikation einer Zeile mit $\lambda \in \mathbb{R} \setminus \{0\}$
- Vertauschen zweier Zeilen
- Addition des Vielfachen einer Zeile zu einer anderen Zeile

Bemerkung 2.3.2. Die Lösungsmenge eines LGS bleibt bei Anwendung elementarer Zeilenumformungen auf die erweiterte Koeffizientenmatrix unverändert.

Achtung: Spaltenumformungen sind dagegen im Allgemeinen unzulässig!

Algorithmus zur Transformation einer Matrix in spezielle Zeilenstufenform.

1. Vertausche die Zeilen so, dass in der ersten Zeile der erste Eintrag $\neq 0$ nicht weiter rechts steht, als in allen anderen Zeilen.
2. Multipliziere die Zeilen mit geeigneten $\lambda \in \mathbb{R} \setminus \{0\}$ so, dass in den Zeilen, bei denen der erste Eintrag $\neq 0$ an der gleichen Stelle wie in der ersten Zeile steht, dieser Eintrag 1 wird. (Dies schließt auch die erste Zeile ein!)
3. Subtrahiere die erste Zeile von allen anderen Zeilen, in denen der erste Eintrag $\neq 0$ an derselben Stelle steht wie in der 1. Zeile.
4. Wende die Schritte 1. - 3. auf diejenige Teilmatrix an, die durch Streichung der ersten Zeile entsteht.

Beispiel 2.3.3. Wir führen den Algorithmus einmal an der Matrix

$$\begin{pmatrix} 0 & 1 & 1 \\ 5 & 10 & -20 \\ 2 & 8 & 4 \end{pmatrix}$$

2 Matrixrechnung

durch und verbinden dabei die Transformationsschritte mit dem Symbol „ \rightsquigarrow “:

$$\begin{aligned}
 \begin{pmatrix} 0 & 1 & 1 \\ 5 & 10 & -20 \\ 2 & 8 & 4 \end{pmatrix} &\stackrel{1.}{\rightsquigarrow} \begin{pmatrix} 5 & 10 & -20 \\ 0 & 1 & 1 \\ 2 & 8 & 4 \end{pmatrix} \stackrel{2.}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -4 \\ 0 & 1 & 1 \\ 2 & 8 & 4 \end{pmatrix} \stackrel{2.}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -4 \\ 0 & 1 & 1 \\ 1 & 4 & 2 \end{pmatrix} \\
 &\stackrel{3.}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -4 \\ 0 & 1 & 1 \\ 0 & 2 & 6 \end{pmatrix} \stackrel{4./2.}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -4 \\ 0 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix} \stackrel{4./3.}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -4 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} \\
 &\stackrel{4./2.}{\rightsquigarrow} \begin{pmatrix} 1 & 2 & -4 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}
 \end{aligned}$$

Mit Hilfe dieses Algorithmus' gelingt uns also die Transformation jeder Matrix $A \in \text{Mat}(m \times n, \mathbb{R})$ in eine spezielle Zeilenstufenform. Geben wir noch eine

Definition 2.3.4. Eine Matrix $A \in \text{Mat}(m \times n, \mathbb{R})$ hat genau dann **Zeilenstufenform**, wenn für alle $i \in \{2, \dots, m\}$ gilt: Sind die ersten $(i - 1)$ -ten Einträge der $(i - 1)$ -ten Zeile = 0, so sind auch die ersten i Einträge der i -ten Zeile = 0.
 Ferner hat A **spezielle Zeilenstufenform**, falls zusätzlich für alle $i \in \{1, \dots, m\}$ gilt: Sind $A_{i1} = A_{i2} = \dots = A_{ij} = 0$, und ist $A_{ij+1} \neq 0$, so ist $A_{ij+1} = 1$.

Wir machen uns klar, was diese Definition genau bedeutet:

- **Zeilenstufenform:**

Jede Zeile der Matrix hat mindestens eine führende 0 mehr als die vorherige,

d.h. etwa für $\begin{pmatrix} 1 & 3 & -7 & 5 \\ * & * & * & * \\ * & * & * & * \end{pmatrix}$ wird $\begin{pmatrix} 1 & 3 & -7 & 5 \\ 0 & * & * & * \\ 0 & 0 & * & * \end{pmatrix}$ gefordert.

Ebenso o.k. wäre aber z.B. auch $\begin{pmatrix} 1 & 3 & -7 & 5 \\ 0 & 0 & 0 & * \\ 0 & 0 & 0 & 0 \end{pmatrix}$, nicht aber $\begin{pmatrix} 1 & 3 & -7 & 5 \\ 0 & 0 & 0 & * \\ 0 & 2 & 0 & 0 \end{pmatrix}$.

- **spezielle Zeilenstufenform:**

Hier verschärft sich die Forderung zu $\begin{pmatrix} 1 & 3 & -7 & 5 \\ 0 & 1 & * & * \\ 0 & 0 & 1 & * \end{pmatrix}$ respektive $\begin{pmatrix} 1 & 3 & -7 & 5 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$.

Auch besitzt $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 5 \end{pmatrix}$ Zeilenstufenform, und $\begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ sowie $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ sind in spezieller Zeilenstufenform.

Nach so viel Matrixumgestaltung wollen wir nun endlich sehen, wie uns die Überführung von Matrizen in Zeilenstufenform bei der Lösung von LGS hilft und welche Rolle sie letztlich für den Gauß-Algorithmus spielt. Rufen wir uns dazu nochmals den Zeichenapparat und die wichtigsten Erkenntnisse über LGS der Form $Ax = b$ ins Gedächtnis zurück. Hierbei sind $A \in \text{Mat}(m \times n, \mathbb{R})$ und $b \in \mathbb{R}^m$ gegeben sowie $x \in \mathbb{R}^n$ gesucht. Wir notieren $\text{Lös}(A, b) = \{x \in \mathbb{R}^n \mid Ax = b\}$ für die Lösungsmenge des LGS.

Das LGS nennen wir homogen, falls $b = 0$ ist, ansonsten inhomogen. Homogene LGS

$$Ax = 0$$

sind immer lösbar, etwa durch die Trivillösung $x = \mathbf{0}$. Weiterhin sahen wir, dass $\text{Lös}(A, \mathbf{0})$ stets einen Untervektorraum von \mathbb{R}^n bildet. Für inhomogene LGS

$$Ax = b$$

bekommen wir die Lösungsmenge $\text{Lös}(A, b) = v + \text{Lös}(A, \mathbf{0})$ allein durch Kenntnis einer („speziellen“) Lösung $v \in \text{Lös}(A, b)$ sowie der zum homogenen LGS gehörigen Lösungsmenge $\text{Lös}(A, \mathbf{0})$.

Bisher fehlt noch eine effektive Entscheidungshilfe, für welche Fälle das inhomogene LGS $Ax = b$ lösbar ist:

Bemerkung 2.3.5. Ist (A, b) die erweiterte Koeffizientenmatrix eines LGS, und ist (A, b) in Zeilenstufenform, so gilt genau dann $\text{Lös}(A, b) = \emptyset$, wenn ein Index $i \in \{1, \dots, m\}$ existiert, so dass die i -te Zeile von A nur aus Nullen besteht, aber $b_i \neq 0$ ist.

Denn für ein solches i lautet die i -te Gleichung $A_{i1} \cdot x_1 + A_{i2} \cdot x_2 + \dots + A_{in} \cdot x_n = b_i$ des LGS:

$$0 = 0 \cdot x_1 + 0 \cdot x_2 + \dots + 0 \cdot x_n = b_i \neq 0 \quad \zeta$$

Diese Gleichung ist nicht lösbar.

Bestehen nun alle Zeilen von A nur dann aus Nullen, wenn auch die entsprechende Komponente $b_i = 0$ ist, dann lässt sich die Lösungsmenge des LGS durch *sukzessives Lösen* der einzelnen Gleichungen von unten nach oben ermitteln. Am besten, wir sehen uns die beiden Fälle einmal im Beispiel an:

Beispiel 2.3.6. Wir betrachten einerseits die erweiterte Koeffizientenmatrix von

$$\begin{pmatrix} 1 & 2 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 1 \end{pmatrix},$$

2 Matrixrechnung

d.h.

$$(A, b) = \begin{pmatrix} 1 & 2 & 4 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Aus der letzten Zeile der Matrix wird unmittelbar ersichtlich, dass $Ax = b$ nicht lösbar ist.

Beispiel 2.3.7. Nun wandeln wir das System ab, in dem wir die 1 im dritten Eintrag des b -Vektors durch 0 ersetzen, d.h. wir bekommen andererseits

$$(A, b) = \begin{pmatrix} 1 & 2 & 4 & 4 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Sukzessives Auflösen von unten nach oben liefert:

$$\begin{aligned} \text{Lös}(A, b) &= \{x \in \mathbb{R}^3 \mid 1 \cdot x_1 + 2 \cdot x_2 + 4 \cdot x_3 = 4 \ \& \ 1 \cdot x_3 = 5\} \\ &= \{x \in \mathbb{R}^3 \mid x_3 = 5 \ \& \ x_1 + 2x_2 + 4 \cdot 5 = 4\} \\ &= \{x \in \mathbb{R}^3 \mid x_3 = 5 \ \& \ x_1 = -2x_2 - 16 \ \& \ x_2 \in \mathbb{R} \text{ beliebig}\} \\ &= \left\{ \begin{pmatrix} -2x_2 - 16 \\ x_2 \\ 5 \end{pmatrix} \in \mathbb{R}^3 \mid x_2 \in \mathbb{R} \text{ beliebig} \right\} \end{aligned}$$

Für die konkrete Ermittlung der Lösungsmenge lohnt sich offenbar die Überführung der erweiterten Koeffizientenmatrix eines LGS in Zeilenstufenform. Die Lösungen lassen sich dann, wie gesehen, nach und nach durch „Rückwärtseinsetzen“ ermitteln.

Fassen wir die bisherigen Ergebnisse dieses Abschnitts zusammen und formulieren den

Gauß-Algorithmus zum Lösen von LGS. Seien $A \in \text{Mat}(m \times n, \mathbb{R})$ und $b \in \mathbb{R}^m$ gegeben. Gesucht sind alle $x \in \mathbb{R}^n$, für die $Ax = b$ gilt.

1. Stelle die erweiterte Koeffizientenmatrix (A, b) auf.
2. Überführe (A, b) durch elementare Zeilenumformungen in (\tilde{A}, \tilde{b}) so, dass (\tilde{A}, \tilde{b}) Zeilenstufenform besitzt².
3. Prüfe, ob $\text{Lös}(A, b) = \text{Lös}(\tilde{A}, \tilde{b}) = \emptyset$ gilt.
4. Ist dagegen $\text{Lös}(\tilde{A}, \tilde{b}) \neq \emptyset$, so löse $\tilde{A}x = \tilde{b}$ durch sukzessives Lösen der einzelnen Gleichungen von unten nach oben.

²Wir wissen, dass dann $\text{Lös}(A, b) = \text{Lös}(\tilde{A}, \tilde{b})$ gilt.

Manchmal interessiert man sich nur für die Lösbarkeit eines inhomogenen LGS und nicht für die Lösungsmenge selbst. Ein passendes Hilfsmittel hierfür ist der so genannte Rang einer Matrix. Dieses Konzept beruht auf der Erkenntnis, dass wir die Spalten einer $m \times n$ -Matrix

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1i} & \cdots & A_{1n} \\ A_{21} & & A_{2i} & & A_{2n} \\ \vdots & & \vdots & & \vdots \\ A_{m1} & \cdots & A_{mi} & \cdots & A_{mn} \end{pmatrix}$$

wie folgt handhaben können:

Definition 2.3.8. Für $A \in \text{Mat}(m \times n, \mathbb{R})$ und $i \in \{1, \dots, n\}$ heißt

$$a_i := \begin{pmatrix} A_{1i} \\ A_{2i} \\ \vdots \\ A_{mi} \end{pmatrix} \in \mathbb{R}^m$$

i -ter **Spaltenvektor** von A .

Damit setzt sich A aus Spaltenvektoren a_1, \dots, a_n zusammen.

Bemerkung 2.3.9. Häufig ist es nützlich, den i -ten Spaltenvektor a_i wie folgt auszudrücken. Für jedes $i \in \{1, \dots, n\}$ gilt

$$A \cdot e_i = \begin{pmatrix} \sum_{j=1}^n A_{1j} \cdot (e_i)_j \\ \vdots \\ \sum_{j=1}^n A_{mj} \cdot (e_i)_j \end{pmatrix} = \begin{pmatrix} A_{1i} \\ \vdots \\ A_{mi} \end{pmatrix} = a_i,$$

wobei e_i der uns bekannte i -te Einheitsvektor aus der Standardbasis von \mathbb{R}^n ist.

Definition 2.3.10. Für $A \in \text{Mat}(m \times n, \mathbb{R})$ heißt

$$\begin{aligned} \text{rg}(A) &:= \text{maximale Anzahl linear unabhängiger Spaltenvektoren von } A \\ &= \dim(L(a_1, \dots, a_n)) \end{aligned}$$

der **Rang** (bzw. Spaltenrang) von A .

2 Matrixrechnung

Beispiel 2.3.11. Es gilt

$$\operatorname{rg} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} = 3$$

und

$$\operatorname{rg} \begin{pmatrix} 1 & 0 & 1 & -1 \\ 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix} = 2.$$

Wie angekündigt, folgt nun die Charakterisierung der Lösbarkeit eines (inhomogenen) LGS durch den Rangbegriff:

Satz 2.3.12

Sei $A \in \operatorname{Mat}(m \times n, \mathbb{R})$ und $b \in \mathbb{R}^m$. Dann gilt:

1. $\operatorname{Lös}(A, b) \neq \emptyset \Leftrightarrow b \in L(a_1, \dots, a_n) \Leftrightarrow \operatorname{rg}(A, b) = \operatorname{rg}(A)$
2. $\operatorname{Lös}(A, b) = \emptyset \Leftrightarrow b \notin L(a_1, \dots, a_n) \Leftrightarrow \operatorname{rg}(A, b) = \operatorname{rg}(A) + 1$

Beweis.

Stets gilt:

$$\operatorname{rg}(A, b) = \begin{cases} \operatorname{rg}(A) & , \text{ falls } b \in L(a_1, \dots, a_n) \\ \operatorname{rg}(A) + 1 & , \text{ falls } b \notin L(a_1, \dots, a_n) \end{cases}$$

Es genügt nun zu zeigen, dass $\operatorname{Lös}(A, b) \neq \emptyset \Leftrightarrow b \in L(a_1, \dots, a_n)$ gilt, denn die Aussage $\operatorname{Lös}(A, b) = \emptyset \Leftrightarrow b \notin L(a_1, \dots, a_n)$ ist dazu logisch äquivalent.

„ \Leftarrow “:

Sei $b \in L(a_1, \dots, a_n)$, d.h. es existieren $x_1, \dots, x_n \in \mathbb{R}$, so dass

$$\begin{aligned} b &= \sum_{i=1}^n x_i \cdot a_i \\ &= \sum_{i=1}^n x_i \cdot A e_i \\ &= \sum_{i=1}^n A \cdot x_i \cdot e_i \end{aligned}$$

$$\begin{aligned}
 &= A \cdot \left(\sum_{i=1}^n x_i \cdot e_i \right) \\
 &= A \cdot x
 \end{aligned}$$

mit:

$$\begin{aligned}
 x &:= \sum_{i=1}^n x_i \cdot e_i \\
 &= x_1 e_1 + \dots + x_n e_n \\
 &= x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \\
 &= \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}
 \end{aligned}$$

Insbesondere löst $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ damit das LGS $Ax = b$. Also ist die Lösungsmenge nicht leer.

„ \Rightarrow “:

Sei umgekehrt $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n x_i \cdot e_i$ eine Lösung von $Ax = b$, d.h. $x \in \text{Lös}(A, b)$. Dann gilt genau wie eben:

$$\begin{aligned}
 b &= A \cdot x \\
 &= A \sum_{i=1}^n x_i \cdot e_i \\
 &= \sum_{i=1}^n x_i \cdot A e_i \\
 &= \sum_{i=1}^n x_i \cdot a_i \\
 &\in \text{L}(a_1, \dots, a_n)
 \end{aligned}$$

□

Beschließen wir diesen Abschnitt mit einem komplexeren Beispiel.

2 Matrixrechnung

Beispiel 2.3.13. Die TNT-Reaktionsgleichung (2.2) führte uns auf das homogene LGS $Ax = \mathbf{0}$ der konkreten Gestalt:

$$\begin{pmatrix} 7 & 0 & -7 & 0 \\ 8 & 1 & -5 & -2 \\ 0 & 1 & -3 & 0 \\ 0 & 3 & -6 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Bekanntlich bildet $\text{Lös}(A, \mathbf{0})$ ein Untervektorraum von \mathbb{R}^4 . Wir ermitteln eine Basis von $\text{Lös}(A, \mathbf{0})$. Mit dem Gauß-Algorithmus überführen wir dazu A in Zeilenstufenform \tilde{A} , denn: $\text{Lös}(A, \mathbf{0}) = \text{Lös}(\tilde{A}, \mathbf{0})$. Es gilt:

$$\begin{aligned} A &= \begin{pmatrix} 7 & 0 & -7 & 0 \\ 8 & 1 & -5 & -2 \\ 0 & 1 & -3 & 0 \\ 0 & 3 & -6 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 8 & 1 & -5 & -2 \\ 0 & 1 & -3 & 0 \\ 0 & 3 & -6 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 8 & 0 & -8 & 0 \\ 8 & 1 & -5 & -2 \\ 0 & 1 & -3 & 0 \\ 0 & 3 & -6 & -1 \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} 8 & 0 & -8 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 1 & -3 & 0 \\ 0 & 3 & -6 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 1 & -3 & 0 \\ 0 & 3 & -6 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 3 & -6 & -1 \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 3 & 9 & -6 \\ 0 & 0 & -6 & 2 \\ 0 & 3 & -6 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 3 & 9 & -6 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -15 & 5 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -15 & 5 \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & -15 & 5 \\ 0 & 0 & -15 & 5 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & -15 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\ &=: \tilde{A} \end{aligned}$$

Beginnend bei der untersten Zeile von \tilde{A} lösen wir das LGS nun sukzessive auf und bekommen:

$$\begin{aligned} \text{Lös}(A, \mathbf{0}) &= \text{Lös}(\tilde{A}, \mathbf{0}) \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 \mid \begin{array}{l} x_1 - x_3 = 0 \\ x_2 + 3x_3 - 2x_4 = 0 \\ -3x_3 + x_4 = 0 \end{array} \right\} \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathbb{R}^4 \mid \begin{array}{l} x_1 - x_3 = 0 \\ x_2 + 3x_3 - 2x_4 = 0 \\ x_4 = 3x_3 \end{array} \right\} \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 3x_3 \end{pmatrix} \in \mathbb{R}^4 \mid \begin{array}{l} x_3 = x_1 \\ x_2 - 3x_3 = 0 \end{array} \right\} \end{aligned}$$

$$\begin{aligned}
&= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_1 \\ 3x_1 \end{pmatrix} \in \mathbb{R}^4 \mid x_2 = 3x_1 \right\} \\
&= \left\{ \begin{pmatrix} x_1 \\ 3x_1 \\ x_1 \\ 3x_1 \end{pmatrix} \in \mathbb{R}^4 \mid x_1 \in \mathbb{R} \text{ beliebig} \right\} \\
&= \left\{ x_1 \cdot \begin{pmatrix} 1 \\ 3 \\ 1 \\ 3 \end{pmatrix} \in \mathbb{R}^4 \mid x_1 \in \mathbb{R} \text{ beliebig} \right\} \\
&= \text{L} \left(\begin{pmatrix} 1 \\ 3 \\ 1 \\ 3 \end{pmatrix} \right)
\end{aligned}$$

Eine Basis von $\text{Lös}(A, \mathbf{0})$ ist dann offenbar gegeben durch:

$$X = \left\{ \begin{pmatrix} 1 \\ 3 \\ 1 \\ 3 \end{pmatrix} \right\}$$

2.4 Geometrie der Ebene

Wir betreiben Geometrie unter der Maßgabe, dass lineare Algebra und analytische Geometrie keine isolierten Themenfelder darstellen, sondern vielmehr eine Synthese zwischen ihnen vorliegt. Vektor- und Matrixrechnung werden sich als außerordentlich nützliches Hilfsmittel zur Behandlung geometrischer Fragestellungen erweisen.

Als mathematisches Modell für die Ebene verwenden wir \mathbb{R}^2 und fassen die Ebene als ein unbegrenzt ausgedehntes flaches zweidimensionales Objekt auf. Dies bedeutet, dass wir die Ebene koordinatisieren, d.h. wir identifizieren Punkte der Ebene anhand zwei reeller Zahlen, ihren Koordinaten³, die gerade die beiden Komponenten des entsprechenden Vektors in \mathbb{R}^2 sind.

Geometrische Beweise erleichtern sich durch den Einsatz von Vektorrechnung z.T. erheblich. Wir beginnen mit einer Begriffsbestimmung.

Definition 2.4.1

Eine **Gerade** in der Ebene ist eine Menge der Form

$$G_{a,v} = \{a + t \cdot v \mid t \in \mathbb{R}\} \subset \mathbb{R}^2.$$

Hierbei ist $a \in \mathbb{R}^2$ der **Aufpunkt** und $v \in \mathbb{R}^2 \setminus \{0\}$ der **Richtungsvektor**. Ferner heißt $t \in \mathbb{R}$ **Streckungsfaktor**.

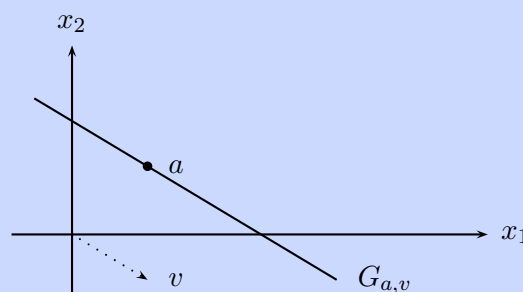


Abb. 13

Wäre $v = 0$ zugelassen, so führte das zu einer Menge $G_{a,0} = \{a\}$, welche offenbar nur einen Punkt enthält - nicht gerade das, was wir uns unter einer Gerade vorstellen.

Wir zeigen nun, dass der Aufpunkt einer Geraden gegen jeden beliebigen Punkt der Gerade austauschbar ist.

Lemma 2.4.2

Ist $G_{a,v} \subset \mathbb{R}^2$ eine Gerade und $p \in G_{a,v}$, so gilt:

$$G_{a,v} = G_{p,v}$$

³Die Vorgehensweise, Vektorrechnung für geometrische Überlegungen zu nutzen und Koordinaten in der Ebene einzuführen, geht auf den französischen Philosoph René Descartes (1596 (La Haye/Touraine, Frankreich) - 1650 (Stockholm, Schweden)), siehe http://de.wikipedia.org/wiki/Rene_Descartes, zurück. Man spricht daher auch von kartesischen Koordinaten.

Beweis.

„ $G_{a,v} \supset G_{p,v}$ “:

Sei $q \in G_{p,v}$ vorgegeben. Wir müssen zeigen, dass nun auch $q \in G_{a,v}$ zutrifft.

Wegen $p \in G_{a,v}$ gibt es ein bestimmtes $t_0 \in \mathbb{R}$, so dass

$$p = a + t_0 \cdot v \quad (2.4)$$

gilt. Andererseits impliziert $q \in G_{p,v}$ die Existenz eines $t_1 \in \mathbb{R}$, so dass

$$q = p + t_1 \cdot v$$

gilt und wir (2.4) einsetzen können:

$$\begin{aligned} q = p + t_1 \cdot v &= a + t_0 \cdot v + t_1 \cdot v \\ &= a + (t_0 + t_1) \cdot v \\ &\in G_{a,v} \end{aligned}$$

„ $G_{a,v} \subset G_{p,v}$ “:

Wir zeigen: Gilt $x \in G_{a,v}$, so auch $x \in G_{p,v}$.

Zunächst existiert wegen $p \in G_{a,v}$ ein $t_0 \in \mathbb{R}$, so dass

$$p = a + t_0 \cdot v \quad (2.5)$$

gilt. Ist nun $x \in G_{a,v}$, dann gibt es ein $t_2 \in \mathbb{R}$, so dass

$$x = a + t_2 \cdot v$$

gilt. Umstellen von (2.5) nach a und Einsetzen führt schließlich auf:

$$\begin{aligned} x = a + t_2 \cdot v &= p - t_0 \cdot v + t_2 \cdot v \\ &= p + (t_2 - t_0) \cdot v \\ &\in G_{p,v} \end{aligned}$$

□

Wir beweisen nun, dass der Richtungsvektor aus zwei verschiedenen Punkten der Gerade gebildet werden kann. Insbesondere legen somit zwei verschiedene Punkte eine Gerade eindeutig fest.

Lemma 2.4.3

Ist $G \subset \mathbb{R}^2$ eine Gerade und sind $a, b \in G$ mit $a \neq b$, so gilt:

$$G = G_{a,b-a}$$

2 Matrixrechnung

Beweis.

Laut Definition gibt es für die Gerade G einen Aufpunkt $p \in \mathbb{R}^2$ und einen Richtungsvektor $v \in \mathbb{R}^2 \setminus \{0\}$, so dass

$$G = G_{p,v}$$

gilt. Wegen Lemma 2.4.2 können wir p durch a ersetzen, d.h. es gilt auch

$$G = G_{a,v}.$$

Ist nun $b \in G_{a,v}$ ein weiterer Punkt der Gerade, dann existiert ein $t_0 \in \mathbb{R}$, so dass

$$b = a + t_0 \cdot v \tag{2.6}$$

und wegen $a \neq b$ insbesondere $t_0 \neq 0$ gilt. Mittels dieser Grundlage zeigen wir $G_{a,v} = G_{a,b-a}$.

„ $G_{a,v} \subset G_{a,b-a}$ “:

Zu $x \in G_{a,v}$ existiert ein $t_1 \in \mathbb{R}$, so dass

$$x = a + t_1 \cdot v$$

gilt. Da $t_0 \neq 0$ ist, kann die Gleichung (2.6) nach v umgestellt und das Resultat

$$v = \frac{1}{t_0} \cdot (b - a)$$

eingesetzt werden:

$$\begin{aligned} x = a + t_1 \cdot v &= a + t_1 \cdot \frac{1}{t_0} \cdot (b - a) \\ &= a + \frac{t_1}{t_0} \cdot (b - a) \\ &\in G_{a,b-a} \end{aligned}$$

„ $G_{a,v} \supset G_{a,b-a}$ “:

Umgekehrt zieht $y \in G_{a,b-a}$ die Existenz eines $t_2 \in \mathbb{R}$ nach sich, so dass

$$y = a + t_2 \cdot (b - a)$$

gilt. Auch lässt sich die Gleichung (2.6) nach

$$b - a = t_0 \cdot v$$

überführen und damit ist:

$$\begin{aligned} y = a + t_2 \cdot (b - a) &= a + t_2 \cdot t_0 \cdot v \\ &= a + (t_2 t_0) \cdot v \\ &\in G_{a,v} \end{aligned}$$

□

Wir gehen nun der Frage nach, wie man eigentlich Abstände in der Ebene misst. Wie kann die Länge eines Vektors berechnet werden? Wir werden uns ausgiebig mit dem Thema Längenmessung und dessen Anwendungen beschäftigen. Wichtigstes Hilfsmittel hierfür liefert

Definition 2.4.4. Für $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ heißt

$$\langle x, y \rangle := x_1 \cdot y_1 + x_2 \cdot y_2$$

Skalarprodukt der Vektoren x und y .

Bemerkung 2.4.5. Offenbar ist das Skalarprodukt eine Abbildung:

$$\langle \cdot, \cdot \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$$

Ferner gelten für alle $x, x', y \in \mathbb{R}^2$ und jedes $t \in \mathbb{R}$ folgende Rechenregeln:

1. $\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle$
2. $\langle t \cdot x, y \rangle = t \cdot \langle x, y \rangle$
3. $\langle x, y \rangle = \langle y, x \rangle$
4. $\langle x, x \rangle \geq 0$ und $\langle x, x \rangle = 0 \Leftrightarrow x = \mathbf{0}$

Definition 2.4.6

Zu $x \in \mathbb{R}^2$ heißt

$$\begin{aligned} \|x\| &:= \sqrt{\langle x, x \rangle} \\ &= \sqrt{x_1^2 + x_2^2} \end{aligned}$$

Norm von x .

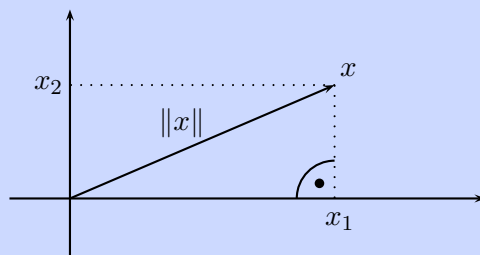


Abb. 14

2 Matrixrechnung

Berücksichtigt man das Entstehen rechtwinkliger Dreiecke, wie in der Abbildung ersichtlich, so kann die Norm eines Vektors auf Grund des Satzes von Pythagoras unmittelbar mit dessen Länge identifiziert werden.

Aus der Rechenregel (4.) für das Skalarprodukt ergibt sich für die Norm weiterhin

$$\|x\| \geq 0 \text{ und } \|x\| = 0 \Leftrightarrow x = \mathbf{0}$$

für alle $x \in \mathbb{R}^2$. Aus den Rechenregeln (2.) und (3.) finden wir für alle $x \in \mathbb{R}^2$ und jedes $t \in \mathbb{R}$:

$$\begin{aligned} \|t \cdot x\| &= \sqrt{\langle t \cdot x, t \cdot x \rangle} \\ &= \sqrt{t \cdot \langle x, t \cdot x \rangle} \\ &= \sqrt{t \cdot \langle t \cdot x, x \rangle} \\ &= \sqrt{t^2 \cdot \langle x, x \rangle} \\ &= \sqrt{t^2} \cdot \sqrt{\langle x, x \rangle} \\ &= |t| \cdot \|x\| \end{aligned}$$

Der noch ausstehende, allgemeine Abstandsbegriff erfordert nun zunächst einen kleinen Abstecher in die Analysis:

Definition 2.4.7. Eine Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$ heißt **monoton**, falls für alle $t_1, t_2 \in \mathbb{R}$ gilt:

$$t_1 \leq t_2 \Rightarrow f(t_1) \leq f(t_2)$$

So ist etwa die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x$ monoton, während $g : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$ nicht monoton ist, da sich z.B. für $-2 \leq 0$ der Widerspruch $g(-2) = 4 \leq 0 = g(0)$ ergibt. Ferner ist auch $h : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+, x \mapsto \sqrt{x}$ monoton. Diese Eigenschaft wird sich im Beweis des folgenden Satzes und dessen Folgerung als hilfreich erweisen.

Satz 2.4.8 (Cauchy-Schwarz-Ungleichung [CSU])

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

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

Beweis.

Wir berechnen die Differenz der Quadrate beider Seiten. Es gilt:

$$\begin{aligned}
 (\|x\| \cdot \|y\|)^2 - (|\langle x, y \rangle|)^2 &= \|x\|^2 \cdot \|y\|^2 - \langle x, y \rangle^2 \\
 &= \left(\sqrt{\langle x, x \rangle}\right)^2 \cdot \left(\sqrt{\langle y, y \rangle}\right)^2 - \langle x, y \rangle^2 \\
 &= \langle x, x \rangle \cdot \langle y, y \rangle - \langle x, y \rangle^2 \\
 &= (x_1^2 + x_2^2) \cdot (y_1^2 + y_2^2) - (x_1 y_1 + x_2 y_2)^2 \\
 &= (x_1^2 + x_2^2) \cdot (y_1^2 + y_2^2) - (x_1^2 y_1^2 + 2x_1 x_2 y_1 y_2 + x_2^2 y_2^2) \\
 &= x_1^2 y_1^2 + x_1^2 y_2^2 + x_2^2 y_1^2 + x_2^2 y_2^2 - x_1^2 y_1^2 - 2x_1 x_2 y_1 y_2 - x_2^2 y_2^2 \\
 &= x_1^2 y_2^2 - 2x_1 x_2 y_1 y_2 + x_2^2 y_1^2 \\
 &= (x_1 y_2 - x_2 y_1)^2 \\
 &\geq 0
 \end{aligned}$$

Insgesamt haben wir also:

$$\begin{aligned}
 (\|x\| \cdot \|y\|)^2 - (|\langle x, y \rangle|)^2 &\geq 0 \\
 \Leftrightarrow (\|x\| \cdot \|y\|)^2 &\geq (|\langle x, y \rangle|)^2
 \end{aligned}$$

Aufgrund der Monotonie der Wurzelfunktion bleibt die Ungleichung erhalten, wenn man auf beiden Seiten die Wurzel zieht. Dies liefert $\|x\| \cdot \|y\| \geq |\langle x, y \rangle|$. \square

Folgerung 2.4.9 (Dreiecksungleichung für $\|\cdot\|$)

Für alle $x, y \in \mathbb{R}^2$ gilt:

$$\|x + y\| \leq \|x\| + \|y\|$$

Beweis.

Wir berechnen:

$$\begin{aligned}
 \|x + y\|^2 &= \left(\sqrt{\langle x + y, x + y \rangle}\right)^2 \\
 &= \langle x + y, x + y \rangle \\
 &= \langle x, x + y \rangle + \langle y, x + y \rangle \\
 &= \langle x + y, x \rangle + \langle x + y, y \rangle \\
 &= \langle x, x \rangle + \langle y, x \rangle + \langle x, y \rangle + \langle y, y \rangle \\
 &= \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle \\
 &= \|x\|^2 + 2\langle x, y \rangle + \|y\|^2
 \end{aligned}$$

2 Matrixrechnung

$$\begin{aligned} & \stackrel{\text{CSU}}{\leq} \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ & \leq (\|x\| + \|y\|)^2. \end{aligned}$$

Wieder führt das Anwenden der Wurzelfunktion auf beiden Seiten zum gewünschten Resultat. \square

An Hand folgender Abbildung erklären wir, weshalb die eben bewiesene Ungleichung den Namen „Dreiecksungleichung“ trägt:

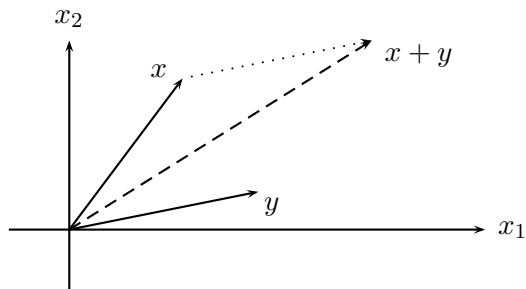


Abb. 15

Offenbar ist der direkte Weg $\|x + y\|$ stets höchstens so lang, wie der „Umweg“ $\|x\| + \|y\|$. Passend angeordnet, bilden x , y und $x + y$ ein Dreieck, welches nur dann konstruierbar ist, wenn zwei Seiten zusammen mindestens so lang sind, wie die dritte Seite.

Nun können wir sinnvoll den Abstandsbegriff zweier Punkte in der Ebene definieren:

Definition 2.4.10. Zu $x, y \in \mathbb{R}^2$ heißt

$$d(x, y) := \|x - y\|$$

euklidischer Abstand von x und y .

Die Definition der Norm als Längenmaß für Ortsvektoren (Vektoren, die im Nullpunkt angetragen sind) ordnet sich prima in den Abstandsbegriff ein. Man erhält sie für den Abstand eines Vektors zum Nullvektor:

$$\begin{aligned} \|x\| &= \|x - \mathbf{0}\| \\ &= d(x, \mathbf{0}) \end{aligned}$$

Bemerkung 2.4.11. Sind $x, y \in \mathbb{R}^2$, so gilt $d(x, y) = d(y, x)$, denn:

$$\begin{aligned} d(x, y) &= \|x - y\| \\ &= \|(-1) \cdot (y - x)\| \\ &= |-1| \cdot \|y - x\| \\ &= \|y - x\| \\ &= d(y, x) \end{aligned}$$

Folgerung 2.4.12

Sind $x, y, z \in \mathbb{R}^2$, so gilt die Dreiecksungleichung:

$$d(x, z) \leq d(x, y) + d(y, z)$$

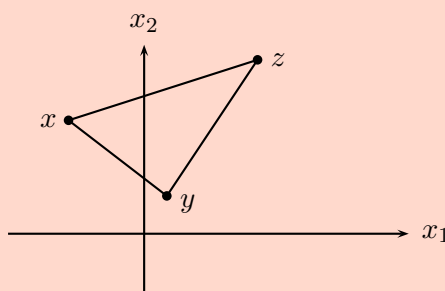


Abb. 16

Beweis.

Wir beginnen mit der linken Seite und rechnen los:

$$\begin{aligned} d(x, z) &= \|x - z\| \\ &= \|x - y + y - z\| \\ &= \|(x - y) + (y - z)\| \\ &\stackrel{\text{Folg. (2.4.9)}}{\leq} \|x - y\| + \|y - z\| \\ &= d(x, y) + d(y, z) \end{aligned}$$

□

Lemma 2.4.13 (Mittelpunkts-Lemma)

Sei $G \subset \mathbb{R}^2$ eine Gerade und $a, b \in G$, $a \neq b$. Dann existiert genau ein Punkt $c \in G$, so dass gilt:

$$d(a, c) = d(b, c)$$

2 Matrixrechnung

Für diesen Punkt c gilt ferner:

$$c = \frac{1}{2}(a + b) \quad \text{und} \quad d(a, c) = d(b, c) = \frac{1}{2}d(a, b)$$

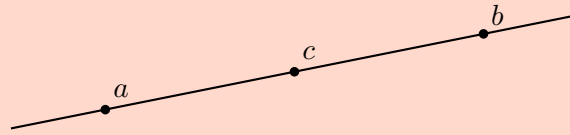


Abb. 17

Beweis.

Nach Lemma (2.4.3) wissen wir, dass G durch die auf ihr befindlichen Punkte $a \neq b$ eindeutig festgelegt ist, d.h. es gilt:

$$G = G_{a,b-a} = \{a + t \cdot (b - a) \mid t \in \mathbb{R}\}$$

Sei nun $q = a + t_0 \cdot (b - a) \in G$ für ein bestimmtes $t_0 \in \mathbb{R}$. Dann gilt

$$\begin{aligned} d(a, q) &= \|a - q\| \\ &= \|a - (a + t_0 \cdot (b - a))\| \\ &= \|(-t_0) \cdot (b - a)\| \\ &= |t_0| \cdot \|b - a\| \end{aligned}$$

und analog:

$$\begin{aligned} d(b, q) &= \|b - q\| \\ &= \|b - (a + t_0 \cdot (b - a))\| \\ &= \|b - a - t_0 \cdot b + t_0 \cdot a\| \\ &= \|(1 - t_0) \cdot (b - a)\| \\ &= |1 - t_0| \cdot \|b - a\| \end{aligned}$$

Damit erhalten wir ein Kriterium dafür, dass q von a und b denselben Abstand hat:

$$\begin{aligned} d(a, q) = d(b, q) &\Leftrightarrow |t_0| \cdot \|b - a\| = |1 - t_0| \cdot \|b - a\| \\ &\stackrel{a \neq b}{\Leftrightarrow} |t_0| = |1 - t_0| \\ &\Leftrightarrow t_0^2 = (1 - t_0)^2 \\ &\Leftrightarrow t_0^2 = 1 - 2t_0 + t_0^2 \\ &\Leftrightarrow t_0 = \frac{1}{2} \end{aligned}$$

Damit ist

$$c := a + \frac{1}{2}(b - a) = a + \frac{1}{2}b - \frac{1}{2}a = \frac{1}{2}(a + b).$$

der einzige Punkt auf G , der von a und b denselben Abstand hat. Dieser Abstand berechnet sich dann zu

$$\begin{aligned} d(b, c) &= d(a, c) \\ &= d\left(a, \frac{1}{2}(a + b)\right) \\ &= \left\| a - \frac{1}{2}(a + b) \right\| \\ &= \left\| a - \frac{1}{2}a - \frac{1}{2}b \right\| \\ &= \left\| \frac{1}{2}(a - b) \right\| \\ &= \frac{1}{2} \|a - b\| \\ &= \frac{1}{2} d(a, b). \end{aligned}$$

□

Definition 2.4.14. Der Punkt $\frac{1}{2}(a + b)$ heißt **Mittelpunkt** von a und b .
(Diese Definition ist auch im Fall $a = b$ zulässig.)

Definition 2.4.15
Zwei Geraden $G_{a,v}$ und $G_{b,w}$ heißen genau dann **parallel**, wenn ihre Richtungsvektoren v, w linear abhängig sind.

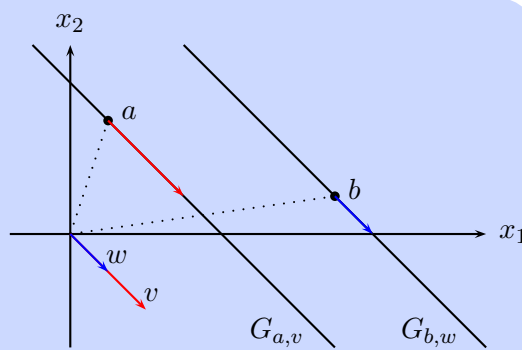


Abb. 18

2 Matrixrechnung

Bemerkung 2.4.16. Für parallele Geraden $G_{a,v}$ und $G_{b,w}$ gilt:

$$G_{b,w} = G_{b,v}$$

und

$$G_{a,v} = G_{a,w},$$

denn mit $w = \alpha v$ für $\alpha \in \mathbb{R} \setminus \{0\}$ ist:

$$\begin{aligned} G_{b,w} &= \{b + tw \mid t \in \mathbb{R}\} \\ &= \{b + t(\alpha v) \mid t \in \mathbb{R}\} \\ &= \{b + (t\alpha)v \mid t \in \mathbb{R}\} \\ &= \{b + t'v \mid t' \in \mathbb{R}\} \\ &= G_{b,v} \end{aligned}$$

Mit $G_{a,v} = G_{a,w}$ verfährt man analog.

Anders als im funktionenzentrierten Analysisunterricht der Schule, drücken wir Geraden $G_{a,v}$ durch Punktmengen im \mathbb{R}^2 aus. Auf dieser Erkenntnis basiert das folgende Lemma, welches eine andere, mengentheoretische Charakterisierung der Parallelität von Geraden bereitstellt.

Lemma 2.4.17

Zwei Geraden G und G' in der Ebene sind genau dann parallel, wenn entweder $G = G'$ oder $G \cap G' = \emptyset$ gilt.

Beweis.

a) Seien G und G' parallel. Wir zeigen, dass dann $G = G'$ oder $G \cap G' = \emptyset$ gilt.

Nehmen wir also an, dass $G \cap G' \neq \emptyset$, denn sonst sind wir schon fertig. Dann gibt es ein $a \in G \cap G'$. Wegen Lemma 2.4.2 können wir dann $G = G_{a,v}$ und $G' = G_{a,w}$ für bestimmte $v, w, \in \mathbb{R}^2 \setminus \{0\}$ schreiben. Da G und G' parallel sind, gilt wegen Bemerkung 2.4.16:

$$G_{a,w} = G_{a,v},$$

d.h.

$$G = G'.$$

b) Sei nun umgekehrt $G = G'$ oder $G \cap G' = \emptyset$. Wir zeigen, dass dann G und G' parallel sind. Falls $G = G'$ gilt, so ist dies klar. Sei also $G \cap G' = \emptyset$. Untersuchen wir, was diese Bedingung

bedeutet. Ein Punkt x in $G \cap G'$ lässt sich sowohl in der Form $x = a + t_1v$ als auch in der Form $x = b + t_2w$ schreiben, wobei wir $G = G_{a,v}$ und $G' = G_{b,w}$ geschrieben haben. Also gilt

$$\begin{aligned} G \cap G' \neq \emptyset &\Leftrightarrow \exists t_1, t_2 \in \mathbb{R} : a + t_1v = b + t_2w \\ &\Leftrightarrow \exists t_1, t_2 \in \mathbb{R} : (v, w) \cdot \begin{pmatrix} t_1 \\ -t_2 \end{pmatrix} = b - a \\ &\Leftrightarrow \text{Lös}((v, w), a - b) \neq \emptyset \end{aligned}$$

Da wir $G \cap G' = \emptyset$ vorausgesetzt hatten, heißt das also $\text{Lös}((v, w), a - b) = \emptyset$. Somit ist $\text{rg}(v, w) < \text{rg}(v, w, a - b)$. Wegen $v, w, a - b \in \mathbb{R}^2$ ist $\text{rg}(v, w, a - b) \leq 2$ und somit $\text{rg}(v, w) \leq 1$. Also sind v und w linear abhängig und somit G und G' parallel. \square

Definition 2.4.18. Ein **Parallelogramm** ist ein 4-Tupel (a, b, c, d) von Punkten $a, b, c, d \in \mathbb{R}^2$, so dass gilt:

- $G_{a,b-a}$ ist parallel zu $G_{c,d-c}$
- $G_{a,c-a}$ ist parallel zu $G_{b,d-b}$

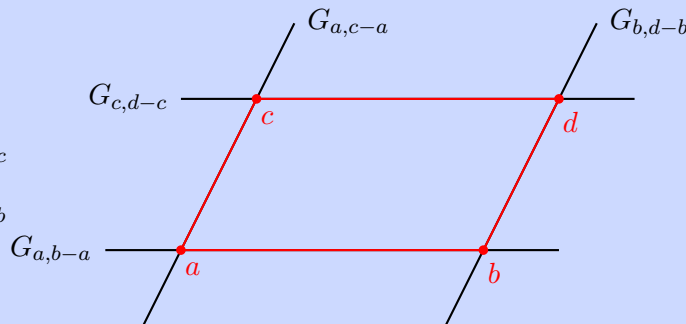


Abb. 19

Ferner heißt ein Parallelogramm (a, b, c, d) *nicht entartet*, falls keine drei Punkte auf einer Geraden liegen.

Indem wir verlangen, dass ein Parallelogramm nicht entartet sein soll, schließt es stets eine Fläche ein. Etwa für $a = b$ und $c = d$ ist dies nicht der Fall. Als entartet sehen wir insbesondere den Fall an, bei dem alle vier Punkte auf einer Geraden liegen:

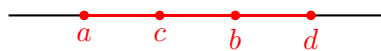


Abb. 20

Bemerkung 2.4.19. Ist (a, b, c, d) ein nicht entartetes Parallelogramm, so sind die Geraden $G_{a,b-a}, G_{c,d-c}, G_{a,c-a}, G_{b,d-b}$ paarweise verschieden und $b - a, c - a$ sind linear unabhängig.

2 Matrixrechnung

Beweis.

Den Nachweis hierfür genehmigen wir uns als Übungsaufgabe. □

Lemma 2.4.20

Ist (a, b, c, d) ein nicht entartetes Parallelogramm, so gilt:

$$c - a = d - b \quad \text{und} \quad b - a = d - c$$

Beweis.

Betrachten wir hierfür die parallelen Geraden $G_{a,c-a}$ und $G_{b,d-b}$. Dann gilt

$$c - a = \lambda(d - b) \tag{2.7}$$

für ein $\lambda \in \mathbb{R}$. Derselbe Ansatz für die Parallelen $G_{a,b-a}$ und $G_{c,d-c}$ liefert

$$b - a = \mu(d - c) \tag{2.8}$$

für ein $\mu \in \mathbb{R}$. Nun zeigen wir $\lambda = \mu = 1$. Hierbei führt Subtraktion der Gleichung (2.8) von (2.7) auf:

$$\begin{aligned} c - b &= \lambda(d - b) - \mu(d - c) \\ \Leftrightarrow c - d + d - b &= \lambda(d - b) - \mu(d - c) \\ \Leftrightarrow -(d - c) + (d - b) &= \lambda(d - b) - \mu(d - c) \\ \Leftrightarrow \mu(d - c) - (d - c) + (d - b) - \lambda(d - b) &= \mathbf{0} \\ \Leftrightarrow (\mu - 1)(d - c) + (1 - \lambda)(d - b) &= \mathbf{0} \end{aligned}$$

Da das Parallelogramm nicht entartet ist, sind die Vektoren $d - c$ und $d - b$ linear unabhängig. Somit gibt es nur die triviale Linearkombination des Nullvektor und es folgt $\mu - 1 = 0$ sowie $1 - \lambda = 0$, also insgesamt $\mu - \lambda = 1$. □

Zwei interessante Strecken in einem Parallelogramm sind die Diagonalen. Für sie gilt der

Satz 2.4.21 (Diagonalensatz)

In einem nicht entarteten Parallelogramm halbieren sich die Diagonalen gegenseitig.

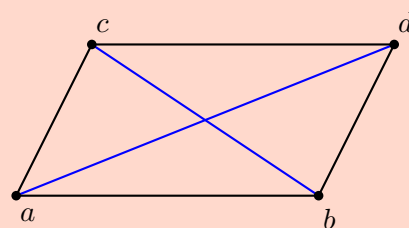


Abb. 21

Beweis.

Wie wir wissen, ist $\frac{1}{2}(a + d)$ der Mittelpunkt der Diagonalen durch a und d sowie $\frac{1}{2}(b + c)$ der Mittelpunkt der Diagonalen durch b und c . Wir zeigen nun die Gleichheit $\frac{1}{2}(a + d) = \frac{1}{2}(b + c)$ mittels direkter Rechnung. Dazu subtrahieren wir, wie schon oft zuvor, die rechte Seite von der linken und steuern auf den Nullvektor als Ergebnis zu:

$$\begin{aligned} \frac{1}{2}(a + d) - \frac{1}{2}(b + c) &= \frac{1}{2}(a + d - b + c) \\ &= \frac{1}{2}(-(c - a) + (d - b)) \\ &\stackrel{\text{Lemma (2.4.20)}}{=} \frac{1}{2}(-(d - b) + (d - b)) \\ &= \mathbf{0} \end{aligned}$$

□

Man beachte, dass der Diagonalsatz für allgemeine Vierecke falsch ist, wie folgende Abbildung illustriert:

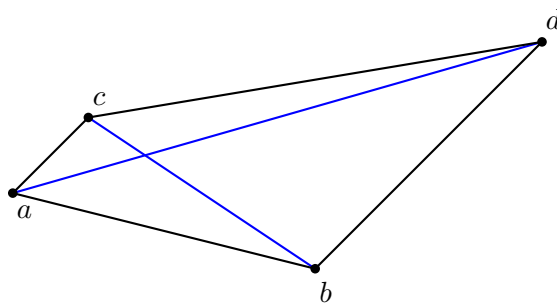


Abb. 22

Nachfolgend sparen wir eine Ecke ein und wenden uns Dreiecken zu.

Definition 2.4.22. Ein **Dreieck** ist ein Tripel (a, b, c) von Punkten $a, b, c \in \mathbb{R}^2$. Ein Dreieck (a, b, c) heißt *nicht entartet*, falls a, b, c nicht auf einer Geraden liegen.

Waren für Parallelogramme im Wesentlichen die Diagonalen interessant, so gibt es im Zusammenhang mit Dreiecken eine Fülle interessanter Geraden bzw. Strecken, etwa Seiten- und Winkelhalbierende, Mittelsenkrechte und Höhen.

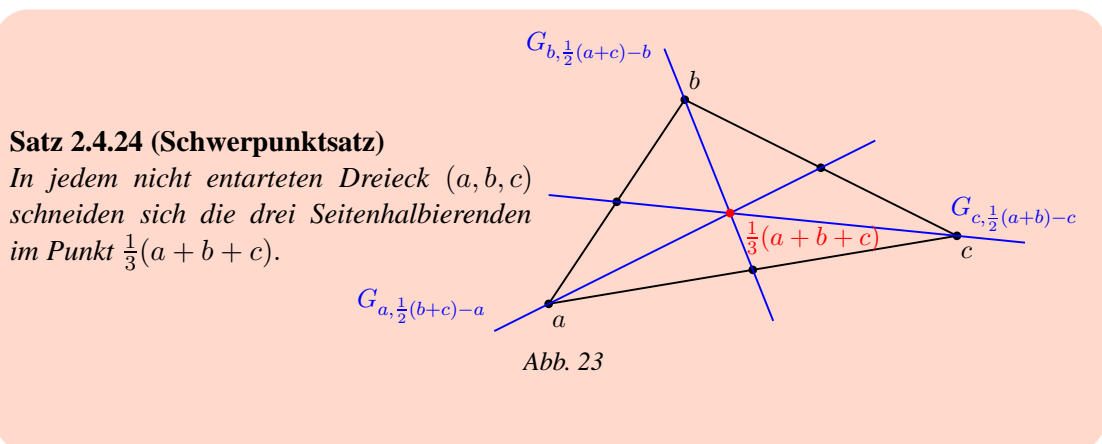
2 Matrixrechnung

Definition 2.4.23. Sei (a, b, c) ein nicht entartetes Dreieck. Eine **Seitenhalbierende** ist eine Gerade durch eine der Ecken a, b, c des Dreiecks und den Mittelpunkt der gegenüberliegenden Seite.

Damit haben wir die:

- Seitenhalbierende durch a und $\frac{1}{2}(b+c)$: $G_{a, \frac{1}{2}(b+c)-a} = \{a + t(\frac{1}{2}(b+c) - a)\}$
- Seitenhalbierende durch b und $\frac{1}{2}(a+c)$: $G_{b, \frac{1}{2}(a+c)-b} = \{b + t(\frac{1}{2}(a+c) - b)\}$
- Seitenhalbierende durch c und $\frac{1}{2}(a+b)$: $G_{c, \frac{1}{2}(a+b)-c} = \{c + t(\frac{1}{2}(a+b) - c)\}$

Bemerkenswerterweise treffen sich die drei Seitenhalbierenden in einem Punkt:



Beweis.

Wir zeigen, dass $\frac{1}{3}(a+b+c)$ auf allen Seitenhalbierenden liegt.

Zu $p \in G_{a, \frac{1}{2}(b+c)-a}$, $q \in G_{b, \frac{1}{2}(a+c)-b}$ und $r \in G_{c, \frac{1}{2}(a+b)-c}$ existieren $t_i \in \mathbb{R}$, so dass

$$\begin{aligned} p &= a + t_0\left(\frac{1}{2}(b+c) - a\right), \\ q &= b + t_1\left(\frac{1}{2}(a+c) - b\right), \\ r &= c + t_2\left(\frac{1}{2}(a+b) - c\right). \end{aligned}$$

Wir wählen $t_0 = t_1 = t_2 = \frac{2}{3}$ und erhalten

$$\begin{aligned} p &= a + \frac{2}{3}\left(\frac{1}{2}(b+c) - a\right) = a + \frac{1}{3}b + \frac{1}{3}c - \frac{2}{3}a = \frac{1}{3}(a+b+c), \\ q &= b + \frac{2}{3}\left(\frac{1}{2}(a+c) - b\right) = b + \frac{1}{3}a + \frac{1}{3}c - \frac{2}{3}b = \frac{1}{3}(a+b+c), \\ r &= c + \frac{2}{3}\left(\frac{1}{2}(a+b) - c\right) = c + \frac{1}{3}a + \frac{1}{3}b - \frac{2}{3}c = \frac{1}{3}(a+b+c). \end{aligned}$$

Die uniforme Wahl der t_i mag überraschen, jedoch ist der Ausdruck $\frac{1}{3}(a+b+c)$ symmetrisch in a, b, c . Daher folgt die Behauptung bereits, indem man schlicht die Bezeichnung der Ecken vertauscht. \square

Definition 2.4.25. Der Punkt $\frac{1}{3}(a+b+c)$ heißt **Schwerpunkt** des Dreiecks (a, b, c) .

Tatsächlich entspricht dieser Schwerpunkt auch dem physikalischen Schwerpunkt des Dreiecks.

Bemerkung 2.4.26. Die Seitenhalbierenden dritteln sich, d.h. sie teilen sich im Verhältnis $2 : 1$, also in Formeln:

$$d\left(a, \frac{1}{3}(a+b+c)\right) = 2 \cdot d\left(\frac{1}{2}(b+c), \frac{1}{3}(a+b+c)\right)$$

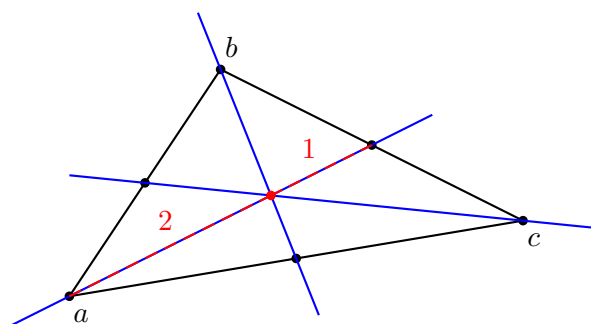


Abb. 24

Wir berechnen hierzu einfach den Abstand der Ecken a, b, c zum Schwerpunkt $\frac{1}{3}(a+b+c)$,

2 Matrixrechnung

etwa:

$$\begin{aligned}d\left(a, \frac{1}{3}(a+b+c)\right) &= \left\| a - \frac{1}{3}(a+b+c) \right\| \\ &= \left\| \frac{2}{3}a - \frac{1}{3}b - \frac{1}{3}c \right\| \\ &= \frac{1}{3} \|2a - b - c\| \\ &= 2 \cdot \frac{1}{6} \|2a - b - c\|\end{aligned}$$

Ferner gilt für den Abstand des gegenüberliegenden Seitenmittelpunktes $\frac{1}{2}(b+c)$ zum Schwerpunkt $\frac{1}{3}(a+b+c)$:

$$\begin{aligned}d\left(\frac{1}{2}(b+c), \frac{1}{3}(a+b+c)\right) &= \left\| \frac{1}{2}(b+c) - \frac{1}{3}(a+b+c) \right\| \\ &= \left\| \frac{1}{2}b + \frac{1}{2}c - \frac{1}{3}a - \frac{1}{3}b - \frac{1}{3}c \right\| \\ &= \left\| -\frac{1}{3}a + \frac{1}{6}b + \frac{1}{6}c \right\| \\ &= \left\| \left(-\frac{1}{6}\right)(2a - b - c) \right\| \\ &= \frac{1}{6} \|2a - b - c\|\end{aligned}$$

Zur Komplettierung dieses Abschnitts wollen wir noch das Konzept von Winkeln, genauer gesagt, Winkelgrößen kennenlernen. Hierfür benötigen wir folgende zwei Beobachtungen. Einerseits kann für $x, y \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$ der Zähler des Quotienten

$$\left| \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \right| = \frac{|\langle x, y \rangle|}{\|x\| \cdot \|y\|}$$

mittels der Cauchy-Schwarz-Ungleichung $|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$ abgeschätzt werden, so dass gilt:

$$\frac{|\langle x, y \rangle|}{\|x\| \cdot \|y\|} \leq \frac{\|x\| \cdot \|y\|}{\|x\| \cdot \|y\|} = 1$$

Also ist:

$$\frac{|\langle x, y \rangle|}{\|x\| \cdot \|y\|} \in [-1, 1]$$

Andererseits kennen wir aus der Analysis die Kosinus-Funktion

$$\cos : [0, \pi] \rightarrow [-1, 1],$$

welche auf dem Intervall $[0, \pi]$ sowohl surjektiv als auch injektiv, also bijektiv ist.

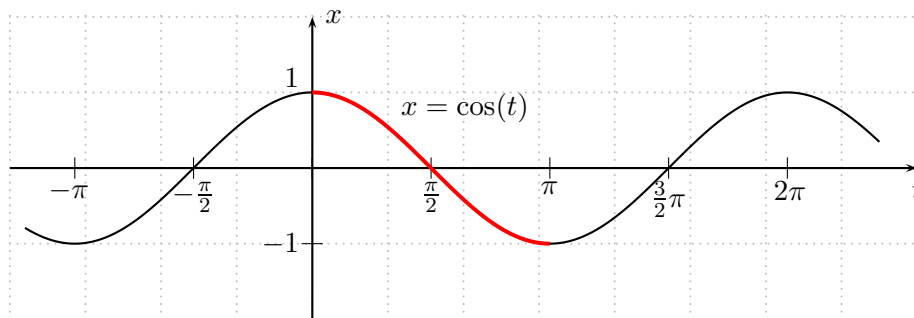


Abb. 25

Wie wir wissen, existiert daher eine Umkehrabbildung:

$$\arccos : [-1, 1] \rightarrow [0, \pi]$$

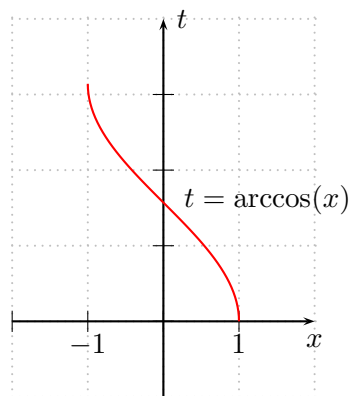


Abb. 26

Somit ist $\cos(t) = x \Leftrightarrow \arccos(x) = t$ und wir können folgende Definition aussprechen.

Definition 2.4.27. Die Zahl

$$\sphericalangle(x, y) := \arccos \left(\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \right)$$

heißt **Innenwinkel** von x und y .

2 Matrixrechnung

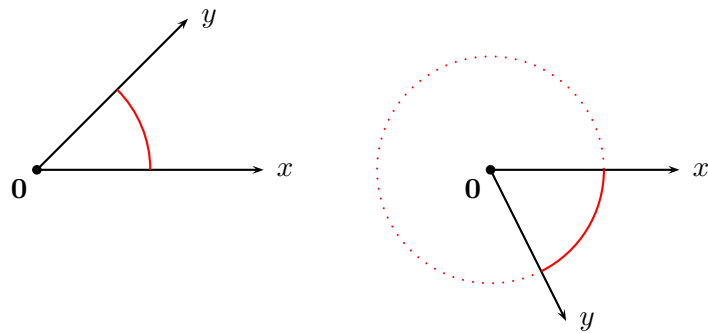


Abb. 27

Der Kosinus des Innenwinkels von x und y entspricht also gerade dem Quotienten $\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}$. Insbesondere wird der Innenwinkel demnach durch das Skalarprodukt definiert. Zu beachten ist, dass auch wirklich stets der Innenwinkel genommen wird, d.h. derjenige Winkel kleiner oder gleich π .

Winkel werden hier immer im **Bogenmaß** angegeben. Das Bogenmaß ist proportional zum Winkelmaß in Grad. Die Umrechnung in das **Gradmaß** kann mittels der Relation

$$1^\circ = \frac{\pi}{180}$$

erfolgen. Sehen wir uns ein paar Beispiele an.

Beispiel 2.4.28. Für $x = y$ gilt:

$$\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} = \frac{\langle x, x \rangle}{\|x\| \cdot \|x\|} = \frac{\|x\|^2}{\|x\|^2} = 1$$

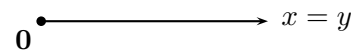


Abb. 28

Folglich ist:

$$\sphericalangle(x, x) = \arccos(1) = 0$$

Beispiel 2.4.29. Für $x = -y$ gilt:

$$\begin{aligned} \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} &= \frac{\langle x, -x \rangle}{\|x\| \cdot \|-x\|} = \frac{-\langle x, x \rangle}{\|x\| \cdot \|x\|} \\ &= \frac{-\|x\|^2}{\|x\|^2} = -1. \end{aligned}$$

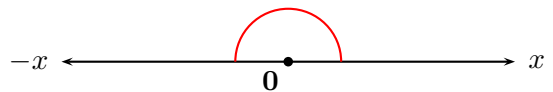


Abb. 29

Folglich ist:

$$\sphericalangle(x, -x) = \arccos(-1) = \pi = 180^\circ$$

Beispiel 2.4.30. Ist $\langle x, y \rangle = 0$, so gilt:

$$\sphericalangle(x, y) = \arccos(0) = \frac{\pi}{2} = 90^\circ$$

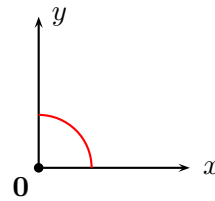


Abb. 30

Definition 2.4.31. Falls $\langle x, y \rangle = 0$ gilt, so heißen x und y **orthogonal** zueinander. Wir sagen auch, dass x und y **aufeinander senkrecht** stehen. Wir schreiben in diesem Fall $x \perp y$.

Eine nützliche Anwendung ist folgender Satz, mit dem wir z.B. aus vorgegebenen drei Seiten eines nicht entarteten euklidischen Dreiecks die Innenwinkel desselben berechnen können. Sind dagegen zwei Seiten und der durch diese Seiten eingeschlossene Winkel vorgegeben, so kann mit dem Satz die dem Winkel gegenüberliegende Seite berechnet werden.

Satz 2.4.32 (Kosinussatz)

Sei (a, b, c) ein nicht entartetes Dreieck und sei $\alpha := \sphericalangle(b - a, c - a)$ der Innenwinkel in der Ecke α . Dann gilt:

$$d(b, c)^2 = d(a, b)^2 + d(a, c)^2 - 2 \cdot d(a, b) \cdot d(a, c) \cdot \cos(\alpha)$$

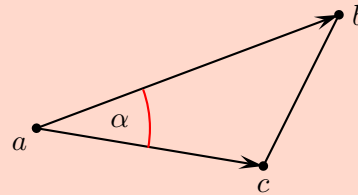


Abb. 31

Entsprechend gelten dann natürlich für $\beta := \sphericalangle(a - b, c - b)$ und $\gamma := \sphericalangle(a - c, b - c)$ die analogen Aussagen

$$d(a, c)^2 = d(b, a)^2 + d(b, c)^2 - 2 \cdot d(b, a) \cdot d(b, c) \cdot \cos(\beta),$$

$$d(a, b)^2 = d(c, a)^2 + d(c, b)^2 - 2 \cdot d(c, a) \cdot d(c, b) \cdot \cos(\gamma).$$

Beweis.

Wir weisen die umgestellte Version

$$d(b, c)^2 - d(a, b)^2 - d(a, c)^2 = -2 \cdot d(a, b) \cdot d(a, c) \cdot \cos(\alpha)$$

2 Matrixrechnung

nach. Es gilt:

$$\begin{aligned}d(b, c)^2 - d(a, b)^2 - d(a, c)^2 &= \|b - c\|^2 - \|a - b\|^2 - \|a, c\|^2 \\&= \langle b - c, b - c \rangle - \langle a - b, a - b \rangle - \langle a - c, a - c \rangle \\&= \langle b, b \rangle - 2 \langle b, c \rangle + \langle c, c \rangle - (\langle a, a \rangle - 2 \langle a, b \rangle + \langle b, b \rangle) \\&\quad - (\langle a, a \rangle - 2 \langle a, c \rangle + \langle c, c \rangle) \\&= -2 \langle a, a \rangle - 2 \langle b, c \rangle + 2 \langle a, b \rangle + 2 \langle a, c \rangle \\&= -2 \cdot (\langle c, b \rangle - \langle a, b \rangle + \langle a, a \rangle - \langle c, a \rangle) \\&= -2 \cdot (\langle c - a, b \rangle + \langle a - c, a \rangle) \\&= -2 \cdot (\langle c - a, b \rangle + \langle c - a, -a \rangle) \\&= -2 \cdot \langle c - a, b - a \rangle \\&= -2 \cdot \|c - a\| \|b - a\| \cdot \frac{\langle c - a, b - a \rangle}{\|c - a\| \|b - a\|} \\&= -2 \cdot d(a, c) \cdot d(a, b) \cdot \cos(\alpha)\end{aligned}$$

□

Folgerung 2.4.33

Die Seitenlängen eines nicht entarteten Dreiecks legen die Innenwinkel fest.

Die umgekehrte Aussage ist falsch, denn Multiplikation der Seitenlängen mit einem Faktor ungleich 0 liefern ein Dreieck mit gleichen Winkelgrößen.

Folgerung 2.4.34 (Satz des Pythagoras)

Ist $\alpha = \frac{\pi}{2}$, so gilt

$$d(b, c)^2 = d(a, b)^2 + d(a, c)^2.$$

Nun noch zu seinem Satz über Parallelogramme.

Satz 2.4.35 (Rhombensatz)

Die vier Seitenlängen eines nicht entarteten Parallelogramms sind genau dann gleich, wenn die beiden Diagonalen sich senkrecht schneiden.

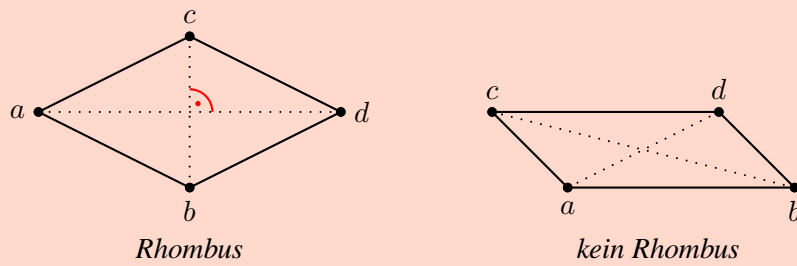


Abb. 32

Beweis.

Für ein nicht entartetes Parallelogramm (a, b, c, d) setze:

$$v := b - a \quad \text{und} \quad w := c - a$$

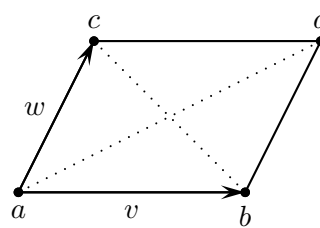


Abb. 33

Damit sind die Diagonalen des Parallelogramms gegeben durch

$$d - a = v + w \quad \text{und} \quad c - b = v - w$$

Um schließlich den Innenwinkel $\sphericalangle(v + w, v - w)$ der beiden Diagonalen zu erhalten, berechnen wir:

$$\begin{aligned} \langle v + w, v - w \rangle &= \langle d - a, c - b \rangle & (2.9) \\ &= \langle d, c \rangle - \langle a, c \rangle + \langle a, b \rangle - \langle d, c \rangle \end{aligned}$$

Wegen der Gültigkeit von $d - c = b - a$ und $d - b = c - a$ in jedem Parallelogramm haben wir andererseits:

$$\begin{aligned} \|v\|^2 - \|w\|^2 &= \langle v, v \rangle - \langle w, w \rangle & (2.10) \\ &= \langle b - a, b - a \rangle - \langle c - a, c - a \rangle \\ &= \langle d - c, b - a \rangle - \langle d - b, c - a \rangle \\ &= \langle d, b - a \rangle - \langle c, b - a \rangle - \langle d, c - a \rangle + \langle b, c - a \rangle \\ &= \langle d, b \rangle - \langle d, a \rangle - \langle c, b \rangle + \langle c, a \rangle - \langle d, c \rangle + \langle d, a \rangle + \langle b, c \rangle - \langle b, a \rangle \\ &= \langle d, b \rangle + \langle c, a \rangle - \langle d, c \rangle - \langle b, a \rangle \\ &= -\langle d - a, c - b \rangle \end{aligned}$$

2 Matrixrechnung

Offenbar ergibt die Bedingung an den Innenwinkel (2.9) bis auf ein Vorzeichen denselben Ausdruck wie die Längendifferenz (2.10). Für den Fall, dass (2.9) und (2.10) gleich 0 sind, spielt jedoch das Vorzeichen keine Rolle. Damit stehen die Diagonalen genau dann senkrecht aufeinander, wenn gilt:

$$\begin{aligned} -\langle d-a, c-b \rangle &= 0 \\ \Leftrightarrow \|v\|^2 - \|w\|^2 &= 0 \\ \Leftrightarrow \|v\| &= \|w\| \end{aligned}$$

Die Parallelogramm-Eigenschaft liefert die Äquivalenz zur Gleichheit aller Seitenlängen. \square

Definition 2.4.36. Ein nicht entartetes Parallelogramm mit gleich langen Seiten heißt **Rhombus** bzw. **Raute**.

2.5 Die komplexen Zahlen

Die natürlichen Zahlen bekommen wir durch Zählen. Wir betrachten \mathbb{N} daher als gegeben. Ein Problem, welches mit natürlichen Zahlen auftreten kann, ist z.B. die Lösung der Gleichung $x + 5 = 0$. Diese besitzt in \mathbb{N} keine Lösungen. Die ganzen Zahlen beseitigen diesen Mangel. Allerdings sind in \mathbb{Z} Gleichungen, wie etwa $2x - 1 = 0$ nicht lösbar. Dies führt uns zur Menge der rationalen Zahlen \mathbb{Q} . Nun gibt es jedoch auch hier wiederum Problemgleichungen, z.B. $x^2 - 2 = 0$. Schließlich ist in der nächsten Erweiterung, den reellen Zahlen \mathbb{R} schon die Gleichung $x^2 + 1 = 0$ nicht lösbar. Hierin besteht unser Motiv für die Konstruktion der komplexen Zahlen. Wir konstruieren diese „neuen“ Zahlen, indem wir

$$\mathbb{C} := \mathbb{R}^2$$

setzen. Die Menge der komplexen Zahlen \mathbb{C} heißt daher auch **Gauß'sche Zahlenebene**.

Die Addition komplexer Zahlen ist die gewöhnliche Vektoraddition, d.h. für alle $x, y \in \mathbb{C}$ gilt:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}$$

Ferner kennen wir bereits das

1. Assoziativgesetz der Addition: $\forall x, y, z \in \mathbb{C} : (x + y) + z = x + (y + z)$
2. Kommutativgesetz der Addition: $\forall x, y, z \in \mathbb{C} : x + y = y + x$
3. neutrale Element der Addition: $\forall x \in \mathbb{C} : x + \mathbf{0} = x$, wobei $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \mathbb{C}$

4. inverse Element der Addition: $\forall x \in \mathbb{C} : x + (-x) = \mathbf{0}$

Eine entscheidende Neuerung bildet die Multiplikation zweier komplexer Zahlen. Für alle $x, y \in \mathbb{C}$ definieren wir:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} := \begin{pmatrix} x_1 \cdot y_1 - x_2 \cdot y_2 \\ x_1 \cdot y_2 + x_2 \cdot y_1 \end{pmatrix}$$

Wir verifizieren, dass die Multiplikation im Wesentlichen den Rechenregeln der reellen Zahlen genügt und prüfen dafür die Gesetze (1.) bis (4.) nach:

1. Assoziativgesetz der Multiplikation: $\forall x, y, z \in \mathbb{C} : (x \cdot y) \cdot z = x \cdot (y \cdot z)$

2. Kommutativgesetz der Multiplikation: $\forall x, y, z \in \mathbb{C} : x \cdot y = y \cdot x$

3. neutrales Element der Multiplikation: $\forall x \in \mathbb{C} : x \cdot \mathbf{1} = x$, wobei $\mathbf{1} := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{C}$

4. inverses Element der Multiplikation: $\forall x \in \mathbb{C} \setminus \{\mathbf{0}\} : x \cdot x^{-1} = \mathbf{1}$, wobei

$$x^{-1} := \frac{1}{x_1^2 + x_2^2} \cdot \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} \in \mathbb{C}$$

ist. Man beachte, dass wegen $x \neq \mathbf{0}$ auch $x_1^2 + x_2^2 > 0$ und damit insbesondere $x_1^2 + x_2^2 \neq 0$ gilt.

Die „Verträglichkeit“ von Addition und Multiplikation ist Aussage eines weiteren Gesetzes, welches wir bereits von den reellen Zahlen kennen. Es stellt sicher, dass man ausmultiplizieren bzw. ausklammern darf:

5. Distributivgesetz: $\forall x, y, z \in \mathbb{C} : x \cdot (y + z) = x \cdot y + x \cdot z$

Beweis.

Zu 1.)

$$\begin{aligned} (x \cdot y) \cdot z &= \begin{pmatrix} x_1 y_1 - x_2 y_2 \\ x_1 y_2 + x_2 y_1 \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\ &= \begin{pmatrix} (x_1 y_1 - x_2 y_2) z_1 - (x_1 y_2 + x_2 y_1) z_2 \\ (x_1 y_1 - x_2 y_2) z_2 + (x_1 y_2 + x_2 y_1) z_1 \end{pmatrix} \\ &= \begin{pmatrix} x_1 y_1 z_1 - x_2 y_2 z_1 - x_1 y_2 z_2 - x_2 y_1 z_2 \\ x_1 y_1 z_2 - x_2 y_2 z_2 + x_1 y_2 z_1 + x_2 y_1 z_1 \end{pmatrix} \\ &= \begin{pmatrix} x_1 y_1 z_1 - x_1 y_2 z_2 - x_2 y_1 z_2 - x_2 y_2 z_1 \\ x_1 y_1 z_2 + x_1 y_2 z_1 + x_2 y_1 z_1 - x_2 y_2 z_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 (y_1 z_1 - y_2 z_2) - x_2 (y_1 z_2 + y_2 z_1) \\ x_1 (y_1 z_2 + y_2 z_1) + x_2 (y_1 z_1 - y_2 z_2) \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} y_1 z_1 - y_2 z_2 \\ y_1 z_2 + y_2 z_1 \end{pmatrix} \\ &= x \cdot (y \cdot z) \end{aligned}$$

2 Matrixrechnung

Zu 2.)

$$\begin{aligned}x \cdot y &= \begin{pmatrix} x_1 y_1 - x_2 y_2 \\ x_1 y_2 + x_2 y_1 \end{pmatrix} \\ &= \begin{pmatrix} y_1 x_1 - y_2 x_2 \\ y_2 x_1 + y_1 x_2 \end{pmatrix} \\ &= \begin{pmatrix} y_1 x_1 - y_2 x_2 \\ y_1 x_2 + y_2 x_1 \end{pmatrix} \\ &= x \cdot y\end{aligned}$$

Zu 3.)

$$\begin{aligned}x \cdot \mathbf{1} &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \cdot 1 - x_2 \cdot 0 \\ x_1 \cdot 0 + x_2 \cdot 1 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= x\end{aligned}$$

Zu 4.)

$$\begin{aligned}x \cdot x^{-1} &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} \frac{x_1}{x_1^2+x_2^2} \\ \frac{-x_2}{x_1^2+x_2^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{x_1^2}{x_1^2+x_2^2} - \frac{-x_2^2}{x_1^2+x_2^2} \\ x_1 \frac{-x_2}{x_1^2+x_2^2} + x_2 \frac{x_1}{x_1^2+x_2^2} \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \mathbf{1}\end{aligned}$$

Zu 5.)

$$\begin{aligned}x \cdot (y + z) &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} y_1 + z_1 \\ y_2 + z_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 \cdot (y_1 + z_1) - x_2 \cdot (y_2 + z_2) \\ x_1 \cdot (y_2 + z_2) + x_2 \cdot (y_1 + z_1) \end{pmatrix} \\ &= \begin{pmatrix} x_1 y_1 + x_1 z_1 - x_2 y_2 - x_2 z_2 \\ x_1 y_2 + x_1 z_2 + x_2 y_1 + x_2 z_1 \end{pmatrix} \\ &= \begin{pmatrix} x_1 y_1 - x_2 y_2 + x_1 z_1 - x_2 z_2 \\ x_1 y_2 + x_2 y_1 + x_1 z_2 + x_2 z_1 \end{pmatrix} \\ &= \begin{pmatrix} x_1 y_1 - x_2 y_2 \\ x_1 y_2 + x_2 y_1 \end{pmatrix} + \begin{pmatrix} x_1 z_1 - x_2 z_2 \\ x_1 z_2 + x_2 z_1 \end{pmatrix} \\ &= x \cdot y + x \cdot z\end{aligned}$$

□

Wir sprachen einführend über die Erweiterung der Zahlbereiche. Dabei war jeder Zahlbereich im nächstgrößeren enthalten. Genauso wollen wir dies nun im Fall der reellen und komplexen Zahlen organisieren. Die reelle Achse (der reelle Zahlenstrahl) wird hierzu in die Gauß'sche Zahlenebene eingebettet und bildet die so genannte **reelle Achse**. Die andere, dazu senkrechte Achse heißt **imaginäre Achse** und besitzt die Grundeinheit:

$$i := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Wie man sich nun denken kann, steht i dabei für *imaginär* und tatsächlich hing der ursprüngliche Gedanke damit zusammen, sich i einfach als eine nicht reelle Zahl vorzustellen. Nichtsdestoweniger sind komplexe Zahlen durchaus real und besitzen vielfältige Anwendungen, etwa in der Physik.

Sehen wir, was uns die Einführung von imaginären Zahlen einbringt. Hierfür quadrieren wir:

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

Mit dieser Erkenntnis können wir nun jede komplexe Zahl $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{C}$ wie folgt notieren:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \end{pmatrix} = x_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = x_1 \cdot \mathbf{1} + x_2 \cdot i$$

Hiermit lässt sich die komplexe Multiplikation leicht merken, denn für $x, y \in \mathbb{C}$ gilt:

$$\begin{aligned} x \cdot y &= (x_1 \cdot \mathbf{1} + x_2 \cdot i) \cdot (y_1 \cdot \mathbf{1} + y_2 \cdot i) \\ &= x_1 y_1 \cdot \mathbf{1} + x_1 y_2 \cdot i + x_2 y_1 \cdot i + x_2 y_2 \cdot i^2 \\ &= (x_1 y_1 - x_2 y_2) \cdot \mathbf{1} + (x_1 y_2 + x_2 y_1) \cdot i \end{aligned}$$

Von nun an identifizieren wir $\mathbf{0} \in \mathbb{C}$ mit $0 \in \mathbb{R}$ sowie $\mathbf{1} \in \mathbb{C}$ mit $1 \in \mathbb{R}$ und schreiben kurz

$$x = x_1 + x_2 \cdot i \quad \text{bzw. gelegentlich auch} \quad x = x_1 + ix_2.$$

Somit kann jede reelle Zahl t mit der komplexen Zahl $t = t + 0 \cdot i$ identifiziert werden, also gilt

$$\mathbb{R} \subset \mathbb{C}.$$

Oft werden z und w statt x und y als Platzhalter für komplexe Zahlen verwendet. Auch wir werden uns diesem Wandel nun unterwerfen.

Definition 2.5.1. Für $z = z_1 + iz_2 \in \mathbb{C}$ heißt $z_1 =: \operatorname{Re}(z)$ **Realteil** und $z_2 =: \operatorname{Im}(z)$ **Imaginärteil** von z .

2 Matrixrechnung

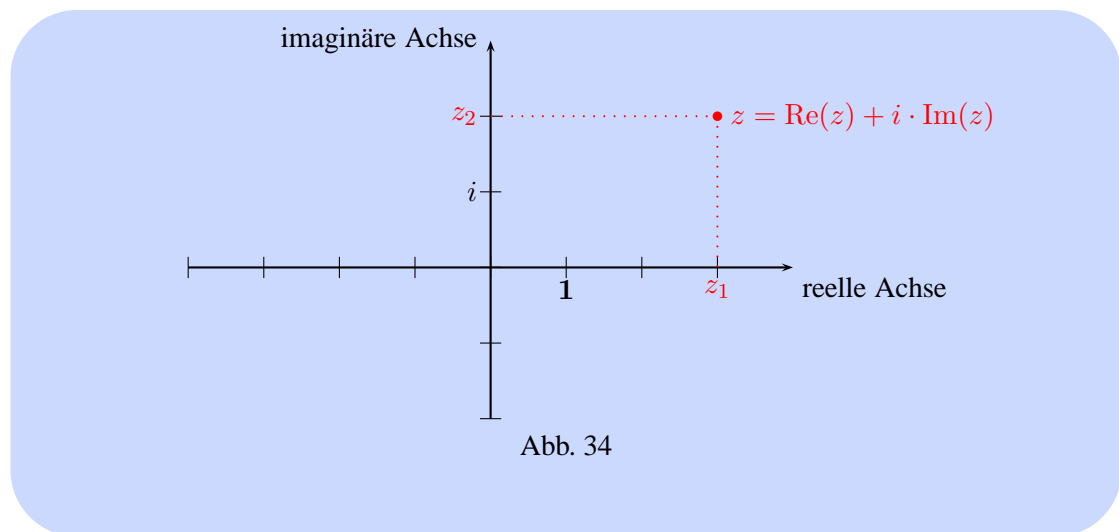


Abb. 34

Insbesondere sind für $z \in \mathbb{C}$ die Werte $\operatorname{Re}(z)$ und $\operatorname{Im}(z)$ reelle Zahlen.

Beispiel 2.5.2. Wir betrachten $z = 1 + i\sqrt{2}$ sowie $w = 2 + i\sqrt{3}$ und berechnen

$$\begin{aligned}\frac{z}{w} &= \frac{1 + i\sqrt{2}}{2 + i\sqrt{3}} \\ &= \frac{(1 + i\sqrt{2}) \cdot (2 - i\sqrt{3})}{(2 + i\sqrt{3}) \cdot (2 - i\sqrt{3})} \\ &= \frac{2 - i\sqrt{3} + i2\sqrt{2} + \sqrt{2}\sqrt{3}}{4 - (i\sqrt{3})^2} \\ &= \frac{(2 + \sqrt{6}) + i(2\sqrt{2} - \sqrt{3})}{4 + 3} \\ &= \frac{2 + \sqrt{6}}{7} + i\frac{2\sqrt{2} - \sqrt{3}}{7},\end{aligned}$$

also gilt

$$\operatorname{Re}\left(\frac{z}{w}\right) = \frac{2 + \sqrt{6}}{7} \quad \text{und} \quad \operatorname{Im}\left(\frac{z}{w}\right) = \frac{2\sqrt{2} - \sqrt{3}}{7}.$$

Die im letzten Beispiel verwendete Umformung, das Erweitern des Bruches $\frac{z}{w}$ mit derjenigen komplexen Zahl, bei der der Imaginärteil das entgegengesetzte Vorzeichen (des Imaginärteils von w) besitzt, erhält folgenden Namen.

Definition 2.5.3. Für $z = z_1 + iz_2 \in \mathbb{C}$ heißt $\bar{z} := z_1 - iz_2$ die **komplexe Konjugation** zu z .

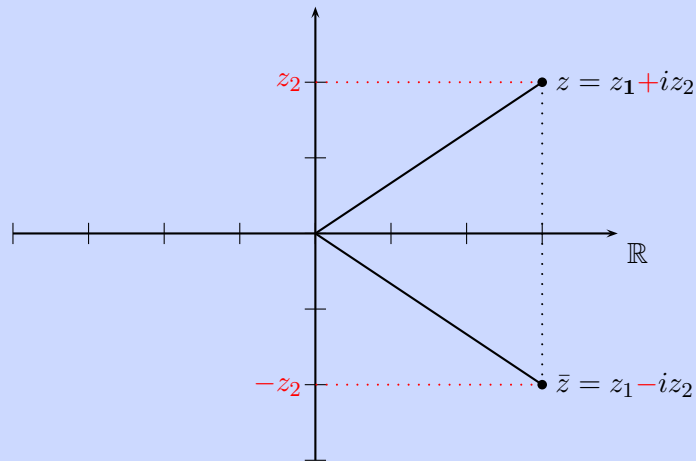


Abb. 35

In geometrischer Hinsicht ist die Abbildung $z \mapsto \bar{z}$ die Spiegelung an der reellen Achse $\mathbb{R} = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\}$. Insbesondere gilt für alle $z = z_1 + iz_2 \in \mathbb{C}$:

1. Die komplexe Konjugation ist selbstinvers, d.h. es gilt:

$$\overline{(\bar{z})} = \overline{\overline{(z_1 + iz_2)}} = \overline{(z_1 - iz_2)} = z_1 + iz_2 = z.$$

2. Es ist $z = \bar{z} \Leftrightarrow z \in \mathbb{R}$.

Satz 2.5.4

Für alle $z, w \in \mathbb{C}$ gilt:

1. $\overline{z + w} = \bar{z} + \bar{w}$

2. $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$

3. $z \cdot \bar{z} = |z|^2$, wobei $|z| = \|z\| = \sqrt{(\operatorname{Re}(z))^2 + (\operatorname{Im}(z))^2}$

2 Matrixrechnung

Beweis.

Zu 1.)

Es gilt

$$\begin{aligned}\overline{z+w} &= \overline{\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}} \\ &= \overline{\begin{pmatrix} z_1 + w_1 \\ z_2 + w_2 \end{pmatrix}} \\ &= \begin{pmatrix} z_1 + w_1 \\ -(z_2 + w_2) \end{pmatrix} \\ &= \begin{pmatrix} z_1 + w_1 \\ -z_2 - w_2 \end{pmatrix} \\ &= \begin{pmatrix} z_1 \\ -z_2 \end{pmatrix} + \begin{pmatrix} w_1 \\ -w_2 \end{pmatrix} \\ &= \bar{z} + \bar{w}\end{aligned}$$

Zu 2.)

Für komplexe Zahlen $z = z_1 + iz_2$ und $w = w_1 + iw_2$ bekommen wir:

$$\begin{aligned}\overline{z \cdot w} &= \overline{(z_1 + iz_2) \cdot (w_1 + iw_2)} \\ &= \overline{z_1w_1 - z_2w_2 + i(z_1w_2 + z_2w_1)} \\ &= z_1w_1 - z_2w_2 - i(z_1w_2 + z_2w_1) \\ &= (z_1 - iz_2) \cdot (w_1 - iw_2) \\ &= \bar{z} \cdot \bar{w}\end{aligned}$$

Zu 3.)

Wir berechnen:

$$\begin{aligned}z \cdot \bar{z} &= (z_1 + iz_2) \cdot (z_1 - iz_2) \\ &= z_1^2 - (iz_2)^2 \\ &= z_1^2 - i^2 z_2^2 \\ &= z_1^2 + z_2^2 \\ &= |z|^2\end{aligned}$$

□

Bemerkung 2.5.5. Im Beispiel (2.5.2) erweiterten wir den Bruch $\frac{z}{w}$ mit der komplexen Konjugation \bar{w} des Nenners, mit der Absicht, einen reellen Nenner zu erhalten. Wie sich nun herausstellt, geschah dies implizit mittels der Regel (3.), denn es gilt

$$\frac{z}{w} = \frac{z \cdot \bar{w}}{w \cdot \bar{w}} = \frac{z \cdot \bar{w}}{|w|^2},$$

wobei $|w|^2 \in \mathbb{R}$ ist.

Blicken wir zurück: Für die Operationen Addition und Konjugation haben wir bereits geometrische Interpretation geliefert. Für die Multiplikation komplexer Zahlen allerdings fehlt uns eine solche. Wir benötigen hierzu eine weitere Darstellungsform komplexer Zahlen z , die durch den komplexen Betrag $|z|$ und folgenden Winkelbegriff festgelegt ist.

Definition 2.5.6. Sei $z \in \mathbb{C} \setminus \{0\}$. Dann heißt

$$\arg(z) := \begin{cases} \sphericalangle(\mathbf{1}, z) & , \text{ falls } \operatorname{Im}(z) \geq 0 \\ 2\pi - \sphericalangle(\mathbf{1}, z) & , \text{ falls } \operatorname{Im}(z) < 0 \end{cases}$$

das **Argument** von z .

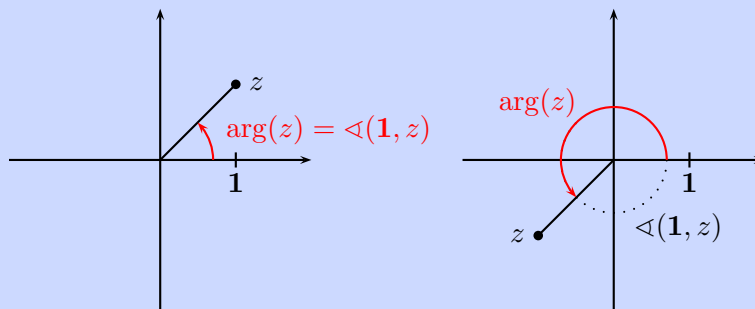


Abb. 36

Damit können wir für komplexe Zahlen $z = \operatorname{Re}(z) + i \cdot \operatorname{Im}(z)$ formulieren:

Satz 2.5.7

Für jedes $z \in \mathbb{C} \setminus \{0\}$ gilt:

1. $\operatorname{Re}(z) = |z| \cdot \cos(\arg(z))$
2. $\operatorname{Im}(z) = |z| \cdot \sin(\arg(z))$

Beweis.

Zu 1.)

Die Definition von $\arg(z)$ gebietet die Betrachtung zweier Fälle:

2 Matrixrechnung

1. Fall: Für $\text{Im}(z) \geq 0$ gilt:

$$\begin{aligned}
 |z| \cdot \cos(\arg(z)) &= |z| \cdot \cos(\angle(\mathbf{1}, z)) \\
 &= |z| \cdot \cos\left(\arccos\left(\frac{\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| \cdot \left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\|}\right)\right) \\
 &= |z| \cdot \frac{\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\| \cdot \left\| \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\|} \\
 &= |z| \cdot \frac{\left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\rangle}{1 \cdot |z|} \\
 &= \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\rangle \\
 &= z_1 \\
 &= \text{Re}(z)
 \end{aligned}$$

2. Fall: Für $\text{Im}(z) < 0$ gilt:

Da die Kosinus-Funktion 2π -periodisch und eine gerade Funktion ist, gilt:

$$\begin{aligned}
 |z| \cdot \cos(\arg(z)) &= |z| \cdot \cos(2\pi - \angle(\mathbf{1}, z)) \\
 &= |z| \cdot \cos(-\angle(\mathbf{1}, z)) \\
 &= |z| \cdot \cos(\angle(\mathbf{1}, z)) \\
 &= \text{Re}(z)
 \end{aligned}$$

Zu 2.)

Aus taktischen Gründen setzen wir zunächst:

$$x := \frac{\text{Im}(z)}{|z|} \tag{2.11}$$

Wegen $|z| = \sqrt{(\text{Re}(z))^2 + (\text{Im}(z))^2}$ gilt

$$\begin{aligned}
 |z|^2 &= (\text{Re}(z))^2 + (\text{Im}(z))^2 \\
 &= (\text{Re}(z))^2 + |z|^2 \frac{(\text{Im}(z))^2}{|z|^2} \\
 &= (\text{Re}(z))^2 + |z|^2 x^2 \\
 &\stackrel{(1.)}{=} (|z| \cos(\arg(z)))^2 + |z|^2 x^2 \\
 &= |z|^2 (\cos(\arg(z)))^2 + |z|^2 x^2.
 \end{aligned}$$

Gut, dass $z \neq \mathbf{0}$ in der Voraussetzung des Satzes steht, denn nun kann durch $|z|^2 \neq 0$ dividiert werden. Wir bekommen:

$$1 = (\cos(\arg(z)))^2 + x^2$$

Der trigonometrische Pythagoras liefert uns:

$$1 = (\cos(\arg(z)))^2 + (\sin(\arg(z)))^2$$

Anschließende Subtraktion dieser beiden Gleichungen führt auf

$$x^2 = (\sin(\arg(z)))^2,$$

also

$$x = \sin(\arg(z)) \quad \text{oder} \quad x = -\sin(\arg(z)).$$

Nun wieder zu den beiden möglichen Fällen:

1. Fall: Für $\operatorname{Im}(z) \geq 0$ ist $x = \frac{\operatorname{Im}(z)}{|z|} \geq 0$. Ferner gilt

$$\arg(z) = \sphericalangle(\mathbf{1}, z) \in [0, \pi]$$

und dem Verlauf der Sinus-Funktion nach damit $\sin(\arg(z)) \geq 0$, also $x = \sin(\arg(z))$. Umstellen von (2.11) und Einsetzen für x ergibt wie gewünscht

$$\operatorname{Im}(z) = x \cdot |z| = \sin(\arg(z)) \cdot |z|.$$

2. Fall: Für $\operatorname{Im}(z) < 0$ ist $x = \frac{\operatorname{Im}(z)}{|z|} < 0$. Weiterhin gilt

$$\arg(z) = 2\pi - \sphericalangle(\mathbf{1}, z) \in [\pi, 2\pi],$$

also dieses Mal $\sin(\arg(z)) \leq 0$ und damit wie im 1. Fall $x = \sin(\arg(z))$. Also folgt wie oben:

$$\operatorname{Im}(z) = |z| \cdot \sin(\arg(z)).$$

□

Im Ergebnis können wir nun komplexe Zahlen $z = \operatorname{Re}(z) + i \cdot \operatorname{Im}(z)$ auch in der Form

$$z = |z| \cdot (\cos(\arg(z)) + i \sin(\arg(z)))$$

schreiben. In der Literatur wird oft abkürzend $\varphi := \arg(z) \in [0, 2\pi)$ gesetzt.

2 Matrixrechnung

Die Darstellung

$$z = |z| \cdot (\cos(\varphi) + i \sin(\varphi))$$

nennt man die **Polardarstellung** für komplexe Zahlen. Sie ist durch den komplexen Betrag $|z|$ und den Winkel φ (ggf. abzüglich $k \cdot 2\pi$ für ein $k \in \mathbb{Z}$) eindeutig bestimmt.

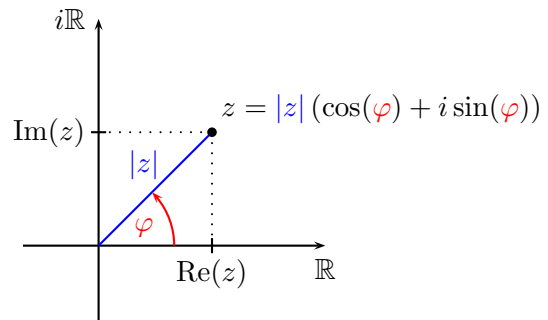


Abb. 37

Nun endlich spendiert uns folgender Satz eine geometrische Interpretation für die Multiplikation:

Satz 2.5.8

Für alle $z, w \in \mathbb{C} \setminus \{0\}$ gilt:

1. $|z \cdot w| = |z| \cdot |w|$
2. $\arg(z \cdot w) = \begin{cases} \arg(z) + \arg(w) & , \text{ falls } \arg(z) + \arg(w) < 2\pi \\ \arg(z) + \arg(w) - 2\pi & , \text{ sonst} \end{cases}$

Beweis.

Zu 1.)

Wegen Satz (2.5.4) gilt:

$$\begin{aligned} |z \cdot w|^2 &= (z \cdot w) \cdot \overline{(z \cdot w)} \\ &= (z \cdot w) \cdot (\bar{z} \cdot \bar{w}) \\ &= (z \cdot \bar{z}) \cdot (w \cdot \bar{w}) \\ &= |z|^2 \cdot |w|^2 \\ &= (|z| \cdot |w|)^2 \end{aligned}$$

Da $|z \cdot w|, |z| \cdot |w| \geq 0$ sind, folgt die Behauptung (1.) durch Wurzelziehen.

Zu 2.)

Wir schreiben abkürzend $\varphi := \arg(z)$ und $\psi := \arg(w)$. Dann gilt

$$\begin{aligned} z &= |z| \cdot (\cos(\varphi) + i \sin(\varphi)) && \text{sowie} \\ w &= |w| \cdot (\cos(\psi) + i \sin(\psi)) \end{aligned}$$

und wir erhalten:

$$\begin{aligned}
 z \cdot w &= |z| \cdot (\cos(\varphi) + i \sin(\varphi)) \cdot |w| \cdot (\cos(\psi) + i \sin(\psi)) \\
 &= |z| \cdot |w| \cdot (\cos(\varphi) + i \sin(\varphi)) \cdot (\cos(\psi) + i \sin(\psi)) \\
 &= |z| \cdot |w| \cdot (\cos(\varphi) \cos(\psi) + i \cos(\varphi) \sin(\psi) + i \sin(\varphi) \cos(\psi) + i^2 \sin(\varphi) \sin(\psi)) \\
 &= |z| \cdot |w| \cdot ((\cos(\varphi) \cos(\psi) - \sin(\varphi) \sin(\psi)) + i(\cos(\varphi) \sin(\psi) + \sin(\varphi) \cos(\psi))) \\
 &\stackrel{(1.)}{=} |z \cdot w| \cdot (\cos(\varphi + \psi) + i \sin(\varphi + \psi))
 \end{aligned}$$

Zudem gehen in die letzte Gleichheit die Additionstheoreme für Kosinus und Sinus ein. Wir werden nochmal darauf zurückkommen.

Falls nun $\varphi + \psi \in [0, 2\pi)$ ist, so gilt

$$\arg(z \cdot w) = \varphi + \psi = \arg(z) + \arg(w).$$

Ist dagegen $\varphi + \psi \geq 2\pi$, so gilt $\varphi + \psi \in [2\pi, 4\pi)$. Wegen der 2π -Periodizität von Kosinus und Sinus erhalten wir dann

$$\begin{aligned}
 z \cdot w &= |z \cdot w| \cdot (\cos(\varphi + \psi) + i \sin(\varphi + \psi)) \\
 &= |z \cdot w| \cdot (\cos(\varphi + \psi - 2\pi) + i \sin(\varphi + \psi - 2\pi))
 \end{aligned}$$

und damit

$$\arg(z \cdot w) = \varphi + \psi - 2\pi \in [0, 2\pi).$$

□

Aus diesem Satz ergibt sich folgende geometrische Interpretation der Multiplikation in \mathbb{C} :

- Die Beträge werden multipliziert: $|z \cdot w| = |z| \cdot |w|$
- Die Argumente werden addiert (und von der Summe ggf. 2π subtrahiert)

Sehen wir uns die beiden Fälle $\arg(z) + \arg(w) \in [0, 2\pi)$ und $\arg(z) + \arg(w) \in [2\pi, 4\pi)$ in je einem Beispiel an:

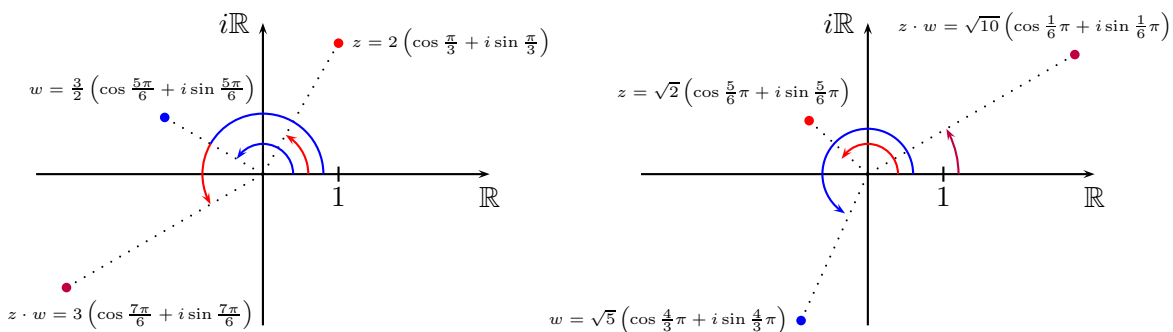


Abb. 38

2 Matrixrechnung

Wir erarbeiten nachfolgend eine Möglichkeit, die eben verwendeten Additionstheoreme herzuleiten. Aus der Analysis ist uns die e -Funktion $\mathbb{R} \rightarrow \mathbb{R}$ mit $x \mapsto e^x$ bekannt:

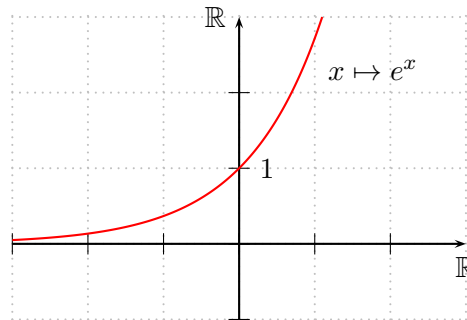


Abb. 39

Für alle $x, x' \in \mathbb{R}$ genügt die e -Funktion der Funktionalgleichung:

$$e^{x+x'} = e^x \cdot e^{x'} \quad (2.12)$$

Im Kontext komplexer Zahlen erhebt sich nun die Frage, ob die e -Funktion zu einer komplexen Funktion $\mathbb{C} \rightarrow \mathbb{C}$ fortsetzbar ist, so dass die Funktionalgleichung (2.12) für alle komplexen Zahlen gültig ist. Ist dies der Fall, so muss für jedes $z = x + iy$ mit $x, y \in \mathbb{R}$ gelten:

$$\begin{aligned} e^z &= e^{x+iy} \\ &= e^x \cdot e^{iy} \end{aligned}$$

Überlegen wir, wie die e -Funktion auf der imaginären Achse definiert werden kann. Wir benötigen also eine Funktion $\mathbb{R} \rightarrow \mathbb{C}$ mit $y \mapsto e^{iy}$, so dass schließlich für alle $y, y' \in \mathbb{R}$ gilt:

$$\begin{aligned} e^{i(y+y')} &= e^{iy+iy'} \\ &= e^{iy} \cdot e^{iy'} \end{aligned} \quad (2.13)$$

Wir definieren:

$$e^{iy} := \cos(y) + i \sin(y) \quad (2.14)$$

Dies liefert eine dritte Darstellungsform $z = |z| \cdot e^{i\varphi}$, die so genannte **Euler'sche Darstellung**⁴ komplexer Zahlen. Hierbei ist $\varphi := \arg(z)$.

⁴benannt nach dem berühmten Mathematiker Leonhard Euler (1707 (Basel, Schweiz) - 1783 (Sankt Petersburg, Russland)), siehe http://de.wikipedia.org/wiki/Leonhard_Euler

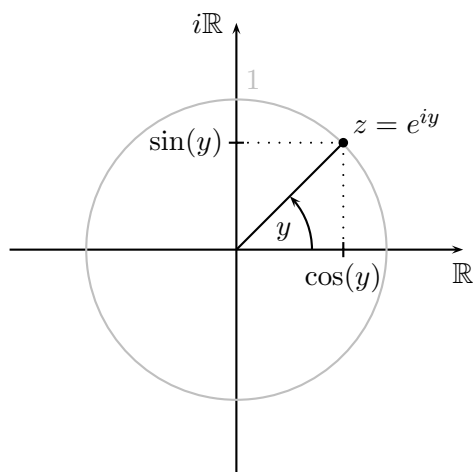


Abb. 40

Die Rechnung aus dem Beweis von Satz (2.5.8) zeigt, dass tatsächlich für alle $y, y' \in \mathbb{R}$ die Funktionalgleichung (2.13) gilt. Mittels (2.14) erhalten wir nun die gewünschte (komplexe) Fortsetzung zu:

$$e^{x+iy} = e^x \cdot (\cos(y) + i \sin(y))$$

Bemerkung 2.5.9. Mit der komplexen e -Funktion lassen sich die Additionstheoreme für Kosinus und Sinus, sollte man sie einmal vergessen haben, wie folgt leicht herleiten. Einerseits gilt nach (2.14)

$$e^{i(y+y')} = \cos(y + y') + i \sin(y + y'). \quad (2.15)$$

Andererseits liefern die Funktionalgleichung (2.13) und (2.14)

$$\begin{aligned} e^{i(y+y')} &= e^{iy} \cdot e^{iy'} \\ &= (\cos(y) + i \sin(y)) \cdot (\cos(y') + i \sin(y')) \\ &= \cos(y) \cos(y') - \sin(y) \sin(y') + i(\sin(y) \cos(y') + \cos(y) \sin(y')). \end{aligned} \quad (2.16)$$

Nun sind zwei komplexe Zahlen genau dann gleich, wenn ihre Real- und Imaginärteile übereinstimmen. Der Vergleich eben dieser in (2.15) und (2.16) ergibt genau die Additionstheoreme für alle $y, y' \in \mathbb{R}$:

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

2 Matrixrechnung

Bemerkung 2.5.10. Im Spezialfall haben wir:

$$e^{i\pi} = -1$$

Dies nennt man auch die **Euler'sche Formel**. Deren Umstellung

$$e^{i\pi} + 1 = 0$$

gilt unter manchen Mathematikern als die schönste Gleichung der Mathematik, da hier die wichtigen mathematischen Größen $0, 1, i, \pi, e$ vorkommen.

Am Anfang dieses Abschnitts standen wir vor dem Problem, bestimmte Gleichungen nicht lösen zu können. Nun haben wir z.B. für $z^2 + 1 = 0$ die Lösungen $z_1 = i$ und $z_2 = -i$ vorzuweisen. Die Anzahl der Lösungen des Polynoms $z \mapsto z^2 + 1$ entspricht also genau seinem Grad. Für komplexe Polynome ist dies immer der Fall. Insbesondere sind in \mathbb{C} alle Polynomgleichungen lösbar. Diese Aussage ist formalisiert im

Satz 2.5.11 (Fundamentalsatz der Algebra)

Sei $n \in \mathbb{N}$. Seien $a_0, \dots, a_n \in \mathbb{C}$ mit $a_0 \neq 0$. Dann existieren $z_1, \dots, z_n \in \mathbb{C}$, so dass für alle $z \in \mathbb{C}$ gilt:

$$a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = a_n \cdot (z - z_1) \cdot \dots \cdot (z - z_n)$$

Dabei ist die linke Seite ein komplexes Polynom vom Grade $n \geq 1$. Die rechte Seite nennt man **Linearfaktorzerlegung** des Polynoms mit Linearfaktoren $(z - z_i)$. Insbesondere sind damit $z = z_1, \dots, z = z_n$ genau die Lösungen der Gleichung:

$$a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = 0$$

In den komplexen Zahlen sind polynomiale Gleichungen also stets lösbar. Jedes komplexe Polynom besitzt eine Linearfaktorzerlegung.

Man beachte, dass die z_1, \dots, z_n nicht verschieden zu sein brauchen, denn etwa für

$$(z - 1)^2 = (z - 1)(z - 1) = z^2 - 2z + 1 = 0$$

erhalten wir die Lösungen $z_1 = z_2 = 1$.

Exakt ausgedrückt hat ein komplexes Polynom n -ten Grades also genau n (nicht notwendigerweise verschiedene) Nullstellen. Daher ist folgende Definition angebracht:

Definition 2.5.12. Für jede Nullstelle ζ von $z \mapsto a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$ heißt die Anzahl der Faktoren $(z - \zeta)$ in der Linearfaktorzerlegung die **Vielfachheit** der Nullstelle ζ .

Zählt man die Anzahl der Nullstellen eines komplexen Polynoms vom Grad n inklusive aller Vielfachheiten zusammen, so erhält man stets genau n .

Beispiel 2.5.13. Wir haben die Linearfaktorzerlegung

$$z \mapsto z^2 + 1 = (z - i)(z - (-i)) = (z - i)(z + i).$$

Also sind $z_1 = i$ und $z_2 = -i$ die beiden Nullstellen, jeweils von der Vielfachheit 1.

Beispiel 2.5.14. Das Polynom

$$z \mapsto z^2 - 2z + 1 = (z - 1)(z - 1)$$

besitzt die Nullstelle $z_1 = 1$ mit der Vielfachheit 2.

Die Konstruktion der Zahlbereiche ist damit in der Tat abgeschlossen. Wir wollen nun abschließend noch das komplexe Analogon zum reellen Wurzelziehen diskutieren und betrachten hierfür die komplexe Gleichung

$$z^n - 1 = 0.$$

Als Spezialfall haben wir etwa

$$\begin{aligned} z^4 - 1 &= (z - 1)(z - i)(z - (-1))(z - (-i)) \\ &= (z - 1)(z - i)(z + 1)(z + i) \end{aligned}$$

Löst $z \in \mathbb{C}$ die Gleichung, d.h. gilt $z^n = 1$, so ist

$$1 = |1| = |z^n| = |z|^n$$

und damit $|z| = 1$. Alle Lösungen der Gleichung $z^n = 1$ liegen somit auf dem Einheitskreis, also dem Kreis um 0 mit Radius 1. Insbesondere ist damit z von der Form $z = e^{iy}$ für ein, noch zu bestimmendes, $y \in \mathbb{R}$. Die Funktionalgleichung (2.13) liefert

$$1 = z^n = (e^{iy})^n = e^{iny} = \cos(ny) + i \sin(ny),$$

2 Matrixrechnung

also muss

$$\cos(ny) = 1 \quad \text{sowie} \quad \sin(ny) = 0$$

gelten. Die $y \in \mathbb{R}$, welche diesen beiden Gleichungen genügen, erhalten wir durch Nullstellenbetrachtung von Sinus und Kosinus. Da der Sinus genau die Vielfachen von π als Nullstellen hat, existiert ein $m \in \mathbb{Z}$, so dass $ny = m \cdot \pi$. Den Wert 1 nimmt der Kosinus bei den geraden Vielfachen von π an, -1 bei den ungeraden Vielfachen. Also ist m von der Form $m = 2k$ für ein $k \in \mathbb{Z}$, so dass $ny = 2k \cdot \pi$ gilt. Wir erhalten folglich insgesamt, dass es ein $k \in \mathbb{Z}$ gibt, so dass

$$y = \frac{2k \cdot \pi}{n}$$

und damit

$$z = e^{i \frac{2k\pi}{n}}$$

gilt.

Nehmen wir umgekehrt $z = e^{i \frac{2k\pi}{n}}$ mit $k \in \mathbb{Z}$ her, so gilt

$$\begin{aligned} z^n &= \left(e^{i \frac{2k\pi}{n}} \right)^n \\ &= e^{i \frac{2k\pi}{n} n} \\ &= e^{i 2k\pi} \\ &= \cos(2k\pi) + i \sin(2k\pi) \\ &= 1, \end{aligned}$$

d.h. die $z \in \mathbb{C}$ von der Form $z = e^{i \frac{2k\pi}{n}}$ sind Lösungen.

Insgesamt haben wir somit alle Lösungen der Gleichung $z^n - 1 = 0$ auf dem Einheitskreis identifiziert. Hier zwei Beispiele für $n = 3$ und $n = 4$:

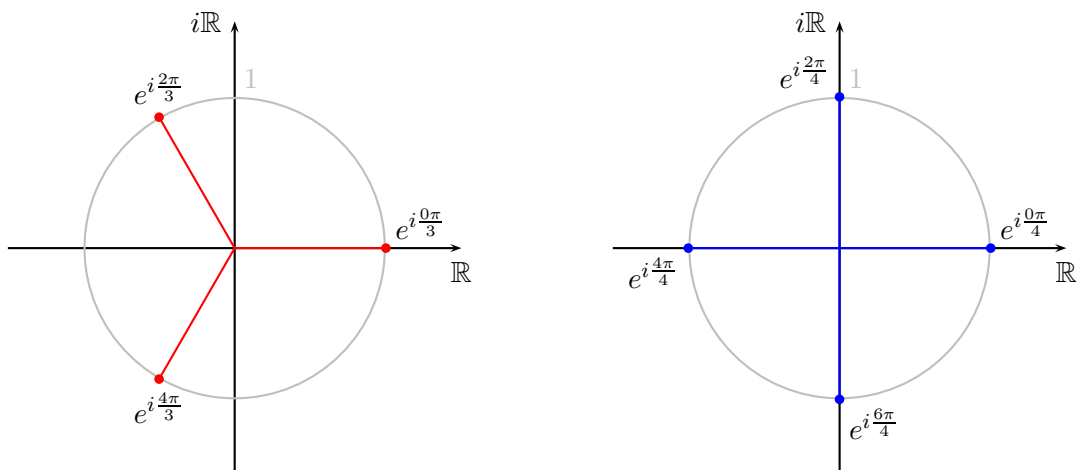


Abb. 41

Da die Nullstellen von $z \mapsto z^n - 1$ den Einheitskreis stets in gleich große Segmente aufteilen, nennt man dieses Polynom auch **Kreisteilungspolynom**.

Definition 2.5.15. Die Lösungen der Gleichung $z^n - 1 = 0$ heißen **n -te Einheitswurzeln**.

Beispiel 2.5.16. Für $n = 3$ ergeben sich die dritten Einheitswurzeln:

$$e^{i\frac{0\pi}{3}} = 1, \quad \rho_+ := e^{i\frac{2\pi}{3}} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i, \quad \rho_- := e^{i\frac{4\pi}{3}} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$$

Sehen wir, was beim Quadrieren, Konjugieren und Aufaddieren passiert:

$$\begin{aligned} \rho_+^2 &= e^{i2 \cdot \frac{2\pi}{3}} = e^{i\frac{4\pi}{3}} = \rho_- \\ \rho_-^2 &= e^{i2 \cdot \frac{4\pi}{3}} = e^{i\frac{8\pi}{3}} = e^{i(\frac{6}{3}\pi + \frac{2}{3}\pi)} = e^{i2\pi} e^{i\frac{2}{3}\pi} = 1 \cdot e^{i\frac{2}{3}\pi} = \rho_+ \\ \bar{\rho}_+ &= \rho_- \\ 1 + \rho_+ + \rho_- &= 0 \end{aligned}$$

Wir wissen nun wie viele Lösungen komplexe polynomiale Gleichungen n -ten Grades besitzen, nicht jedoch, wie diese für eine gegebene Gleichung berechnet werden können. Betrachten wir also allgemein eine Gleichung der Form

$$a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = 0$$

mit gegebenen $a_j \in \mathbb{C}$, $a_n \neq 0$, und gesuchtem z .

Lineare Polynome ($n = 1$): Hier gilt

$$\begin{aligned} a_1 z + a_0 &= 0 \\ \Leftrightarrow z &= -\frac{a_0}{a_1}. \end{aligned}$$

Quadratische Polynome ($n = 2$): Wir bestimmen die Lösungen der Gleichung

$$a_2 z^2 + a_1 z + a_0 = 0. \quad (2.17)$$

Da $a_2 \neq 0$ ist, können wir $p := \frac{a_1}{2a_2}$ und $q := \frac{a_0}{a_2}$ setzen. Gleichung (2.17) ist äquivalent zu

$$z^2 + \frac{a_1}{a_2} z + \frac{a_0}{a_2} = 0,$$

2 Matrixrechnung

und damit zu

$$z^2 + 2pz + q = 0. \quad (2.18)$$

Hierbei heißt $D := p^2 - q \in \mathbb{C}$ *Diskriminante* von (2.18). Sei nun $\sqrt{D} \in \mathbb{C}$ eine Wurzel von D . Gemeint ist eine Lösung w der Gleichung

$$w^2 = D.$$

Aus dem Fundamentalsatz wissen wir, dass es in diesem Fall genau zwei Wurzeln gibt. Ist also \sqrt{D} die eine Wurzel, so ist $-\sqrt{D}$ die andere. Der Bezeichner \sqrt{D} legt also, anders als bei reellen Polynomen, den Wurzelwert nicht eindeutig fest. Man sagt, \sqrt{D} ist nicht wohldefiniert. Dies macht aber im Folgenden nichts.

Satz 2.5.17 (*p-q-Formel, Satz von Vieta für quadratische Gleichungen*)

Die Lösungen von (2.18) sind gegeben durch

$$z_1 = -p + \sqrt{D} \quad \text{und} \quad z_2 = -p - \sqrt{D}.$$

Ferner gilt

$$z_1 + z_2 = -2p, \quad z_1 \cdot z_2 = q \quad \text{sowie} \quad (z_1 - z_2)^2 = 4D. \quad (2.19)$$

Beweis.

Wir überprüfen zunächst, dass (2.19) für unsere Wahl von z_1 und z_2 gilt. Es ist

$$\begin{aligned} z_1 + z_2 &= -p + \sqrt{D} + (-p - \sqrt{D}) \\ &= -2p \end{aligned}$$

und

$$\begin{aligned} z_1 \cdot z_2 &= (-p + \sqrt{D})(-p - \sqrt{D}) \\ &= p^2 - D \\ &= q \end{aligned}$$

sowie

$$\begin{aligned} (z_1 - z_2)^2 &= (-p + \sqrt{D} - (-p - \sqrt{D}))^2 \\ &= (2\sqrt{D})^2 \\ &= 4D. \end{aligned}$$

Nun sehen wir, dass z_1 und z_2 tatsächlich die beiden Nullstellen des Polynoms sind, denn

$$\begin{aligned}(z - z_1)(z - z_2) &= z^2 - z_1z - z_2z + z_1z_2 \\ &= z^2 - (z_1 + z_2)z + z_1z_2 \\ &\stackrel{(2.19)}{=} z^2 + 2pz + q.\end{aligned}$$

Für die letzte Gleichheit haben wir die ersten beiden Formeln aus (2.19) benutzt. \square

Beispiel 2.5.18. Für $z^2 - 2z - 3 = 0$ haben wir $p = -1$, $q = -3$ und $D = p^2 - q = 1 - (-3) = 4$. Damit folgt nach dem Satz

$$\begin{aligned}z_1 &= -p + \sqrt{D} = 1 + \sqrt{4} = 3 \quad \text{und} \\ z_2 &= -p - \sqrt{D} = 1 - \sqrt{4} = -1.\end{aligned}$$

Beispiel 2.5.19. Für $z^2 - (1+i)z + i = 0$ ist $p = -\frac{1+i}{2}$, $q = i$ und $D = \left(-\frac{1+i}{2}\right)^2 - i = \frac{1+2i-1}{4} - i = \frac{1}{2}i - i = -\frac{1}{2}i$. Damit ist $\sqrt{D} = \frac{\sqrt{-i}}{\sqrt{2}}$. Wir bestimmen $\sqrt{-i}$ durch Lösen der Gleichung $w \cdot w = -i$. Hierbei ist unser geometrisches Verständnis von der Multiplikation komplexer Zahlen überaus hilfreich. Wegen $|-i| = 1$ und $\arg(-i) = \frac{3\pi}{2}$ ist

$$-i = e^{i\frac{3\pi}{2}} = e^{i(\frac{3}{4}\pi + \frac{3}{4}\pi)} = e^{i\frac{3}{4}\pi} \cdot e^{i\frac{3}{4}\pi} =: w \cdot w.$$

Wählen wir als eine Lösung also

$$\sqrt{-i} = e^{i\frac{3\pi}{4}} = \cos\left(\frac{3\pi}{4}\right) + i \sin\left(\frac{3\pi}{4}\right) = -\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2} = \frac{\sqrt{2}}{2}(-1 + i),$$

so erhalten wir schließlich

$$\sqrt{D} = \frac{\sqrt{-i}}{\sqrt{2}} = \frac{\frac{\sqrt{2}}{2}(-1 + i)}{\sqrt{2}} = \frac{1}{2}(-1 + i).$$

Folglich sind die beiden Nullstellen gegeben durch

$$\begin{aligned}z_1 &= -p + \sqrt{D} = \frac{1+i}{2} + \frac{1}{2}(1+i) = i \quad \text{und} \\ z_2 &= -p - \sqrt{D} = \frac{1+i}{2} - \frac{1}{2}(1+i) = 1.\end{aligned}$$

Bemerkung 2.5.20. Für $p, q \in \mathbb{R}$ ist auch $D \in \mathbb{R}$ und wir erhalten folgende Fallunterscheidung:

reelle Diskriminante	Anzahl der Lösungen	
$D > 0$	zwei verschiedene reelle Lösungen	$z_1 \neq z_2 \in \mathbb{R}$
$D = 0$	eine doppelte reelle Lösung	$z_1 = z_2 \in \mathbb{R}$
$D < 0$	zwei verschiedene, zueinander komplex konjugierte Lösungen	$\bar{z}_1 = z_2 \in \mathbb{C} \setminus \mathbb{R}$

2 Matrixrechnung

Kubische Polynome ($n = 3$): Da $a_3 \neq 0$ ist, kann für die kubische Gleichung

$$a_3 z^3 + a_2 z^2 + a_1 z + a_0 = 0. \quad (2.20)$$

wie folgt substituiert werden:

$$\begin{aligned} w &= z + \frac{a_2}{3a_3} \\ \Leftrightarrow z &= w - \frac{a_2}{3a_3}. \end{aligned}$$

Damit ist (2.20) äquivalent zu

$$\begin{aligned} 0 &= a_3 \left(w - \frac{a_2}{3a_3} \right)^3 + a_2 \left(w - \frac{a_2}{3a_3} \right)^2 + a_1 \left(w - \frac{a_2}{3a_3} \right) + a_0 \\ &= a_3 \left(w^3 - 3 \cdot w^2 \cdot \frac{a_2}{3a_3} + 3 \cdot w \cdot \left(\frac{a_2}{3a_3} \right)^2 - \left(\frac{a_2}{3a_3} \right)^3 \right) \\ &\quad + a_2 \left(w^2 - 2 \cdot w \cdot \frac{a_2}{3a_3} + \left(\frac{a_2}{3a_3} \right)^2 \right) + a_1 \left(w - \frac{a_2}{3a_3} \right) + a_0 \\ &= a_3 \left(w^3 - \frac{a_2}{a_3} w^2 + \frac{a_2^2}{3a_3^2} w - \frac{a_2^3}{3a_3^3} \right) \\ &\quad + a_2 \left(w^2 - \frac{2a_2}{3a_3} w + \frac{a_2^2}{3a_3^2} \right) + a_1 \left(w - \frac{a_2}{3a_3} \right) + a_0 \\ &= \left(a_3 w^3 - a_2 w^2 + \frac{a_2^2}{3a_3} w - \frac{a_2^3}{27a_3^3} \right) \\ &\quad + \left(a_2 w^2 - \frac{2a_2^2}{3a_3} w + \frac{a_2^3}{9a_3^2} \right) + \left(a_1 w - \frac{a_1 a_2}{3a_3} \right) + a_0 \\ &= a_3 w^3 + \left(\frac{a_2^2}{3a_3} - \frac{2a_2^2}{3a_3} + a_1 \right) w - \frac{a_2^3}{27a_3^3} + \frac{a_2^3}{9a_3^2} - \frac{a_1 a_2}{3a_3} + a_0 \\ &= a_3 w^3 + \left(-\frac{a_2^2}{3a_3} + a_1 \right) w + \left(\frac{2a_2^3}{27a_3^3} - \frac{a_1 a_2}{3a_3} + a_0 \right). \end{aligned}$$

Division durch a_3 liefert

$$w^3 + \underbrace{\left(-\frac{a_2^2}{3a_3^2} + \frac{a_1}{a_3} \right)}_{:=3p} w + \underbrace{\left(\frac{2a_2^3}{27a_3^3} - \frac{a_1 a_2}{3a_3^2} + \frac{a_0}{a_3} \right)}_{:=2q} = 0$$

und damit

$$w^3 + 3pw + 2q = 0. \quad (2.21)$$

Die Substitutionen haben bewirkt, dass der quadratische Term aus der Gleichung verschwunden ist. Im kubischen Fall ist die Diskriminante von (2.21) nun definiert durch $D := p^3 + q^2$. Die Formeln für die allgemeinen Lösungen kubischer Polynomgleichungen heißen Cardanische Gleichungen. Wir fassen sie in folgendem

Satz 2.5.21 (Cardano, Vieta)

Seien die dritten Wurzeln $u_{\pm} := \sqrt[3]{-q \pm \sqrt{D}}$ so gewählt, dass $u_+u_- = -p$.⁵ Dann sind

$$\begin{aligned}w_1 &= u_+ + u_-, \\w_2 &= \rho_+u_+ + \rho_-u_-, \\w_3 &= \rho_-u_+ + \rho_+u_-\end{aligned}$$

die Lösungen von 2.21. Dabei sind ρ_+ und ρ_- die dritten Einheitswurzeln wie in Beispiel 2.5.16. Ferner gilt

$$w_1 + w_2 + w_3 = 0, \quad (2.22)$$

$$w_1w_2 + w_2w_3 + w_1w_3 = 3p, \quad (2.23)$$

$$w_1w_2w_3 = -2q, \quad (2.24)$$

$$(w_1 - w_2)^2(w_2 - w_3)^2(w_1 - w_3)^2 = -108D. \quad (2.25)$$

Beweis.

Unter Benutzung einer Formel aus Beispiel 2.5.16 rechnen wir (2.22) nach:

$$\begin{aligned}w_1 + w_2 + w_3 &= u_+ + u_- + \rho_+u_+ + \rho_-u_- + \rho_-u_+ + \rho_+u_- \\&= (1 + \rho_+ + \rho_-)u_+ + (1 + \rho_+ + \rho_-)u_- \\&= 0.\end{aligned}$$

Ähnlich folgt (2.23):

$$\begin{aligned}w_1w_2 + w_2w_3 + w_1w_3 &= w_1(w_1 + w_3) + w_2w_3 \\&= \underbrace{(\rho_+ + \rho_-)}_{=-1}(u_+ + u_-)^2 + u_+^2 + \underbrace{(\rho_+^2 + \rho_-^2)}_{=\rho_- + \rho_+ = -1}u_+u_- + u_-^2 \\&= -(u_+^2 + 2u_+u_- + u_-^2) + u_+^2 + -u_+u_- + u_-^2 \\&= -3u_+u_- \\&= 3p,\end{aligned}$$

sowie (2.24):

$$\begin{aligned}w_1w_2w_3 &= (u_+ + u_-)(\rho_+u_+ + \rho_-u_-)(\rho_-u_+ + \rho_+u_-) \\&= \rho_+\rho_-u_+^3 + (\rho_+^2 + \rho_-^2 + \rho_+\rho_-)u_+^2u_- + (\rho_+^2 + \rho_-^2 + \rho_+\rho_-)u_+u_-^2 + \rho_-\rho_+u_-^3 \\&= u_+^3 + u_-^3 \\&= -q + \sqrt{D} + (-q - \sqrt{D}) \\&= -2q.\end{aligned}$$

2 Matrixrechnung

Damit können wir nachprüfen, dass w_1, w_2 und w_3 tatsächlich die drei Nullstellen des Polynoms sind:

$$\begin{aligned}(w - w_1)(w - w_2)(w - w_3) &= w^3 - (w_1 + w_2 + w_3)w^2 \\ &\quad + (w_1w_2 + w_1w_3 + w_2w_3)w - w_1w_2w_3 \\ &= w^3 + 3pw + 2q.\end{aligned}$$

Für die letzte Gleichung haben wir die Formeln (2.22), (2.23) und (2.24) eingesetzt. Die noch fehlende Formel (2.25) rechnet man auf ähnliche Weise nach. Die Einzelheiten überlassen wir dem interessierten Leser. \square

Beispiel 2.5.22. Wir suchen die Lösungen der kubische Gleichung mit reellen Koeffizienten:

$$z^3 - 6z^2 + 21z - 52 = 0 \tag{2.26}$$

Hierbei ist $a_0 = -52$, $a_1 = 21$, $a_2 = -6$ und $a_3 = 1$. Zunächst eliminieren wir den quadratischen Term mittels der Substitution

$$w = z + \frac{a_2}{3a_3} = z + 2.$$

Dann ist (2.26) äquivalent zu

$$w^3 + 3pw + 2q = 0, \tag{2.27}$$

wobei

$$\begin{aligned}3p &= -\frac{a_2^2}{3a_3^2} + \frac{a_1}{a_3} = 9 \\ \Leftrightarrow p &= 3\end{aligned}$$

und

$$\begin{aligned}2q &= \frac{2a_2^3}{27a_3^3} - \frac{a_1a_2}{3a_3^2} + \frac{a_0}{a_3} = -26 \\ \Leftrightarrow q &= -13\end{aligned}$$

ist. Für die Diskriminante erhalten wir also

$$D = p^3 + q^2 = 196.$$

Nun können die Cardanischen Lösungsformeln angewendet werden. Es ist

$$u_{\pm} = \sqrt[3]{-q \pm \sqrt{D}} = \sqrt[3]{13 \pm \sqrt{196}}$$

und damit

$$u_+ = \sqrt[3]{27} = 3 \quad \text{und} \quad u_- = \sqrt[3]{-1} = -1.$$

Man beachte, dass hierbei die Forderung

$$u_+ \cdot u_- = 3 \cdot (-1) = -3 = -p$$

erfüllt ist. Insbesondere bei kubischen Gleichungen mit komplexen Koeffizienten darf man nicht vergessen zu überprüfen, dass man die „richtigen“ dritten Wurzeln gewählt hat.

Mit den dritten Einheitswurzeln $\rho_+ = -\frac{1}{2}(1 - \sqrt{3}i)$ und $\rho_- = -\frac{1}{2}(1 + \sqrt{3}i)$ können wir hier direkt die Lösungen von (2.27) angeben:

$$\begin{aligned} w_1 &= u_+ + u_- = 3 - 1 = 2, \\ w_2 &= \rho_+ u_+ + \rho_- u_- = -\frac{3}{2}(1 - \sqrt{3}i) + \frac{1}{2}(1 + \sqrt{3}i) = -1 + 2\sqrt{3}i, \\ w_3 &= \rho_- u_+ + \rho_+ u_- = -\frac{3}{2}(1 + \sqrt{3}i) + \frac{1}{2}(1 - \sqrt{3}i) = -1 - 2\sqrt{3}i. \end{aligned}$$

Die Rücksubstitution

$$z_i = w_i - \frac{a_2}{3a_3} = w_i + 2$$

liefert schließlich die Lösungen von (2.26) zu

$$z_1 = 4 \quad \text{sowie} \quad z_2 = 1 + 2\sqrt{3}i \quad \text{und} \quad z_3 = 1 - 2\sqrt{3}i.$$

Hierbei sind z_2 und z_3 zueinander komplex konjugiert. Die Lösungsmenge ist symmetrisch, da z_3 aus z_2 durch Konjugation, d.h. durch Spiegelung an der reellen Achse hervorgeht. Polynomiale Gleichungen mit reellen Koeffizienten $a_i \in \mathbb{R}$ für alle $i \in \{0, \dots, n\}$, $n \geq 1$, besitzen stets symmetrische Lösungsmengen. Diese Aussage wird als Übungsaufgabe gezeigt.

Bemerkung 2.5.23. Für Polynome vom Grad 4 kann man ähnliche Lösungsformeln herleiten. Für Grad 5 und größer hatte man lange nach ähnlichen Formeln gesucht, bis der norwegische Mathematiker Niels Henrik Abel bewies, dass es solche geschlossenen Formeln dafür nicht geben kann. Abel starb 1829 mit 26 Jahren an Lungentuberkulose. Aus seiner Schulzeit soll es einen Klassenbucheintrag seines Lehrers Holmboe geben: „... dass er der größte Mathematiker der Welt werden kann, wenn er lange genug lebt“. Anlässlich des 200. Geburtstages von Abel richtete die norwegische Regierung im Jahr 2002 eine Stiftung zur Verleihung eines Preises für außergewöhnliche wissenschaftliche Arbeiten auf dem Gebiet der Mathematik ein. Dieser **Abelpreis** wird seit 2003 jährlich durch die Norwegische Akademie der Wissenschaften verliehen und ist mit 6 Millionen norwegischen Kronen (derzeit etwa 625.000 Euro) dotiert. Er entspricht dem Nobelpreis in anderen wissenschaftlichen Disziplinen.



Abb. 42

2.6 Graphentheorie - etwas Kombinatorik

Wir betrachten folgende Landkarte Deutschlands mit seinen durchnummerierten Bundesländern.

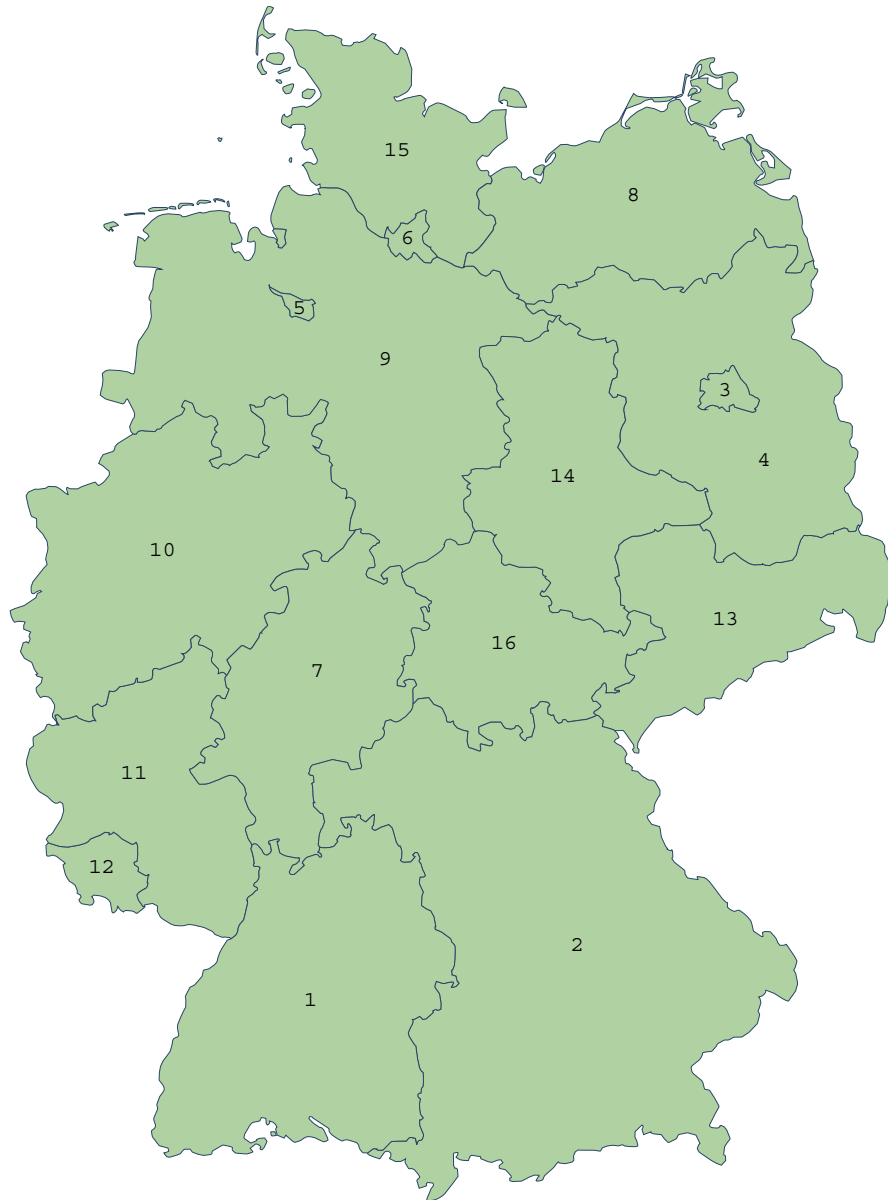


Abb. 43

Es soll uns nachfolgend um Fragestellungen folgenden Typs gehen:

Wie viele Möglichkeiten gibt es, von Brandenburg nach Rheinland-Pfalz zu gelangen und dabei genau 10-mal eine Landesgrenze zu überschreiten? (2.28)

Folgender systematischer Zugang bedient sich der uns inzwischen vertrauten Matrixrechnung. Wir codieren zunächst die wesentlichen Informationen in einem so genannten **Graph**. Anschaulich besteht ein Graph aus zwei Mengen, einer Menge von Knoten (Ecken) und einer Menge von Kanten. Eine Kante ist dabei die Verbindung zweier Knoten.

Im Bundesländer-Beispiel haben wir für den Bundesländer-Graph die Menge $\{\text{Knoten}\} = \{\text{Bundesländer}\}$ und folgende Kantenrelation: Zwei Länder werden genau dann durch eine Kante verbunden, wenn sie aneinander grenzen. Dies liefert folgende Darstellung des Graphen:

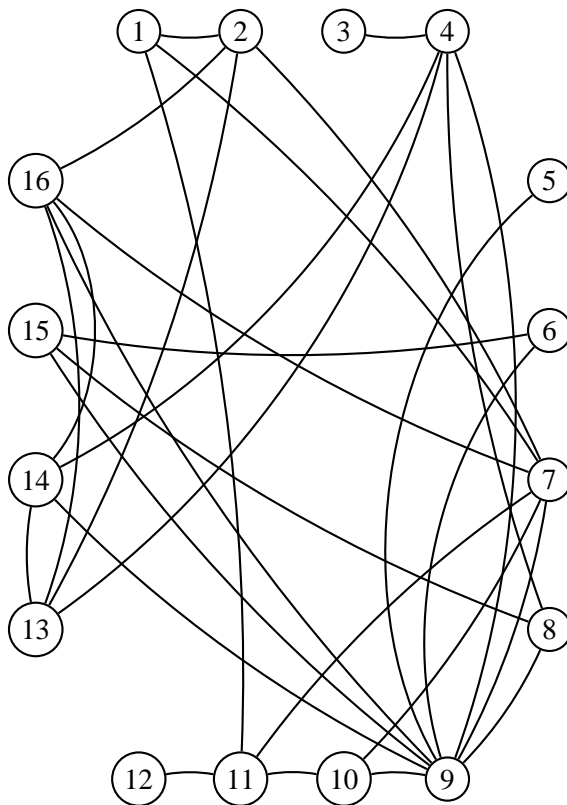


Abb. 44

Offenbar kann ein Graph mit vielen Informationen sehr unübersichtlich werden.

2 Matrixrechnung

Mathematisch beschreiben wir daher einen Graph mit endlich vielen Knoten durch eine so genannte **Adjazenzmatrix** A . Hierzu nummerieren wir die Knoten durch und setzen

$$A_{ij} := \begin{cases} 1 & , \text{ falls } i\text{-ter und } j\text{-ter Knoten durch eine Kante verbunden sind} \\ 0 & , \text{ sonst} \end{cases}$$

Im Bundesländer-Beispiel erhalten wir folgende 16×16 -Matrix:

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

Stets ist A symmetrisch, d.h. für alle i, j gilt $A_{ij} = A_{ji}$.

Definition 2.6.1. Ein **Weg der Länge** m in einem Graphen ist ein $(m+1)$ -Tupel von Knoten (k_0, \dots, k_m) , so dass für alle $i \in \{0, \dots, m-1\}$ gilt:

k_i und k_{i+1} sind durch eine Kante verbunden.

Im Bundesländer-Beispiel sind z.B. $(1, 2, 7)$ und $(5, 9, 5, 9, 6)$ Wege der Länge 2 bzw. 4, $(1, 2, 3)$ hingegen ist kein Weg, da 2 und 3 nicht miteinander verbunden sind (Bayern und Berlin haben keine gemeinsame Landesgrenze).

Die Lösung der Fragestellung (2.28) ist somit durch die Anzahl der Wege der Länge $m = 10$ von Knoten 4 nach Knoten 11 bestimmt. Allgemein gilt folgender

Satz 2.6.2

Ist A die Adjazenzmatrix eines endlichen Graphen (mit nummerierten Knoten), so ist die Anzahl der Wege der Länge m vom i -ten zum j -ten Knoten genau der (i, j) -te Eintrag der Matrix

$$A^m := \underbrace{A \cdot \dots \cdot A}_{m\text{-mal}}$$

Beweis.

Sei $\alpha(i, j, m)$ die Anzahl der Wege der Länge m vom i -ten zum j -ten Knoten. Wir zeigen:

$$(A^m)_{ij} = \alpha(i, j, m) \quad (2.29)$$

mittels vollständiger Induktion nach der Weglänge m .

Im Induktionsanfang erhalten wir für $m = 1$ per Definition

$$A_{ij} = (A^1)_{ij} = \alpha(i, j, 1) = \begin{cases} 1 & , \text{ die Knoten } i \text{ und } j \text{ sind verbunden} \\ 0 & , \text{ die Knoten } i \text{ und } j \text{ sind nicht verbunden} \end{cases} .$$

Sei für den Induktionsschritt die Aussage (2.29) für ein $m \in \mathbb{N}$ gültig. Wir zeigen, dass dann auch

$$(A^{m+1})_{ij} = \alpha(i, j, m + 1)$$

gilt. Wir fixieren man den Startknoten i und den Zielknoten j . Um in $m + 1$ Schritten nach von i nach j zu gelangen, müssen wir zunächst in m Schritten von i zu irgendeinem Knoten k gelangen und dann in einem Schritt von k nach j . Letzteres ist nur dann möglich, wenn k und j verbunden sind, d.h. wenn $A_{kj} = 1$. Es gilt also

$$\alpha(i, j, m + 1) = \sum_k \alpha(i, k, m) \cdot A_{kj}.$$

Mittels der Induktionsvoraussetzung (IV) folgern wir nun:

$$\begin{aligned} \alpha(i, j, m + 1) &= \sum_k \alpha(i, k, m) \cdot A_{kj} \\ &\stackrel{\text{IV}}{=} \sum_k (A^m)_{ik} \cdot A_{kj} \\ &= (A^m \cdot A)_{ij} \\ &= (A^{m+1})_{ij} \end{aligned}$$

□

2 Matrixrechnung

Im Bundesländer-Beispiel multipliziert man die Matrix A mit sich selbst und erhält

$$A^2 = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 & 0 & 2 & 0 & 1 & 2 & 1 & 1 & 1 & 0 & 0 & 2 \\ 1 & 4 & 0 & 1 & 0 & 0 & 2 & 0 & 2 & 1 & 2 & 0 & 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 5 & 1 & 1 & 1 & 1 & 2 & 1 & 0 & 0 & 1 & 2 & 2 & 3 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 & 1 & 2 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 2 & 2 & 0 & 1 & 1 & 1 & 6 & 1 & 2 & 2 & 2 & 1 & 2 & 2 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 & 2 & 1 & 3 & 2 & 1 & 0 & 0 & 1 & 2 & 1 & 1 \\ 1 & 2 & 1 & 2 & 0 & 1 & 2 & 2 & 9 & 1 & 2 & 0 & 3 & 2 & 2 & 2 \\ 2 & 1 & 0 & 1 & 1 & 1 & 2 & 1 & 1 & 3 & 1 & 1 & 0 & 1 & 1 & 2 \\ 1 & 2 & 0 & 0 & 0 & 0 & 2 & 0 & 2 & 1 & 4 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 2 & 1 & 3 & 0 & 0 & 0 & 4 & 2 & 0 & 2 \\ 0 & 2 & 1 & 2 & 1 & 1 & 2 & 2 & 2 & 1 & 0 & 0 & 2 & 4 & 1 & 2 \\ 0 & 0 & 0 & 2 & 1 & 1 & 1 & 1 & 2 & 1 & 0 & 0 & 0 & 1 & 3 & 1 \\ 2 & 2 & 0 & 3 & 1 & 1 & 2 & 1 & 2 & 2 & 1 & 0 & 2 & 2 & 1 & 5 \end{pmatrix}$$

So ist zum Beispiel der 1-1-Eintrag gleich 3, was bedeutet, dass man auf genau 3 Weisen in zwei Schritten von Land 1 zu sich selbst reisen kann. In der Tat hat Baden-Württemberg genau 3 Nachbarländer. Der 3-14-Eintrag ist gleich 1, denn man kann auf genau eine Weise in zwei Schritten von Berlin nach Sachsen-Anhalt reisen (nämlich über Brandenburg). Schließlich ist der 2-5-Eintrag gleich 0, da man nicht in zwei Schritten von Bayern nach Bremen kommt.

Um die Ausgangsfrage zu beantworten, müssen wir die 10-te Potenz von A berechnen. Wenn man kühlen Kopf bewahrt und dies tut, so erhält man

$$A^{10} = \begin{pmatrix} 122133 & 175608 & 41206 & 191539 & 69876 & 95617 & 251161 & 136823 & \dots \\ 175608 & 256948 & 61837 & 283567 & 102476 & 141044 & 365281 & 202881 & \dots \\ 41206 & 61837 & 16181 & 70772 & 25306 & 35425 & 87950 & 51606 & \dots \\ 191539 & 283567 & 70772 & 327963 & 118896 & 163586 & 405312 & 234358 & \dots \\ 69876 & 102476 & 25306 & 118896 & 43734 & 59813 & 147740 & 85119 & \dots \\ 95617 & 141044 & 35425 & 163586 & 59813 & 82359 & 202852 & 117784 & \dots \\ 251161 & 365281 & 87950 & 405312 & 147740 & 202852 & 522812 & 290802 & \dots \\ 136823 & 202881 & 51606 & 234358 & 85119 & 117784 & 290802 & 169390 & \dots \\ 319249 & 473106 & 118896 & 532707 & 190956 & 265744 & 672308 & 384640 & \dots \\ 152981 & 221828 & 53153 & 248598 & 91066 & 124615 & 318354 & 177768 & \dots \\ 122267 & 177363 & 41752 & \boxed{191066} & 69033 & 95186 & 251872 & 136938 & \dots \\ 27621 & 38805 & 8757 & 41752 & 15557 & 21055 & 56168 & 29812 & \dots \\ 178571 & 263718 & 65943 & 296414 & 106589 & 147676 & 375339 & 213619 & \dots \\ 204096 & 301686 & 75337 & 344079 & 124737 & 172005 & 431041 & 247342 & \dots \\ 119833 & 177465 & 44690 & 205927 & 74788 & 103242 & 254324 & 147932 & \dots \\ 247727 & 362837 & 88468 & 409963 & 148901 & 204426 & 518378 & 292894 & \dots \end{pmatrix}$$

Also ergibt sich die Anzahl der Wege der Länge 10 von Brandenburg (4) nach Rheinland-Pfalz (11) zu

$$(A^{10})_{4,11} = (A^{10})_{11,4} = 191066.$$

Wir beschließen unseren Ausflug in die Graphentheorie mit folgenden zwei bekannten graphentheoretischen Problemen.

Das Königsberger Brückenproblem (1736 gelöst von Euler).

Rechts sehen wir die Brücken der Stadt Königsberg (heute Kaliningrad). Die Frage ist nun, ob es einen Rundweg gibt, bei dem man alle sieben Brücken der Stadt über den Pregel genau einmal überquert und wieder zum Ausgangspunkt gelangt. Kurz gefragt: Gibt es einen Rundweg durch Königsberg, bei dem man über jede Brücke genau einmal geht?

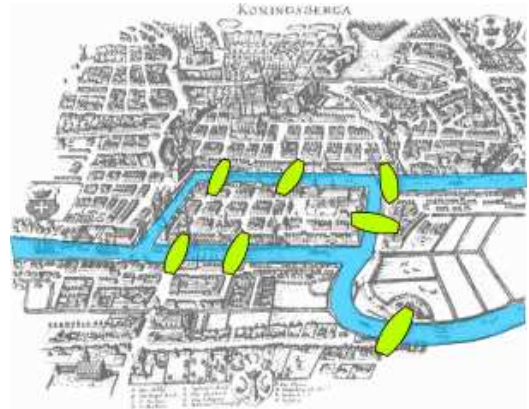


Abb. 45

Für die mathematische Handhabung ordnen wir dem Stadtplan einen Graphen zu. Dabei entspricht jeder Knoten einem zusammenhängenden Stadtgebiet und die Brücken, wobei werden durch Kanten realisiert. Wir erlauben jetzt also, dass zwei Knoten durch mehr als eine Kante verbunden werden können. Somit ist das Problem auf folgende Frage reduziert: Finden wir einen Weg der Länge 7 durch den Graphen, bei dem jede Kante genau einmal benutzt wird.

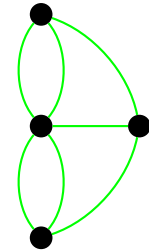


Abb. 46

Eulers Argument geht nun wie folgt: Bei jedem Besuch eines Knotens werden genau 2 Kanten „verbraucht“, nämlich zum einen die Kante, über die wir zum Knoten gelangen und die, über die wir den Knoten wieder verlassen. Dies bedeutet aber, dass falls ein Rundweg existiert, so muss in jedem Knoten eine gerade Anzahl von Kanten vorliegen. Im Königsberger Graphen liegen jedoch jeweils ungerade viele Kanten in jedem Knoten an. Folglich gibt es keinen Rundweg durch Königsberg.

Das Vierfarbenproblem.

Wie viele verschiedene Farben benötigt man, um jede Landkarte so zu färben, dass zwei benachbarte Länder stets verschieden gefärbt sind?

Drei Farben sind im Allgemeinen zu wenig, d.h. es werden mindestens vier Farben benötigt:

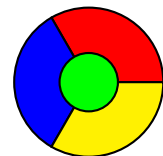


Abb. 47

Auch in folgendem Beispiel genügen vier Farben:

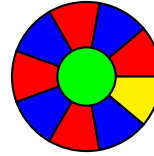


Abb. 48

Der Engländer Francis Guthrie vermutete 1852 erstmals, dass vier Farben für jede mögliche Konstellation ausreichend sind. Dies ist genau die Aussage des Vier-Farben-Satzes, welcher tatsächlich korrekt ist, wie Kenneth Appel und Wolfgang Haken 1977 im ersten Computer-gestützten Beweis überhaupt herausfanden. Man kann sich denken, dass dieses nichtklassische Beweismittel anfangs sehr umstritten war.

2.7 Determinanten

Wir bezeichnen mit $\text{Mat}(n \times m, \mathbb{C})$ die Menge der $n \times m$ -Matrizen mit komplexen Einträgen.

Definition 2.7.1. Sei $n \in \mathbb{N}$. Eine **Determinanten-Funktion** ist eine Abbildung

$$\det : \text{Mat}(n \times n, \mathbb{C}) \rightarrow \mathbb{C},$$

so dass gilt:

(D1) \det ist linear in jeder Zeile, d.h. es gilt für alle λ und jedes $i \in \{1, \dots, n\}$:

$$\det \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ \lambda \cdot A_{i1} & \cdots & \lambda \cdot A_{in} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} = \lambda \cdot \det \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{i1} & \cdots & A_{in} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}$$

und

$$\det \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{i1} + A'_{i1} & \cdots & A_{in} + A'_{in} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} = \det \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{i1} & \cdots & A_{in} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} + \det \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A'_{i1} & \cdots & A'_{in} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}.$$

(D2) \det ist alternierend, d.h. stimmen zwei (verschiedene) Zeilen von A überein, so ist $\det(A) = 0$.

2 Matrixrechnung

(D3) \det ist normiert, d.h.

$$\det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = 1.$$

Satz 2.7.2

Für jedes $n \in \mathbb{N}$ gibt es genau eine Determinanten-Funktion.

Wir führen den Nachweis der Existenz und Eindeutigkeit der Determinanten-Funktion an dieser Stelle nicht vor. Stattdessen betrachten wir zunächst folgenden Spezialfall:

Beispiel 2.7.3. Für $n = 1$ gilt

$$\begin{aligned} \det((a)) &= \det((a \cdot 1)) \\ &\stackrel{(D1)}{=} a \cdot \det((1)) \\ &\stackrel{(D3)}{=} a \cdot 1 \\ &= a. \end{aligned}$$

Nur aus den Bedingungen (D1) bis (D3) und ohne Kenntnis konkreter Rechenvorschriften lassen sich folgende Eigenschaften der Determinanten-Funktion herleiten:

Satz 2.7.4

Für alle $A, B \in \text{Mat}(n \times n, \mathbb{C})$ und $\lambda \in \mathbb{C}$ gilt:

(D4) $\det(\lambda \cdot A) = \lambda^n \cdot \det(A)$.

(D5) Ist eine Zeile von A gleich 0, so ist $\det(A) = 0$.

(D6) Entsteht B aus A durch Vertauschung zweier Zeilen, so gilt $\det(B) = -\det(A)$.

(D7) Entsteht B durch Addition des Vielfachen einer Zeile zu einer anderen, so gilt $\det(B) = \det(A)$.

(D8) Ist A eine obere Dreiecksmatrix mit den Diagonaleinträgen $\lambda_1, \dots, \lambda_n$, d.h.

$$A = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix},$$

so gilt $\det(A) = \lambda_1 \cdot \dots \cdot \lambda_n$.

(D9) $\det(A) = 0 \Leftrightarrow \operatorname{rg}(A) < n$.

Beweis.

Zu (D4): Es gilt

$$\begin{aligned} \det(\lambda \cdot A) &= \det \begin{pmatrix} \lambda \cdot A_{11} & \cdots & \lambda \cdot A_{1n} \\ \vdots & & \vdots \\ \lambda \cdot A_{n1} & \cdots & \lambda \cdot A_{nn} \end{pmatrix} \\ &\stackrel{(D1)}{=} \lambda \cdot \det \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \lambda \cdot A_{21} & \cdots & \lambda \cdot A_{2n} \\ \vdots & & \vdots \\ \lambda \cdot A_{n1} & \cdots & \lambda \cdot A_{nn} \end{pmatrix} \\ &\stackrel{(D1)}{=} \dots \\ &\stackrel{(D1)}{=} \lambda^n \cdot \det(A). \end{aligned}$$

Zu (D5): Es gilt

$$\begin{aligned} \det \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} &= \det \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ 0 \cdot 0 & \cdots & 0 \cdot 0 \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} \\ &\stackrel{(D1)}{=} 0 \cdot \det \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix} \\ &= 0. \end{aligned}$$

2 Matrixrechnung

Zu (D6): Seien a_1, \dots, a_n die Zeilen von A , d.h. es sei

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} \quad \text{sowie} \quad B = \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix}.$$

Dann gilt

$$\begin{aligned} 0 &\stackrel{(D2)}{=} \det \begin{pmatrix} a_1 \\ \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \\ a_n \end{pmatrix} &\stackrel{(D1)}{=} \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_i + a_j \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i + a_j \\ \vdots \\ a_n \end{pmatrix} \\ &\stackrel{(D1)}{=} \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} \\ &\stackrel{(D2)}{=} \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} \\ &= \det(A) + \det(B). \end{aligned}$$

Somit ist $\det(B) = -\det(A)$.

Zu (D7): Seien wieder a_1, \dots, a_n die Zeilen von A . Für $i \neq j$ betrachten wir

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda \cdot a_j \\ \vdots \\ a_n \end{pmatrix}.$$

Dann gilt

$$\begin{aligned} \det(B) &\stackrel{(D1)}{=} \det \begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det \begin{pmatrix} a_1 \\ \vdots \\ \lambda \cdot a_j \\ \vdots \\ a_n \end{pmatrix} \\ &\stackrel{(D1)}{=} \det(A) + \lambda \cdot \det \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} \\ &\stackrel{(D2)}{=} \det(A). \end{aligned}$$

Zu (D8).

1. Fall:

Ist $\lambda_i \neq 0$ für alle $i \in \{1, \dots, n\}$, so addieren wir im ersten Schritt von der vorletzten Zeile ein gewisses Vielfaches der letzten Zeile, so dass der letzte Eintrag der $(n-1)$ -ten Zeile zu 0 wird. Durch sukzessives Addieren von Vielfachen gewisser Zeilen zu anderen Zeilen (beginnend von unten) kann A dann in die Matrix

$$B = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

überführt werden. Damit gilt

$$\begin{aligned} \det(A) &\stackrel{(D7)}{=} \det(B) \\ &= \det \begin{pmatrix} \lambda_1 \cdot 1 & 0 & \cdots & 0 \\ 0 & \lambda_2 \cdot 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \cdot 1 \end{pmatrix} \end{aligned}$$

2 Matrixrechnung

$$\stackrel{(D1)}{=} \lambda_1 \cdot \dots \cdot \lambda_n \cdot \det \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

$$\stackrel{(D3)}{=} \lambda_1 \cdot \dots \cdot \lambda_n.$$

2. Fall:

Gibt es ein $i \in \{1, \dots, n\}$, so dass $\lambda_i = 0$ gilt, dann sei o.B.d.A. i die größte Zahl aus $\{1, \dots, n\}$, für die $\lambda_i = 0$ ist. Wie im Fall 1 kann nun A durch Addition des Vielfachen gewisser Zeilen zu anderen Zeilen in die Form

$$B = \begin{pmatrix} \lambda_1 & * & \dots & \dots & \dots & \dots & * \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & \lambda_{i-1} & * & \dots & \dots & * \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & \lambda_{i+1} & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & \lambda_n \end{pmatrix}$$

überführt werden. Insbesondere gilt dann

$$\det(A) \stackrel{(D7)}{=} \det(B) \stackrel{(D5)}{=} 0.$$

Andererseits ist auch $\lambda_1 \cdot \dots \cdot \lambda_n = 0$, da wenigstens einer der Faktoren gleich 0 ist. Also gilt (D8) auch in diesem Fall.

Zu (D9):

Wir nennen Zeilenvertauschungen und Additionen des Vielfachen einer Zeile zu einer anderen Zeile **spezielle Zeilenumformungen**. Es entstehe A' aus A durch Anwendung spezieller Zeilenumformungen. Dann folgt aus (D6) und (D7)

$$\det(A') = \pm \det(A).$$

Wie im Gauß-Algorithmus kann A somit in eine obere Dreiecksmatrix

$$A' = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

überführt werden und wir erhalten nach (D8)

$$\det(A') = \lambda_1 \cdot \dots \cdot \lambda_n.$$

Dann gilt

$$\begin{aligned}
 \det(A) = 0 &\Leftrightarrow \det(A') = 0 \\
 &\Leftrightarrow \lambda_1 \cdot \dots \cdot \lambda_n = 0 \\
 &\Leftrightarrow \exists i \in \{1, \dots, n\} : \lambda_i = 0 \\
 &\Leftrightarrow \operatorname{rg}(A') < n \\
 &\Leftrightarrow \operatorname{rg}(A) < n.
 \end{aligned}$$

□

Nun wollen wir eine Auswahl der Eigenschaften (D1) bis (D9) gleich einmal anwenden und berechnen die Determinante der folgenden Matrix:

Beispiel 2.7.5. Wir betrachten

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

und berechnen

$$\begin{aligned}
 \det(A) &\stackrel{(D6)}{=} -\det \begin{pmatrix} 2 & 3 & 4 \\ 1 & i & 1 \\ 0 & 1 & i \end{pmatrix} \\
 &\stackrel{(D1)}{=} -\frac{1}{-2} \cdot \det \begin{pmatrix} 2 & 3 & 4 \\ -2 & -2i & -2 \\ 0 & 1 & i \end{pmatrix} \\
 &\stackrel{(D7)}{=} \frac{1}{2} \cdot \det \begin{pmatrix} 2 & 3 & 4 \\ 0 & 3-2i & 2 \\ 0 & 1 & i \end{pmatrix} \\
 &\stackrel{(D1)}{=} \frac{1}{2} \cdot \frac{1}{3-2i} \cdot \det \begin{pmatrix} 2 & 3 & 4 \\ 0 & 3-2i & 2 \\ 0 & 3-2i & 3i+2 \end{pmatrix} \\
 &\stackrel{(D7)}{=} \frac{1}{2} \cdot \frac{1}{3-2i} \cdot \det \begin{pmatrix} 2 & 3 & 4 \\ 0 & 3-2i & 2 \\ 0 & 0 & 3i \end{pmatrix} \\
 &\stackrel{(D8)}{=} \frac{1}{2(3-2i)} \cdot 2(3-2i)3i \\
 &= 3i.
 \end{aligned}$$

Im Allgemeinen ist es mühsam, die Determinante durch Anwendung der Eigenschaften der Determinanten-Funktion zu bestimmen. Wir wollen daher eine Möglichkeit kennenlernen, für eine gegebene Matrix eine Berechnungsformel für die zugehörige Determinante zu entwickeln, den Entwicklungssatz von Laplace. Dabei kommt folgende Technik zum Einsatz: Das Ausgangsproblem der Berechnung einer $n \times n$ -Matrix ($n > 2$) wird auf die Berechnung von

2 Matrixrechnung

$(n - 1) \times (n - 1)$ -Matrizen zurückgeführt, dann, falls erforderlich, auf die Berechnung von $(n - 2) \times (n - 2)$ -Matrizen, usw. Für 2×2 -Matrizen merken wir uns schließlich eine explizite Formel und brechen somit die Rückführung spätestens bei 2×2 -Matrizen ab.

Wie werden nun die Matrizen verkleinert?

Definition 2.7.6. Sei $A \in \text{Mat}(n \times n, \mathbb{C})$ und $i, j \in \{1, \dots, n\}$. Dann heißt

$$A_{ij}^{\text{Str}} := \begin{pmatrix} A_{11} & \cdots & A_{1j} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ A_{i1} & \cdots & A_{ij} & \cdots & A_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nj} & \cdots & A_{nn} \end{pmatrix} \begin{matrix} j\text{-te Spalte} \\ \\ i\text{-te Zeile} \\ \end{matrix}$$

Streichungsmatrix. Offenbar ist $A_{ij}^{\text{Str}} \in \text{Mat}((n - 1) \times (n - 1), \mathbb{C})$.

Stattdessen

Satz 2.7.7 (Spaltenentwicklungssatz von Laplace)

Sei $A \in \text{Mat}(n \times n, \mathbb{C})$, $n \geq 2$. Dann gilt

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} \cdot A_{ij} \cdot \det(A_{ij}^{\text{Str}}).$$

Dies nennt man Entwicklung der Determinante von A nach der j -ten Spalte.

Bemerkung 2.7.8. Prinzipiell besteht die Wahlfreiheit, nach welcher Spalte entwickelt wird. Es ist jedoch ratsam, Spalten zu bevorzugen, die möglichst viele Nullen enthalten, da dann von vornherein ganze Summanden entfallen und der Rechenaufwand sinkt.

Beispiel 2.7.9. Der Entwicklungssatz liefert im Speziellen eine explizite Berechnungsformel für Determinanten von 2×2 -Matrizen. Hierfür entwickeln wir

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{Mat}(2 \times 2, \mathbb{C})$$

einmal nach der 1. Spalte:

$$\begin{aligned}
 \det(A) &= \sum_{i=1}^2 (-1)^{i+1} \cdot A_{i1} \cdot \det(A_{i1}^{\text{Str}}) \\
 &= (-1)^{1+1} \cdot A_{11} \cdot \det(A_{11}^{\text{Str}}) + (-1)^{1+2} \cdot A_{21} \cdot \det(A_{21}^{\text{Str}}) \\
 &= a \cdot \det \begin{pmatrix} \cancel{a} & \cancel{b} \\ c & d \end{pmatrix} - c \cdot \det \begin{pmatrix} a & b \\ \cancel{c} & \cancel{d} \end{pmatrix} \\
 &= a \cdot \det(d) - c \cdot \det(b) \\
 &= ad - cb.
 \end{aligned}$$

Die Determinanten-Formel für 2×2 -Matrizen merken wir uns: Man bilde das Produkt der Hauptdiagonalelemente $a \cdot d$ und subtrahiere das Produkt der Nebendiagonalelemente $c \cdot b$.

Beispiel 2.7.10. Bei der Spaltenentwicklung der Determinante von

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

entscheiden wir uns bewusst für die erste Spalte, da im ersten Eintrag eine 0 vorkommt. Es gilt

$$\begin{aligned}
 \det(A) &= \sum_{i=1}^3 (-1)^{i+1} A_{i1} \det(A_{i1}^{\text{Str}}) \\
 &= (-1)^{1+1} A_{11} \det(A_{11}^{\text{Str}}) + (-1)^{2+1} A_{21} \det(A_{21}^{\text{Str}}) + (-1)^{3+1} A_{31} \det(A_{31}^{\text{Str}}) \\
 &= 0 \cdot \det \begin{pmatrix} \cancel{0} & \cancel{1} & \cancel{i} \\ 1 & i & 1 \\ 2 & 3 & 4 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 0 & 1 & i \\ \cancel{1} & \cancel{i} & \cancel{1} \\ 2 & 3 & 4 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 0 & 1 & i \\ 1 & i & 1 \\ \cancel{2} & \cancel{3} & \cancel{4} \end{pmatrix} \\
 &= 0 \cdot \det \begin{pmatrix} i & 1 \\ 3 & 4 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 1 & i \\ 3 & 4 \end{pmatrix} + 2 \cdot \det \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} \\
 &= 0 \cdot (i \cdot 4 - 3 \cdot 1) - 1 \cdot (1 \cdot 4 - 3 \cdot i) + 2 \cdot (1 \cdot 1 - i \cdot i) \\
 &= -(4 - 3i) + 2 \cdot 2 \\
 &= 3i.
 \end{aligned}$$

Bemerkung 2.7.11. Die bei der Laplace-Entwicklung auftretenden Vorzeichen $(-1)^{i+j}$ bilden ein Schachbrett-Muster. Etwa im Fall einer 5×5 -Matrix muss man sich somit die Vorzeichen

2 Matrixrechnung

nicht gesondert überlegen:

$$\begin{pmatrix} + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \\ - & + & - & + & - \\ + & - & + & - & + \end{pmatrix}$$

Determinanten-Eigenschaft (D1) besagt, dass \det linear in jeder Zeile ist. Außerdem gilt:

Folgerung 2.7.12

\det ist linear in jeder Spalte, d.h. für Spaltenvektoren $a_1, \dots, a_n, a_j, a'_j \in \mathbb{C}^n$ und $A = (a_1, \dots, a_n)$ gilt

$$\begin{aligned} \det(a_1, \dots, \lambda \cdot a_j, \dots, a_n) &= \lambda \cdot \det(a_1, \dots, a_j, \dots, a_n) \quad \text{und} \\ \det(a_1, \dots, a_j + a'_j, \dots, a_n) &= \det(a_1, \dots, a_j, \dots, a_n) + \det(a_1, \dots, a'_j, \dots, a_n). \end{aligned}$$

Beweis.

Die Laplace-Entwicklung nach der j -ten Spalte liefert

$$\begin{aligned} \det(a_1, \dots, \lambda \cdot a_j, \dots, a_n) &= \sum_{i=1}^n (-1)^{i+j} \cdot \lambda \cdot A_{ij} \cdot \det(A_{ij}^{\text{Str}}) \\ &= \lambda \cdot \sum_{i=1}^n (-1)^{i+j} \cdot A_{ij} \cdot \det(A_{ij}^{\text{Str}}) \\ &= \lambda \cdot \det(a_1, \dots, a_j, \dots, a_n). \end{aligned}$$

Analog erhalten wir für $\tilde{A} := (a_1, \dots, a_j + a'_j, \dots, a_n)$ mit $a_j = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{nj} \end{pmatrix}$, $a'_j = \begin{pmatrix} A'_{1j} \\ \vdots \\ A'_{nj} \end{pmatrix}$

$$\begin{aligned} \det(\tilde{A}) &= \sum_{i=1}^n (-1)^{i+j} \cdot \tilde{A}_{ij} \cdot \det(\tilde{A}_{ij}^{\text{Str}}) \\ &= \sum_{i=1}^n (-1)^{i+j} \cdot (A_{ij} + A'_{ij}) \cdot \det(A_{ij}^{\text{Str}}) \\ &= \sum_{i=1}^n (-1)^{i+j} \cdot A_{ij} \cdot \det(A_{ij}^{\text{Str}}) + \sum_{i=1}^n (-1)^{i+j} \cdot A'_{ij} \cdot \det(A_{ij}^{\text{Str}}) \\ &= \det(a_1, \dots, a_j, \dots, a_n) + \det(a_1, \dots, a'_j, \dots, a_n). \end{aligned}$$

□

Warum heißt der Entwicklungssatz für Determinanten eigentlich Spaltenentwicklungssatz? Gibt es auch einen Zeilenentwicklungssatz? Ja tatsächlich, den gibt es. Folgende Operation vertauscht die Zeilen und Spalten einer Matrix.

Definition 2.7.13. Zu $A \in \text{Mat}(m \times n, \mathbb{C})$ heißt die Matrix $A^t \in \text{Mat}(n \times m, \mathbb{C})$ mit den Einträgen

$$(A^t)_{ij} := A_{ji}$$

Transponierte von A .

Beispiel 2.7.14. Zu $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & i \end{pmatrix} \in \text{Mat}(2 \times 3, \mathbb{C})$ ist $A^t = \begin{pmatrix} 1 & 3 \\ 2 & 1 \\ 3 & i \end{pmatrix} \in \text{Mat}(3 \times 2, \mathbb{C})$.

Satz 2.7.15

Für alle $A \in \text{Mat}(n \times n, \mathbb{C})$ gilt $\det(A^t) = \det(A)$.

Beweis.

Die direkte Rechnung hierfür wird enorm groß. Der Nachweis schlägt daher einen auf den ersten Blick ungewöhnlichen, auf den zweiten Blick aber überaus eleganten Weg ein.

Zunächst definieren wir die „neue“ Determinanten-Funktion

$$\widetilde{\det} : \text{Mat}(n \times n, \mathbb{C}) \rightarrow \mathbb{C} \quad \text{durch die Vorschrift} \quad \widetilde{\det}(A) := \det(A^t).$$

Wir zeigen nun, dass $\widetilde{\det}$ genau die Eigenschaften (D1), (D2) und (D3) einer Determinanten-Funktion erfüllt. Da die Funktion \det nach Satz (2.7.2) eindeutig ist, folgt dann $\det = \widetilde{\det}$, also für jedes $A \in \text{Mat}(n \times n, \mathbb{C})$:

$$\det(A) = \det(A^t).$$

Prüfen wir also nach, dass $\widetilde{\det}$ die definierenden Eigenschaften (D1), (D2) und (D3) von \det erfüllt:

Zu (D1): Wir müssen zeigen, dass $\widetilde{\det}$ linear in jeder Zeile ist.

2 Matrixrechnung

Nach (2.7.12) ist \det linear in jeder Spalte. Die Abbildung $A \mapsto \widetilde{\det}(A) = \det(A^t)$ bildet nun aber gerade die Zeilen von A auf die Spalten von A^t ab. Somit ist $\widetilde{\det}$ linear in jeder Zeile.

Zu (D2): Wir zeigen, dass die Übereinstimmung zweier Zeilen in A dann $\widetilde{\det}(A) = 0$ nach sich zieht.

Sei also A eine Matrix mit zwei gleichen Zeilen. Dann besitzt A^t zwei gleiche Spalten und es gilt $\text{rg}(A^t) < n$. Nach (D9) ist dies äquivalent zu $\widetilde{\det}(A) = \det(A^t) = 0$.

Zu (D3): Es gilt

$$\widetilde{\det} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}^t = \det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \stackrel{(D3)}{=} 1.$$

Insgesamt folgt somit $\det(A) = \widetilde{\det}(A) = \det(A^t)$ für alle $A \in \text{Mat}(n \times n, \mathbb{C})$. \square

Definition 2.7.16. Die Matrix $\mathbb{1}_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ heißt $n \times n$ -**Einheitsmatrix**.

Satz 2.7.17 (Zeilenentwicklungssatz von Laplace)

Sei $A \in \text{Mat}(n \times n, \mathbb{C})$, $n \geq 2$. Dann gilt:

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} \cdot A_{ij} \cdot \det(A_{ij}^{\text{Str}})$$

(Entwicklung nach der i -ten Zeile).

Beweis.

Wie wir wissen, ist $\det(A) = \det(A^t)$. Der Spaltenentwicklungssatz für $\det(A^t)$ liefert somit die Behauptung. \square

Beispiel 2.7.18. Wir betrachten wieder $A = \begin{pmatrix} 0 & 1 & i \\ 1 & i & 1 \\ 2 & 3 & 4 \end{pmatrix}$ und entwickeln nach der 1. Zeile:

$$\begin{aligned}
 \det A &= \sum_{j=1}^3 (-1)^{1+j} A_{1j} \det(A_{1j}^{\text{Str}}) \\
 &= (-1)^{1+1} A_{11} \det(A_{11}^{\text{Str}}) + (-1)^{1+2} A_{12} \det(A_{12}^{\text{Str}}) + (-1)^{1+3} A_{13} \det(A_{13}^{\text{Str}}) \\
 &= 0 \cdot \det \begin{pmatrix} \cancel{0} & \cancel{1} & \cancel{i} \\ 1 & i & 1 \\ 2 & 3 & 4 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} \cancel{0} & \cancel{1} & \cancel{i} \\ 1 & i & 1 \\ 2 & \cancel{3} & \cancel{4} \end{pmatrix} + i \cdot \det \begin{pmatrix} \cancel{0} & \cancel{1} & \cancel{i} \\ 1 & i & 1 \\ 2 & 3 & \cancel{4} \end{pmatrix} \\
 &= 0 \cdot \det \begin{pmatrix} i & 1 \\ 3 & 4 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix} + i \cdot \det \begin{pmatrix} 1 & i \\ 2 & 3 \end{pmatrix} \\
 &= 0 - 1 \cdot (1 \cdot 4 - 2 \cdot 1) + i \cdot (1 \cdot 3 - 2 \cdot i) \\
 &= -2 + i(3 - 2i) \\
 &= 3i.
 \end{aligned}$$

Wir arbeiten nachfolgend auf den Determinanten-Multiplikations-Satz hin, welcher aussagt, wie die Determinante eines Matrixprodukts $A \cdot B$ für $n \times n$ -Matrizen A und B ermittelt werden kann ohne die Multiplikation selbst durchzuführen.

In Analogie zu Eigenschaft (D6) ziehen wir aus (2.7.15) die

Folgerung 2.7.19

Entsteht \tilde{A} aus A durch das Vertauschen zweier Spalten, so ist $\det(\tilde{A}) = -\det(A)$.

Beweis.

Das Vertauschen zweier Spalten von A ist äquivalent zum Vertauschen zweier Zeilen von A^t und es gilt

$$\begin{aligned}
 \det(\tilde{A}) &\stackrel{(2.7.15)}{=} \det(\tilde{A}^t) \\
 &\stackrel{(D6)}{=} -\det(\tilde{A}^t) \\
 &= -\det(A).
 \end{aligned}$$

□

Beispiel 2.7.20. Zu $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ und $\tilde{A} = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$ sind $\det(A) = -2$ und $\det(\tilde{A}) = 2$.

2 Matrixrechnung

Bemerkung 2.7.21. Für $A \in \text{Mat}(n \times n, \mathbb{C})$ gilt genau dann $\text{rg}(A) < n$, wenn $\text{rg}(A^t) < n$.

Beweis.

Nach (D9) ist $\text{rg}(A) < n$ äquivalent zu $\det(A) = 0$. Nach Satz (2.7.15) ist damit $\det(A) = \det(A^t) = 0$. Dies ist wiederum nach (D9) äquivalent zu $\text{rg}(A^t) < n$. \square

Satz 2.7.22

Für alle $A \in \text{Mat}(n \times n, \mathbb{C})$ gilt $\text{rg}(A) = \text{rg}(A^t)$.

Der Nachweis hierfür ist ziemlich technisch. Wir verzichten daher darauf.

Bemerkung 2.7.23. Alle Sätze und Beweise bezüglich linearer Unabhängigkeit, etc. (vgl. Abschnitt 2.2) gelten ohne Einschränkung auch für komplexe Zahlen.

Lemma 2.7.24

Seien $A, B \in \text{Mat}(n \times n, \mathbb{C})$ mit $\text{rg}(A) < n$. Dann existieren $x, y \in \mathbb{C}^n \setminus \{\mathbf{0}\}$, so dass

$$Ax = \mathbf{0} \quad \text{und} \quad y^t A = \mathbf{0}$$

gilt. Ferner ist

$$\text{rg}(BA) < n \quad \text{sowie} \quad \text{rg}(AB) < n.$$

Beweis.

Ist $\text{rg}(A) < n$, so sind die Spaltenvektoren $a_1, \dots, a_n \in \mathbb{C}^n$ von

$$A = (a_1, \dots, a_n)$$

linear abhängig, d.h. es existieren $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ nicht alle gleich 0, so dass

$$\lambda_1 a_1 + \dots + \lambda_n a_n = \mathbf{0}$$

gilt. Wählen wir $x = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ und schreiben

$$Ax = (a_1, \dots, a_n) \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 a_1 + \dots + \lambda_n a_n = \mathbf{0},$$

also ist

$$Ax = \mathbf{0}$$

und es folgt

$$(BA)x = B(Ax) = B\mathbf{0} = \mathbf{0}.$$

Hierbei ist $BA = (c_1, \dots, c_n)$ mit $c_1, \dots, c_n \in \mathbb{C}^n$ linear abhängig. Folglich ist $\text{rg}(BA) < n$. Aus $\text{rg}(A) < n$ resultiert nach Bemerkung (2.7.21) zudem $\text{rg}(A^t) < n$, d.h. die Spaltenvektoren von A^t sind linear abhängig. Dies ist jedoch genau dann der Fall, wenn die Zeilenvektoren von A linear abhängig sind. Eine analoge Argumentation wie eben ergibt ein y^t , so dass $y^t A = \mathbf{0}$ gilt. Damit ist dann $\text{rg}(AB) < n$. \square

Satz 2.7.25 (Determinanten-Multiplikations-Satz)

Für alle $A, B \in \text{Mat}(n \times n, \mathbb{C})$ gilt:

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

Beweis.

1. Fall: Sei $\det(B) = 0$.

Nach (D9) gilt dann $\text{rg}(B) < n$ und folglich $\text{rg}(AB) < n$. Wiederum nach (D9) ist nun

$$\det(AB) = 0 = \det(A) \cdot 0.$$

Es folgt

$$\det(AB) = \det(A) \cdot \det(B).$$

2. Fall: Wir fixieren B mit $\det(B) \neq 0$.

Wie im Beweis zu Satz (2.7.15) definieren wir nun für alle $A \in \text{Mat}(n \times n, \mathbb{C})$ eine Funktion

$$\widetilde{\det} : \text{Mat}(n \times n, \mathbb{C}) \rightarrow \mathbb{C} \quad \text{durch die Vorschrift} \quad \widetilde{\det}(A) := \frac{\det(AB)}{\det(B)}.$$

2 Matrixrechnung

Wir zeigen wieder, dass $\widetilde{\det}$ die Eigenschaften (D1), (D2), (D3) der Determinanten-Funktion erfüllt. Wegen der Eindeutigkeit von \det folgt dann $\widetilde{\det} = \det$.

Zu (D1): Wir müssen zeigen, dass $\widetilde{\det}$ linear in jeder Zeile ist.

Teil 1: Sei

$$\hat{A} := \underbrace{\begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & \lambda & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}}_{=:\Delta} \cdot A.$$

die Matrix, die durch Multiplikation der i -ten Zeile von A mit $\lambda \in \mathbb{C}$ entsteht. Dann resultiert

$$\hat{A}B = \Delta(AB),$$

also das Produkt AB multipliziert mit λ in der i -ten Zeile. Damit folgt aus (D1)

$$\det(\hat{A}B) = \det(\Delta AB) = \lambda \det(AB).$$

Ferner ist

$$\widetilde{\det}(\hat{A}) = \frac{\det(\hat{A}B)}{\det(B)} = \lambda \frac{\det(AB)}{\det(B)} = \lambda \widetilde{\det}(A).$$

Teil 2: Seien $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{C}^n$. Wir drücken das Matrixprodukt $A \cdot B$ mit den Zeilen $a_i^t = (a_{i1}, \dots, a_{in})$ von A und den Spalten b_j von B für $i, j \in \{1, \dots, n\}$ aus. Dies ergibt:

$$\begin{aligned} A \cdot B &= \begin{pmatrix} a_1^t \\ \vdots \\ a_n^t \end{pmatrix} \cdot (b_1, \dots, b_n) \\ &= \begin{pmatrix} a_1^t b_1 & \cdots & a_1^t b_n \\ \vdots & & \vdots \\ a_i^t b_1 & \cdots & a_i^t b_n \\ \vdots & & \vdots \\ a_n^t b_1 & \cdots & a_n^t b_n \end{pmatrix}. \end{aligned}$$

Für $a'_i, a''_i \in \mathbb{C}^n$ betrachten wir nun $a_i = a'_i + a''_i$ und erhalten⁶

$$\begin{aligned} A \cdot B &= \begin{pmatrix} a_1^t b_1 & \cdots & a_1^t b_n \\ \vdots & & \vdots \\ (a'_i + a''_i)^t b_1 & \cdots & (a'_i + a''_i)^t b_n \\ \vdots & & \vdots \\ a_n^t b_1 & \cdots & a_n^t b_n \end{pmatrix} \\ &= \begin{pmatrix} a_1^t b_1 & \cdots & a_1^t b_n \\ \vdots & & \vdots \\ (a'_i)^t b_1 + (a''_i)^t b_1 & \cdots & (a'_i)^t b_n + (a''_i)^t b_n \\ \vdots & & \vdots \\ a_n^t b_1 & \cdots & a_n^t b_n \end{pmatrix}. \end{aligned}$$

Für

$$A' := \begin{pmatrix} a_1^t \\ \vdots \\ (a'_i)^t \\ \vdots \\ a_n^t \end{pmatrix} \quad \text{und} \quad A'' := \begin{pmatrix} a_1^t \\ \vdots \\ (a''_i)^t \\ \vdots \\ a_n^t \end{pmatrix}$$

gilt folglich

$$\det(A \cdot B) = \det(A' \cdot B) + \det(A'' \cdot B).$$

Somit ist

$$\widetilde{\det}(A) = \frac{\det(AB)}{\det(B)} = \frac{\det(A'B) + \det(A''B)}{\det(B)} = \widetilde{\det}(A') + \widetilde{\det}(A'').$$

Zu (D2): Wir zeigen: Die Übereinstimmung zweier Zeilen in A zieht $\widetilde{\det}(A) = 0$ nach sich.

Hat A zwei gleiche Zeilen, so gilt $\text{rg}(A) < n$ ebenso wie $\text{rg}(AB) < n$. Nun ist nach (D9) $\det(AB) = 0$ und es gilt

$$\widetilde{\det}(A) = \frac{\det(AB)}{\det(B)} = 0.$$

Zu (D3):

Es gilt

$$\widetilde{\det}(\mathbf{1}_n) = \frac{\det(\mathbf{1}_n B)}{\det(B)} = \frac{\det(B)}{\det(B)} = 1.$$

Insgesamt folgt somit

$$\det(A) = \widetilde{\det}(A) = \frac{\det(AB)}{\det(B)}$$

⁶Man überlege sich hierzu, dass für alle Matrizen $A_1, A_2 \in \text{Mat}(m \times n, \mathbb{C})$ gilt: $(A_1 + A_2)^t = A_1^t + A_2^t$.

2 Matrixrechnung

für alle $A \in \text{Mat}(n \times n, \mathbb{C})$. Multiplikation dieser Zeile mit $\det(B)$ liefert

$$\det(AB) = \det(A) \cdot \det(B).$$

□

In den komplexen Zahlen kann die Multiplikation mit $z \neq 0$ stets durch die Division durch z bzw. die Multiplikation mit z^{-1} rückgängig gemacht werden, d.h. es gilt dann $z \cdot z^{-1} = 1$. Hierbei nennt man z^{-1} auch das zu z inverse Element. Für eine bestimmte Teilmenge von $\text{Mat}(n \times n, \mathbb{C})$ liefert uns der Determinanten-Multiplikations-Satz vergleichbare Möglichkeiten.

Definition 2.7.26. Eine Matrix $A \in \text{Mat}(n \times n, \mathbb{C})$ heißt **invertierbar**, falls eine Matrix $B \in \text{Mat}(n \times n, \mathbb{C})$ existiert, so dass

$$AB = BA = \mathbb{1}_n$$

ist. Die Matrix B nennt man **inverse Matrix** zu A und erhält das neue Symbol A^{-1} . Die Menge aller invertierbaren $n \times n$ -Matrizen bezeichnen wir mit $\text{GL}(n, \mathbb{C})$. Dabei steht GL für General Linear und $\text{GL}(n, \mathbb{C})$ heißt auch **allgemeine lineare Gruppe**.

Bemerkung 2.7.27. Ist $A \in \text{GL}(n, \mathbb{C})$, so ist $\det(A) \neq 0$ und $\det(A^{-1}) = \frac{1}{\det(A)}$.

Beweis.

Es gilt

$$1 = \det(\mathbb{1}_n) = \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1}).$$

Damit ist $\det(A) \neq 0$ und $\det(A^{-1}) = \frac{1}{\det(A)}$.

□

Lemma 2.7.28

Ist $A \in \text{GL}(n, \mathbb{C})$, so ist die inverse Matrix A^{-1} eindeutig.

Beweis.

Seien A^{-1} und \tilde{A}^{-1} inverse Matrizen zu A . Dann gilt

$$A\tilde{A}^{-1} = \mathbb{1}_n$$

Multiplikation mit A^{-1} liefert

$$\underbrace{A^{-1} \cdot A}_{=\mathbb{1}_n} \tilde{A}^{-1} = A^{-1} \cdot \mathbb{1}_n$$

und somit

$$\tilde{A}^{-1} = A^{-1}.$$

□

Welche Anwendungen für Determinanten ergeben sich uns? Wir beginnen mit linearen Gleichungssystemen. Ist $A \in \text{GL}(n, \mathbb{C})$, so hat das LGS

$$A \cdot x = b \tag{2.30}$$

genau eine Lösung. Sofern wir Kenntnis von der zu A inversen Matrix A^{-1} haben, kann die Gleichung (2.30) wie folgt nach x aufgelöst werden:

$$\begin{aligned} A \cdot x &= b \\ \Leftrightarrow A^{-1} \cdot A \cdot x &= A^{-1} \cdot b \\ \Leftrightarrow \mathbb{1}_n \cdot x &= A^{-1} \cdot b \\ \Leftrightarrow x &= A^{-1} \cdot b. \end{aligned}$$

Ob eine quadratische Matrix A invertierbar ist, erkennt man daran, ob $\det(A) \neq 0$ gilt. Die konkrete Berechnung von A^{-1} kann man ebenfalls mittels Determinanten bewerkstelligen.

Satz 2.7.29

Sei $A \in \text{Mat}(n \times n, \mathbb{C})$. Sei $B \in \text{Mat}(n \times n, \mathbb{C})$ definiert durch

$$B_{ij} := (-1)^{i+j} \cdot \det(A_{ji}^{\text{Str}}).$$

Dann gilt:

$$A \cdot B = B \cdot A = \det(A) \cdot \mathbb{1}_n.$$

Ist A invertierbar, d.h. gilt $A \in \text{GL}(n, \mathbb{C})$, dann gilt also:

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

Beispiel 2.7.30. Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ erhalten wir

$$B = \begin{pmatrix} +d & -b \\ -c & +a \end{pmatrix} \quad \text{und} \quad \det(A) = ad - bc,$$

2 Matrixrechnung

also insgesamt

$$A^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Beweis.

Zunächst zeigen wir $A \cdot B = \det(A) \cdot \mathbb{1}_n$.

1. Fall:

Die Diagonalelemente von $A \cdot B$, d.h. die (i, i) -ten Einträge sind gegeben durch

$$\begin{aligned} (A \cdot B)_{ii} &= \sum_{j=1}^n A_{ij} \cdot B_{ji} \\ &= \sum_{j=1}^n (-1)^{i+j} \cdot A_{ij} \cdot \det(A_{ij}^{\text{Str}}) \\ &\stackrel{2.7.17}{=} \det(A). \end{aligned}$$

2. Fall:

Für $i \neq j$ ist der (i, j) -te Eintrag von $A \cdot B$:

$$\begin{aligned} (A \cdot B)_{ij} &= \sum_{k=1}^n A_{ik} \cdot B_{kj} \\ &= \sum_{k=1}^n (-1)^{k+j} \cdot A_{ik} \cdot \det(A_{jk}^{\text{Str}}) \\ &= \sum_{k=1}^n (-1)^{k+j} \cdot \tilde{A}_{jk} \cdot \det(A_{jk}^{\text{Str}}), \end{aligned}$$

wobei \tilde{A} aus A dadurch hervorgeht, dass die j -te Zeile durch die i -te ersetzt wird. Ferner können wir A_{jk}^{Str} durch $\tilde{A}_{jk}^{\text{Str}}$ ersetzen, denn die j -te Zeile wird ohnehin im nächsten Schritt gestrichen. Wir erhalten somit

$$\begin{aligned} (A \cdot B)_{ij} &= \sum_{k=1}^n (-1)^{k+j} \cdot \tilde{A}_{jk} \cdot \det(\tilde{A}_{jk}^{\text{Str}}) \\ &\stackrel{2.7.17}{=} \det(\tilde{A}) \\ &= 0, \end{aligned}$$

da \tilde{A} zwei identische Zeilen hat. Also folgt $A \cdot B = \det(A) \cdot \mathbb{1}_n$.

Für die Gleichung $B \cdot A = \det(A) \cdot \mathbb{1}_n$ argumentiert man analog mit Spalten- statt Zeilenentwicklung. \square

Bemerkung 2.7.31. Im Falle einer **Diagonalmatrix**

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

gilt $\det(A) = \lambda_1 \cdot \dots \cdot \lambda_n \neq 0$ genau dann, wenn $\lambda_i \neq 0$ für alle $i \in \{1, \dots, n\}$ ist. In diesem Fall gilt

$$A^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{\lambda_n} \end{pmatrix}.$$

Die Möglichkeit, inverse Matrizen auszurechnen, hilft beim konkreten Lösen von LGS:

Satz 2.7.32 (Cramer'sche Regel)

Seien $a_1, \dots, a_n \in \mathbb{C}^n$ und sei $A = (a_1, \dots, a_n) \in \text{GL}(n, \mathbb{C})$. Sei ferner $b \in \mathbb{C}^n$. Dann ist

die eindeutige Lösung $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ des LGS

$$A \cdot x = b$$

gegeben durch

$$x_i = \frac{\det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n)}{\det(A)}.$$

Beweis.

Die Lösung ist gegeben durch $x = A^{-1} \cdot b$. Wir schreiben einerseits

$$\begin{aligned} x_i &= (A^{-1} \cdot b)_i \\ &= \sum_{j=1}^n (A^{-1})_{ij} \cdot b_j \\ &= \sum_{j=1}^n \frac{1}{\det(A)} (-1)^{i+j} \det(A_{ji}^{\text{Str}}) \cdot b_j \end{aligned}$$

2 Matrixrechnung

$$= \frac{1}{\det(A)} \cdot \sum_{j=1}^n (-1)^{i+j} \cdot \det(A_{ji}^{\text{Str}}) \cdot b_j. \quad (2.31)$$

Mittels Determinanten-Entwicklung nach der i -ten Spalte (und der Tatsache, dass die i -te Spalte b ohnehin gestrichen wird) gilt andererseits:

$$\det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n) = \sum_{j=1}^n (-1)^{i+j} \cdot b_j \cdot \det(A_{ji}^{\text{Str}}). \quad (2.32)$$

Setzen wir (2.32) in (2.31) ein, so erhalten wir die Behauptung des Satzes. \square

Beispiel 2.7.33. Wir betrachten das LGS $A \cdot x = b$ der konkreten Gestalt

$$\begin{pmatrix} 0 & 1 & i \\ 1 & i & 1 \\ 2 & 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Aus Beispiel (2.7.5) wissen wir bereits $\det \begin{pmatrix} 0 & 1 & i \\ 1 & i & 1 \\ 2 & 3 & 4 \end{pmatrix} = 3i \neq 0$. Somit ist A invertierbar und wir erhalten:

$$\begin{aligned} x_1 &= \frac{\det \begin{pmatrix} 1 & 1 & i \\ 0 & i & 1 \\ 0 & 3 & 4 \end{pmatrix}}{3i} = \frac{1 \cdot (4i - 3)}{3i} = \frac{4}{3} + i, \\ x_2 &= \frac{\det \begin{pmatrix} 0 & 1 & i \\ 1 & 0 & 1 \\ 2 & 0 & 4 \end{pmatrix}}{3i} = \frac{-1 \cdot 2}{3i} = \frac{2}{3}i, \\ x_3 &= \frac{\det \begin{pmatrix} 0 & 1 & 1 \\ 1 & i & 0 \\ 2 & 3 & 0 \end{pmatrix}}{3i} = \frac{1 \cdot (3 - 2i)}{3i} = -\frac{2}{3} - i, \end{aligned}$$

und damit die eindeutige Lösung des LGS zu

$$x = \begin{pmatrix} \frac{4}{3} + i \\ \frac{2}{3}i \\ -\frac{2}{3} - i \end{pmatrix}.$$

Nun kommen wir zu zwei markanten Typen von 2×2 -Matrizen mit besonderer geometrischer Bewandnis.

Definition 2.7.34. Eine Matrix der Form

$$R_\theta := \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \in \text{Mat}(2 \times 2, \mathbb{R})$$

mit $\theta \in \mathbb{R}$ heißt **Drehmatrix**.

Für sie gilt $\det(R_\theta) = \cos^2(\theta) + \sin^2(\theta) = 1$.

Ferner liefert die Multiplikation einer Matrix mit dem i -ten Standard-Basis-Vektor stets den i -ten Spaltenvektor der Matrix, so auch hier:

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix},$$

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}.$$

Die geometrische Interpretation ist nun die folgende:

R_θ beschreibt eine Drehung im mathematisch positiven Drehsinn (also entgegen dem Uhrzeigersinn) um den Winkel θ und Drehpunkt $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

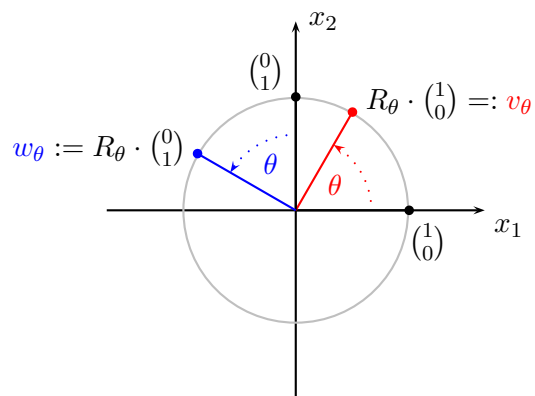


Abb. 49

Die Vektoren v_θ und w_θ helfen bei der Interpretation des folgenden, zu Drehmatrizen eng verwandten Typs von Matrizen:

Definition 2.7.35. Eine Matrix der Form

$$S_\theta := \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} \in \text{Mat}(2 \times 2, \mathbb{R})$$

mit $\theta \in \mathbb{R}$ heißt **Spiegelungsmatrix**.

2 Matrixrechnung

Es gilt $\det(S_\theta) = -\cos^2(2\theta) - \sin^2(2\theta) = -1$.

Sehen wir uns nun an, was die Anwendung von S_θ auf die Vektoren v_θ und w_θ bewirkt. Mittels der Additionstheoreme für Sinus und Cosinus finden wir:

$$\begin{aligned} S_\theta \cdot v_\theta &= \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} \cdot \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \\ &= \begin{pmatrix} \cos(2\theta) \cdot \cos(\theta) + \sin(2\theta) \cdot \sin(\theta) \\ \sin(2\theta) \cdot \cos(\theta) - \cos(2\theta) \cdot \sin(\theta) \end{pmatrix} \\ &= \begin{pmatrix} \cos(2\theta) \cdot \cos(-\theta) - \sin(2\theta) \cdot \sin(-\theta) \\ \sin(2\theta) \cdot \cos(-\theta) + \cos(2\theta) \cdot \sin(-\theta) \end{pmatrix} \\ &= \begin{pmatrix} \cos(2\theta - \theta) \\ \sin(2\theta - \theta) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \\ &= v_\theta \end{aligned}$$

und analog:

$$S_\theta \cdot w_\theta = -w_\theta.$$

Somit wird v_θ auf sich und w_θ auf sein Negatives abgebildet. Ferner bildet $\{v_\theta, w_\theta\}$ offenbar eine Basis des \mathbb{R}^2 , d.h. jedes $x \in \mathbb{R}^2$ kann als

$$x = \lambda_1 v_\theta + \lambda_2 w_\theta$$

mit $\lambda_1, \lambda_2 \in \mathbb{R}$ dargestellt werden. Man rechnet leicht

$$\begin{aligned} S_\theta(x) &= S_\theta(\lambda_1 v_\theta + \lambda_2 w_\theta) \\ &= \lambda_1 S_\theta(v_\theta) + \lambda_2 S_\theta(w_\theta) \\ &= \lambda_1 v_\theta - \lambda_2 w_\theta \end{aligned}$$

nach, womit S_θ genau die Spiegelung an der Achse $\mathbb{R} \cdot v_\theta$ beschreibt:

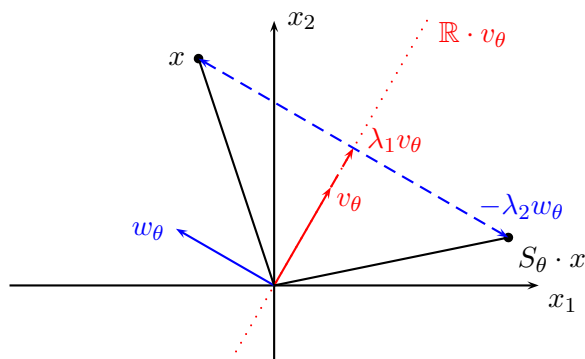


Abb. 50

Als weitere Anwendung von Determinanten wenden wir uns Volumenberechnungen zu. Dies macht folgenden vorbereitenden Abschnitt notwendig.

2.8 Affine Abbildungen und affine Unterräume

Die hier gemachten Aussagen legen den Zahlkörper $K = \mathbb{R}$ bzw. $K = \mathbb{C}$ zu Grunde.

Im Rahmen unserer Überlegungen zu Lösungsmengen von LGS wurden Untervektorräume (UVR) thematisiert. In geometrischer Interpretation stellten sich dabei die 1-dimensionalen UVRe des \mathbb{R}^2 als genau die Ursprungsgeraden und die 2-dimensionalen UVRe des \mathbb{R}^3 als genau die Ursprungsebenen heraus. Ausgehend von UVRen lernen wir nun eine Möglichkeit kennen, auch solche Geraden bzw. Ebenen darzustellen, die die $\mathbf{0}$ nicht enthalten. Das Stichwort hierbei ist „affin“. Dieser Begriff bedeutet sinngemäß soviel wie „verwandt“ und beinhaltet im Speziellen, dass gewisse Eigenschaften geometrischer Objekte erhalten bleiben. Gießen wir die Umgangssprache in eine mathematische Funktion, so kommt das dabei heraus:

Definition 2.8.1. Eine Abbildung

$$F : K^n \rightarrow K^m$$

heißt genau dann **affin**, wenn es ein $A \in \text{Mat}(m \times n, K)$ und ein $b \in K^m$ gibt, so dass für alle $x \in K^n$ gilt:

$$F(x) = A \cdot x + b.$$

Beispiel 2.8.2. Für $K = \mathbb{R}$ und $n = m = 2$ realisiert die Rotationsmatrix R_θ die Drehung um θ zunächst nur um den Drehpunkt $\mathbf{0}$. Soll $x \in \mathbb{R}^2$ um einen anderen Dreh-Mittelpunkt $M \in \mathbb{R}^2$ rotieren, so nehmen wir zunächst eine Verschiebung vor, bei der M in den Ursprung überführt wird, drehen dann x um den Ursprung und verschieben wieder zurück. In Formeln:

$$\begin{aligned} F(x) &= R_\theta \cdot (x - M) + M \\ &= R_\theta \cdot x - R_\theta \cdot M + M \\ &=: A \cdot x + b, \end{aligned}$$

wobei nun $A = R_\theta$ und $b = M - R_\theta \cdot M$ sind.

2 Matrixrechnung

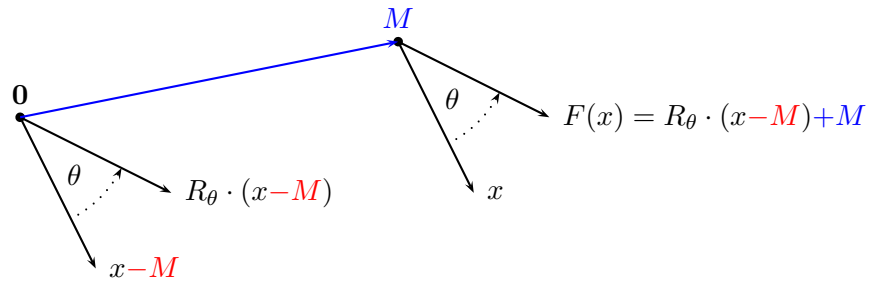


Abb. 51

Ähnlich lässt sich eine affine Abbildung für die Spiegelung an einer Achse formulieren, die nicht durch 0 geht.

Kommen wir schließlich zur Verallgemeinerung von Geraden in Dimension 1 und Ebenen in Dimension 2.

Definition 2.8.3. Eine Teilmenge $U \subset K^n$ heißt genau dann **affiner Unterraum**, wenn es einen UVR $U_0 \subset K^n$ und ein $x \in U$ gibt, so dass

$$U = \{y + x \mid y \in U_0\}$$

gilt.

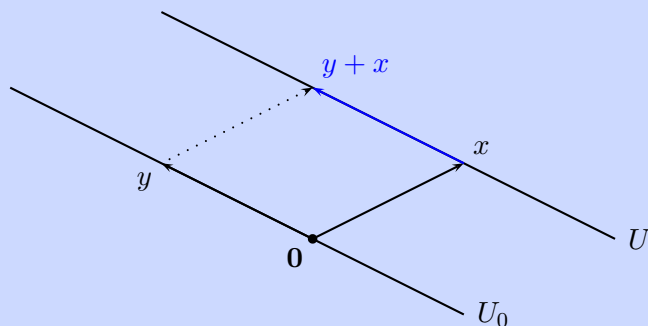


Abb. 52

Im \mathbb{R}^2 ist durch den UVR U_0 eine Ursprungsgerade und mit U ein 1-dimensionaler affiner Unterraum gegeben.

Bemerkung 2.8.4. Stets sind UVRs zugleich affine Unterräume, umgekehrt gilt dies jedoch im Allgemeinen nicht.

Bemerkung 2.8.5. Der UVR U_0 kann aus U durch Bildung der Menge aller Differenzvektoren $u - v$ für $u, v \in U$ (zurück-)gewonnen werden, d.h. es gilt

$$U_0 = \{u - v \mid u, v \in U\}.$$

Insbesondere ist damit U_0 durch U eindeutig festgelegt.

Beweis.

„ \supset “: Sind $u, v \in U$, so existieren $y, y' \in U_0$, so dass $u = y + x$ und $v = y' + x$ gilt. Somit folgt

$$\begin{aligned} u - v &= (y + x) - (y' + x) \\ &= y - y' \\ &= y + (-y') \\ &\in U_0. \end{aligned}$$

„ \subset “: Sei $y \in U_0$. Für $x \in U$ setzen wir $u := y + x$ und $v := \mathbf{0} + x$. Dann sind $u, v \in U$ und wir erhalten

$$\begin{aligned} u - v &= (y + x) - x \\ &= y. \end{aligned}$$

Somit ist y von der Form $u - v$ mit $u, v \in U$. □

Bemerkung 2.8.6. Der „Aufpunkt“ x eines affinen Unterraumes kann durch ein beliebiges $x' \in U$ ersetzt werden und ist damit insbesondere durch U *nicht eindeutig* festgelegt. Für jedes $x' \in U$ gilt demnach

$$U = \{y + x' \mid y \in U_0\}.$$

Beweis.

Zunächst ist $U = \{y + x \mid y \in U_0\}$ per Definition. Nun setzen wir $U' := \{y + x' \mid y \in U_0\}$ und zeigen $U = U'$.

„ \subset “: Ist $u \in U$, so existiert ein $y \in U_0$, so dass

$$\begin{aligned} u &= y + x \\ &= \underbrace{y}_{\in U_0} + \underbrace{(x - x')}_{\in U_0} + x' \\ &\quad \underbrace{\hspace{1.5cm}}_{\in U_0} \\ &\in U'. \end{aligned}$$

„ \supset “: Hier argumentiert man analog durch Vertauschen der Rollen von x und x' . □

Definition 2.8.7. $\dim(U) := \dim(U_0)$ heißt die **Dimension des affinen Unterraums** U .

Bemerkung 2.8.8. Sei $U \subset K^n$ ein affiner Unterraum. Dann ist U ein UVR $\Leftrightarrow \mathbf{0} \in U$.

Beweis.

„ \Rightarrow “: Klar nach Definition eines UVRs.

„ \Leftarrow “: Wir wählen $x = \mathbf{0}$ als Aufpunkt von U und erhalten $U = \{y + \mathbf{0} \mid y \in U_0\} = U_0$. \square

Beispiel 2.8.9. Hier noch einmal zur Zusammenfassung die wichtigen Fälle unserer Anschauung:

dim	affine Unterräume	Untervektorräume
1	Geraden	Ursprungsgeraden
2	Ebenen	Ursprungsebenen

Wie hängen affine Abbildungen und affine Unterräume zusammen?

Bemerkung 2.8.10. Ist $U \subset K^n$ ein affiner Unterraum und ist $F : K^n \rightarrow K^m$ eine affine Abbildung, so ist auch

$$F(U) \subset K^m$$

ein affiner Unterraum.

2.9 Volumina

In diesem Abschnitt verwenden wir Determinanten in geometrischem Kontext, nämlich zur Volumenberechnung. Um unnötige Wiederholungen und überflüssigen Arbeitsaufwand zu vermeiden, behandeln wir alle Dimensionen auf einmal. Dabei haben wir uns unter einem eindimensionalen Volumen eine Länge, einem zweidimensionalen Volumen einen Flächeninhalt und einem dreidimensionalen Volumen das anschauliche Volumen eines Körpers im dreidimensionalen Raum vorzustellen. Die Theorie funktioniert aber auch für Volumina in Dimension ≥ 4 .

Systematisch entwickelt man diese Theorie in der so genannten *Maßtheorie*, einer mathematischen Disziplin, die eng mit der Analysis und der Wahrscheinlichkeitstheorie zusammenhängt. Wir werden hier lediglich einige Eigenschaften der Volumenfunktion benutzen ohne uns um die Einzelheiten ihrer Konstruktion zu kümmern.

Legen wir zunächst fest, von welchen Mengen wir das Volumen untersuchen werden.

Sie \mathcal{K}_n die Menge der n -dimensionalen reellen Kompakta, also

$$\begin{aligned}\mathcal{K}_n &:= \{X \subset \mathbb{R}^n \mid X \text{ ist beschränkt und abgeschlossen}\} \\ &= \{X \subset \mathbb{R}^n \mid X \text{ kompakt}\}.\end{aligned}$$

Hierbei bedeutet beschränkt, dass X sich nicht ins Unendliche erstreckt, d.h.

$$\text{es gibt einen Radius } R > 0, \text{ so dass } X \subset B_R(\mathbf{0}) := \{x \in \mathbb{R}^n \mid \|x\| \leq R\}$$

und abgeschlossen, dass der Rand ∂X zu X dazugehört, also $\partial X \subset X$, d.h. jedes $x \in \mathbb{R}^n$, für das es eine Folge (x_i) von Punkten aus X gibt mit $\lim_{i \rightarrow \infty} (x_i) = x$, muss selbst zu X gehören, $x \in X$.

Definition 2.9.1. Das n -dimensionale Volumen ist eine Abbildung

$$\text{vol}_n : \mathcal{K}_n \rightarrow \mathbb{R}$$

mit $\text{vol}_n(X) \geq 0$ für alle $X \in \mathcal{K}_n$, so dass gilt:

- $\text{vol}_n(W_n) = 1$, wobei $W_n = \underbrace{[0, 1] \times \dots \times [0, 1]}_{n\text{-mal}} \subset \mathbb{R}^n$ der n -dimensionale Einheitswürfel ist.
- Ist $X \subset Y$, so gilt $\text{vol}_n(X) \leq \text{vol}_n(Y)$.
- Es gilt $\text{vol}_n(X \cup Y) = \text{vol}_n(X) + \text{vol}_n(Y) - \text{vol}_n(X \cap Y)$ für alle $X, Y \in \mathcal{K}_n$.
- Ist $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine affine Abbildung, d.h. $F(x) = Ax + b$ mit $A \in \text{Mat}(n \times n, \mathbb{R})$ und $b \in \mathbb{R}^n$, dann gilt für alle $X \in \mathcal{K}_n$:

$$\text{vol}_n(F(X)) = |\det(A)| \cdot \text{vol}_n(X).$$

Die vierte Bedingung stellt sicher, dass sich das Volumen „richtig“ verhält. Bei Verschiebungen, d.h. $A = \mathbb{1}_n$, ändert sich das Volumen nicht. Strecken wir das Kompaktum in eine Richtung mit dem Faktor r , so sollte sich das Volumen mit r multiplizieren.

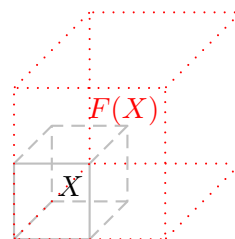
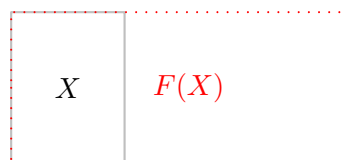


Abb. 53

2 Matrixrechnung

So erwarten wir z.B. für $\det \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} = 3$, dass sich das 2-dimensionale Volumen eines Rechtecks verdreifacht (Streckung in e_1 -Richtung mit dem Faktor 3). Für die Streckung in alle drei Richtungen mit dem Faktor 2, d.h. für $\det \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = 8$, wird sich das 3-dimensionale Volumen verachtfachen, denn $2^3 = 8$.

Definition 2.9.2. Eine Teilmenge $X \subset \mathbb{R}^n$ der Form

$$X = \{q + t_1 b_1 + \dots + t_n b_n \mid t_i \in [0, 1]\}$$

für eine Basis $\{b_1, \dots, b_n\}$ von \mathbb{R}^n und $q \in \mathbb{R}^n$ heißt **n-dimensionales Parallelotop**.

Beispiel 2.9.3. Für $n = 2$ erhalten wir im Allgemeinen ein Parallelogramm, im Falle $n = 3$ ein Parallelepiped.

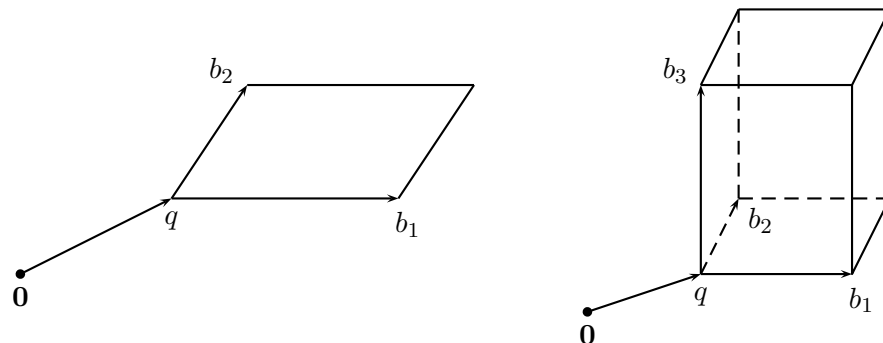


Abb. 54

Für die nun folgende Berechnung des Volumens kommen wir allein mit den Eigenschaften (1) und (4) von vol_n aus. Sei dazu $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $F(x) = Ax + b$ die affine Abbildung mit $A = (b_1, \dots, b_n)$ und $b = q$, wobei b_1, \dots, b_n das n -dimensionale Parallelotop X aufspannen. Mittels der Definition von W_n rechnet man leicht nach:

$$F(W_n) = X.$$

Damit folgt:

$$\begin{aligned}
 \operatorname{vol}_n(X) &= \operatorname{vol}_n(F(W_n)) \\
 &\stackrel{(d)}{=} |\det(A)| \cdot \operatorname{vol}_n(W_n) \\
 &\stackrel{(a)}{=} |\det(A)| \cdot 1 \\
 &= |\det(b_1, \dots, b_n)|.
 \end{aligned}$$

Für weitere Volumenberechnungen benötigen wir folgende Vorbereitung:

Bemerkung 2.9.4. Ist $U \subset \mathbb{R}^n$ ein UVR der Dimension $n - 1$, so können wir die Orthogonal-Projektion P auf U in Matrixschreibweise notieren. Wir wählen dazu einen der beiden normierten, auf U senkrecht stehenden Vektoren $v \in \mathbb{R}^n$ mit Fußpunkt in $\mathbf{0}$, d.h. es soll $v \perp U$ sowie $\|v\| = 1$ gelten. Damit bildet die Orthogonal-Projektion P den i -ten Standard-Basis-Vektor e_i auf

$$\begin{aligned}
 P \cdot e_i &= e_i - \langle e_i, v \rangle \cdot v \\
 &= e_i - v_i \cdot v
 \end{aligned}$$

ab. Also haben wir

$$P = (e_1 - v_1 v, \dots, e_n - v_n v) \in \operatorname{Mat}(n \times n, \mathbb{R})$$

und damit insbesondere

$$\begin{aligned}
 P \cdot v &= (e_1 - v_1 v) \cdot v_1 + \dots + (e_n - v_n v) \cdot v_n \\
 &= \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} - \underbrace{(v_1^2 + \dots + v_n^2)}_{=\|v\|^2=1} \cdot v \\
 &= \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} - v \\
 &= \mathbf{0}.
 \end{aligned}$$

Für Vektoren, die bereits in U liegen, d.h. für alle $x \in U$, also $\langle x, v \rangle = 0$, gilt $P \cdot x = x$, denn

$$\begin{aligned}
 P \cdot x &= (e_1 - v_1 v) \cdot x_1 + \dots + (e_n - v_n v) \cdot x_n \\
 &= \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} - \underbrace{(v_1 x_1 + \dots + v_n x_n)}_{=\langle v, x \rangle = 0} \cdot v \\
 &= x.
 \end{aligned}$$

Das folgende Lemma verallgemeinert die Tatsache, dass etwa ein Zylinder der Höhe 0, welcher dann lediglich eine Kreisscheibe in einer Ebene des \mathbb{R}^3 darstellt, keinerlei 3-dimensionales Volumen besitzt bzw. dieses gleich 0 ist.

2 Matrixrechnung

Lemma 2.9.5

Sei $X \in \mathcal{K}_n$. Sei X ferner enthalten in einem $(n - 1)$ -dimensionalen affinen Unterraum $V \subset \mathbb{R}^n$. Dann ist

$$\text{vol}_n(X) = 0.$$

Beweis.

Mittels einer affinen Abbildung machen wir uns das Leben leichter und verschieben X in eine „günstige“ Lage. Verschiebungen sind affine Abbildungen von der Form $F(x) = \mathbb{1}_n \cdot x + b$. Da $\det(\mathbb{1}_n) = 1$ ändern sie das Volumen nicht. Wir wählen b so, dass $\mathbf{0} \in F(V)$, d.h. $F(V)$ ein UVR ist. Dann ist $F(X) \subset F(V)$.

Nun genügt es, $\text{vol}_n(F(X)) = \mathbf{0}$ nachzuweisen. Sei hierfür $P \in \text{Mat}(n \times n, \mathbb{R})$ die Matrix der Orthogonal-Projektion auf $F(V)$. Wir erhalten $P \cdot x = x$ für alle $x \in F(X)$ und es gilt zunächst

$$\begin{aligned} \text{vol}_n(F(X)) &= \text{vol}_n(P(F(X))) \\ &= |\det(P)| \cdot \text{vol}_n(F(X)). \end{aligned}$$

Wegen $P(\mathbb{R}^n) = F(V)$ ist

$$\begin{aligned} \text{rg}(P) &= \dim(P(\mathbb{R}^n)) \\ &= \dim(F(V)) \\ &= \dim(V) \\ &= n - 1 \\ &< n \end{aligned}$$

und damit $\det(P) = 0$. Mit diesem Ergebnis folgt

$$\text{vol}_n(F(X)) = 0.$$

□

Dreiecke. Wir kommen nun zu 2-dimensionalen Volumina $\text{vol}_2(\Delta)$ von Dreiecken Δ .

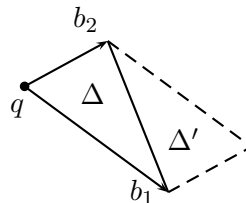


Abb. 55

Für das entstandene Parallelogramm bekommen wir

$$\begin{aligned}
 |\det(b_1, b_2)| &= \text{vol}_2(\Delta \cup \Delta') \\
 &\stackrel{(c)}{=} \text{vol}_2(\Delta) + \text{vol}_2(\Delta') - \underbrace{\text{vol}_2(\Delta \cap \Delta')}_{\stackrel{(2.9.5)}{=} 0} \\
 &= \text{vol}_2(\Delta) + \text{vol}_2(\Delta').
 \end{aligned}$$

Nun ist aber $\text{vol}_2(\Delta') = \text{vol}_2(\Delta)$, denn Δ' geht aus Δ durch Drehung hervor. Wegen $\det(R_\theta) = 1$ folgt somit:

$$\boxed{\text{vol}_2(\Delta) = \frac{1}{2} |\det(b_1, b_2)|}$$

Beispiel 2.9.6. Habe Δ die Ecken $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 2 \\ -2 \end{pmatrix}$. Wir nehmen $q = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ und erhalten

$$b_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \quad \text{sowie}$$

$$b_2 = \begin{pmatrix} 2 \\ -2 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}.$$

Damit ist

$$\text{vol}_2(\Delta) = \frac{1}{2} |\det(b_1, b_2)| = \frac{1}{2} \left| \det \begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix} \right| = \frac{1}{2} |-7| = \frac{7}{2}.$$

Die oben hergeleitete Formel eignet sich gut für Dreiecke, deren Ecken man explizit kennt. Häufig kennt man jedoch nur Seitenlängen und Winkel. Was macht man dann? Betrachten wir dazu das Dreieck Δ mit $0 < \gamma < \pi$.

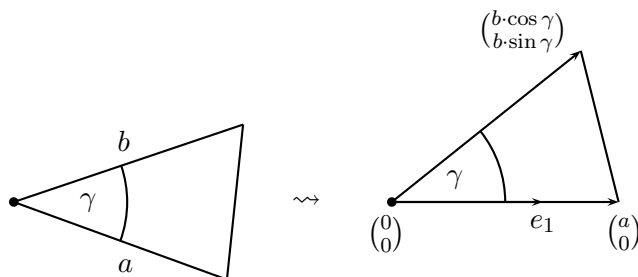


Abb. 56

2 Matrixrechnung

Ohne das 2-dimensionale Volumen von Δ zu ändern erreichen wir durch Verschiebung und Drehung, dass die Ecken von Δ durch $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} a \\ 0 \end{pmatrix}$ und

$$R_\gamma \cdot \begin{pmatrix} b \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) \\ \sin(\gamma) & \cos(\gamma) \end{pmatrix} \begin{pmatrix} b \\ 0 \end{pmatrix} = \begin{pmatrix} b \cdot \cos(\gamma) \\ b \cdot \sin(\gamma) \end{pmatrix}$$

gegeben sind. Wir setzen in die Volumenformel ein und erhalten

$$\begin{aligned} \text{vol}_2(\Delta) &= \frac{1}{2} \left| \det \begin{pmatrix} a & b \cos(\gamma) \\ 0 & b \sin(\gamma) \end{pmatrix} \right| \\ &= \frac{1}{2} |ab \sin(\gamma)|. \end{aligned}$$

Wir können nun also das 2-dimensionale Volumen eines Dreiecks, von dem wir die zwei Seitenlängen a, b und den eingeschlossenen Winkel $\gamma \in (0, \pi)$ kennen, nach folgender Formel berechnen:

$$\boxed{\text{vol}_2(\Delta) = \frac{1}{2} ab \sin(\gamma)}$$

Im Spezialfall rechtwinkliger Dreiecke mit $\gamma = \frac{\pi}{2}$ ergibt sich:

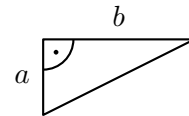


Abb. 57

$$\boxed{\text{vol}_2(\Delta) = \frac{ab}{2}}$$

Daraus erhält man für ein beliebiges Dreieck

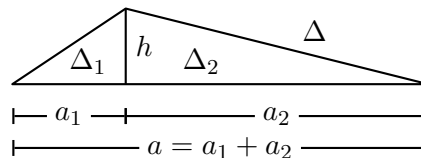


Abb. 58

eine Flächenformel in Termen der Grundlänge und der Höhe:

$$\begin{aligned} \text{vol}_2(\Delta) &= \text{vol}_2(\Delta_1 \cup \Delta_2) \\ &= \text{vol}_2(\Delta_1) + \text{vol}_2(\Delta_2) - \underbrace{\text{vol}_2(\Delta_1 \cap \Delta_2)}_{=0} \\ &= \frac{a_1 h}{2} + \frac{a_2 h}{2} \\ &= \frac{ah}{2}, \end{aligned}$$

also:

$$\text{vol}_2(\Delta) = \frac{1}{2}ah$$

Regelmäßige n -Ecke.

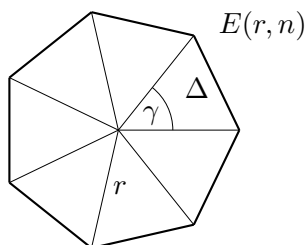


Abb. 59

Wir betrachten das n -Eck $E(r, n)$ und bemerken zunächst:

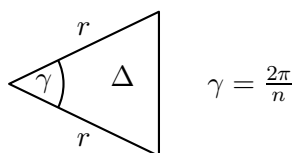


Abb. 60

Mittels der Formel für das 2-dimensionale Volumen dieses Dreiecks erhalten wir:

$$\text{vol}_2(E(r, n)) = \frac{nr^2}{2} \cdot \sin\left(\frac{2\pi}{n}\right)$$

Kreisschreiben. Wir approximieren das 2-dimensionale Volumen der Kreisscheibe von innen und von außen.

Dazu beschreiben wir zunächst der Kreisscheibe ein n -Eck $E(r, n)$ ein und erhalten eine untere Schranke für den Flächeninhalt.

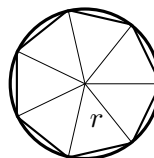


Abb. 61

Für das einbeschriebene n -Eck gilt $E(r, n) \subset D(r)$ für alle n , also

$$\begin{aligned} \text{vol}_2(D(r)) &\stackrel{(b)}{\geq} \text{vol}_2(E(r, n)) \\ &= \frac{nr^2}{2} \cdot \sin\left(\frac{2\pi}{n}\right). \end{aligned}$$

2 Matrixrechnung

Unser Wissen über das Verhalten von Limites und Grenzwerten bringt uns hier wie folgt weiter:

$$\begin{aligned}
 \text{vol}_2(D(r)) &\geq \lim_{n \rightarrow \infty} \left(\frac{nr^2}{2} \cdot \sin n \left(\frac{2\pi}{n} \right) \right) \\
 &\stackrel{x=\frac{2\pi}{n}}{=} \lim_{x \rightarrow 0} \left(\frac{2\pi}{x} \cdot \frac{r^2}{2} \cdot \sin(x) \right) \\
 &= \pi r^2 \cdot \underbrace{\lim_{x \rightarrow 0} \left(\frac{\sin(x)}{x} \right)}_{=1} \\
 &= \pi r^2.
 \end{aligned}$$

Eine untere Schranke für $\text{vol}_2(D(r))$ ist somit durch $\pi r^2 \leq \text{vol}_2(D(r))$ gegeben. Finden wir auch als obere Schranke πr^2 , so ist $\text{vol}_2(D(r)) = \pi r^2$.

Für das umschriebene n -Eck $E(r_n, n)$ erhalten wir aus der Beziehung $\cos\left(\frac{\pi}{n}\right) = \frac{r}{r_n}$ die Gleichung $r_n = \frac{r}{\cos\left(\frac{\pi}{n}\right)}$.

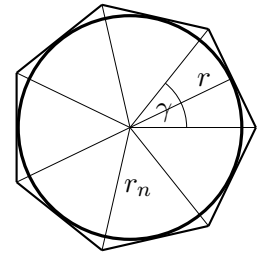


Abb. 62

Wir bekommen

$$\begin{aligned}
 \text{vol}_2(E(r_n, n)) &= \frac{nr_n^2}{2} \cdot \sin\left(\frac{2\pi}{n}\right) \\
 &= \frac{nr^2}{2 \cdot \left(\cos\left(\frac{\pi}{n}\right)\right)^2} \cdot \sin\left(\frac{2\pi}{n}\right).
 \end{aligned}$$

Aus $E(r_n, n) \supset D(r)$ folgt nun

$$\begin{aligned}
 \text{vol}_2(D(r)) &\leq \text{vol}_2(E(r_n, n)) \\
 &= \frac{nr^2 \sin\left(\frac{2\pi}{n}\right)}{2 \cdot \left(\cos\left(\frac{\pi}{n}\right)\right)^2},
 \end{aligned}$$

also

$$\begin{aligned}
 \text{vol}_2(D(r)) &\leq \lim_{n \rightarrow \infty} \frac{nr^2 \sin\left(\frac{2\pi}{n}\right)}{2 \left(\cos\left(\frac{\pi}{n}\right)\right)^2} \\
 &\stackrel{x=\frac{2\pi}{n}}{=} \lim_{x \rightarrow 0} \frac{\frac{2\pi}{x} \sin(x)}{2 \left(\cos\left(\frac{x}{2}\right)\right)^2} \\
 &= \pi r^2 \cdot \underbrace{\lim_{x \rightarrow 0} \frac{\sin(x)}{x}}_{\rightarrow 1} \cdot \underbrace{\frac{1}{\left(\cos\left(\frac{x}{2}\right)\right)^2}}_{\rightarrow 1} \\
 &= \pi r^2
 \end{aligned}$$

und insgesamt:

$$\boxed{\text{vol}_2(D(r)) = \pi r^2}$$

Ellipsen. Vom Inhalt der Einheitskreisscheibe $D(1)$ ausgehend, betrachten wir Gebiete, welche von Ellipsen $\mathcal{E}(a, b)$ mit den halben Hauptachsenlängen a und b berandet werden.

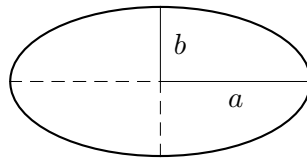


Abb. 63

Hierzu wird $D(1)$ in x_1 -Richtung um den Faktor a und in x_2 -Richtung um b mittels

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

gestreckt. Wir erhalten

$$\begin{aligned} \text{vol}_2(\mathcal{E}(a, b)) &= |\det(A)| \cdot \text{vol}_2(D(1)) \\ &= ab \cdot \pi, \end{aligned}$$

also:

$$\boxed{\text{vol}_2(\mathcal{E}(a, b)) = a \cdot b \cdot \pi}$$

2.10 Eigenwerte und Eigenvektoren

Meistens kann man einer Matrix nicht auf Anhieb ansehen, was sie geometrisch beschreibt. So sahen z.B. die Spiegelungsmatrizen S_θ auf den ersten Blick recht kompliziert aus, obwohl die Achsenspiegelungen, die sie beschreiben, geometrisch einfache Abbildungen sind. Es ist also wünschenswert, ein Verfahren zu entwickeln, das es uns erlaubt, gegebene Matrizen besser zu „verstehen“. Dazu dienen Eigenwerte und Eigenvektoren. Es stehe wieder K für \mathbb{R} oder \mathbb{C} .

Definition 2.10.1. Sei $A \in \text{Mat}(n \times n, K)$. Ein $\lambda \in K$ heißt genau dann **Eigenwert** von A , wenn ein $v \in K^n \setminus \{0\}$ existiert, so dass

$$A \cdot v = \lambda \cdot v \quad (2.33)$$

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

Ließe man $v = 0$ zu, so würde die Eigenwert-Gleichung (2.33) mit $v = 0$ für jedes λ gelten. Somit wäre jede Zahl λ Eigenwert und die Definition wäre unsinnig.

Die Gleichung (2.33) sagt aus, dass die Matrix A den Vektor v um den Faktor λ streckt.

Beispiel 2.10.2. Für die uns wohlbekannte Spiegelungsmatrix $A = S_\theta$ und $v = v_\theta$ gilt

$$\begin{aligned} A \cdot v &= S_\theta \cdot v_\theta \\ &= v_\theta \\ &= 1 \cdot v_\theta. \end{aligned}$$

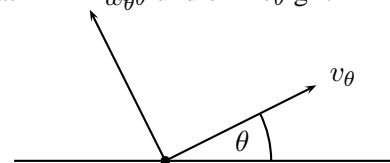


Abb. 64

Folglich ist v_θ Eigenvektor von S_θ zum Eigenwert 1. Für $v = w_\theta$ erhalten wir

$$\begin{aligned} A \cdot v &= S_\theta \cdot w_\theta \\ &= -w_\theta \\ &= (-1) \cdot w_\theta. \end{aligned}$$

Also ist w_θ Eigenvektor von S_θ zum Eigenwert -1 .

Beispiel 2.10.3. Denken wir für $K = \mathbb{R}$ einen Moment über die Drehmatrix $A = R_\theta$ mit $0 < \theta < \pi$ nach. Offenbar bleibt die Länge eines beliebigen Vektors bei der Drehung unverändert, und der Bildvektor ist gerade um den Winkel θ vom Original weggedreht, zeigt also in eine andere Richtung. Damit kann er insbesondere kein Vielfaches des Originals sein. Wir erwarten

also aufgrund dieser geometrischen Überlegung, dass R_θ keine Eigenvektoren und Eigenwerte hat. Wir werden dies in Beispiel 2.10.19 genauer abklären.

Beispiel 2.10.4. Für eine Diagonalmatrix

$$A = \Delta(\lambda_1, \dots, \lambda_n) := \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

beobachten wir

$$\Delta(\lambda_1, \dots, \lambda_n) \cdot e_k = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_k \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_k \cdot e_k.$$

Somit ist e_k Eigenvektor von $\Delta(\lambda_1, \dots, \lambda_n)$ zum Eigenwert λ_k für $k \in \{1, \dots, n\}$.

Woran liegt es nun, dass sich manchmal n , zuweilen 2 oder eben gar keine Eigenwerte ergeben? Wir beschreiten einen systematischen Zugang und beginnen mit folgendem

Satz 2.10.5

Ist $A \in \text{Mat}(n \times n, K)$ und sind v_1, \dots, v_k Eigenvektoren von A zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$, dann sind v_1, \dots, v_k linear unabhängig.

Beweis.

Wir führen eine vollständige Induktion nach dem Index k durch.

Induktionsanfang: Ist $k = 1$ und damit v_1 Eigenvektor zu λ_1 , so ist $v_1 \neq \mathbf{0}$ nach Definition, also linear unabhängig.

Induktionsschritt: Gelte die Aussage wie oben für ein $k \in \mathbb{N}$. Wir zeigen, dass die Aussage dann auch für $k + 1$ gültig ist. Hierzu seien v_1, \dots, v_{k+1} Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_{k+1}$. Wir überprüfen, dass nun v_1, \dots, v_{k+1} linear unabhängig sind, also dass aus

$$\alpha_1 v_1 + \dots + \alpha_{k+1} v_{k+1} = \mathbf{0} \tag{2.34}$$

2 Matrixrechnung

dann

$$\alpha_1 = \dots = \alpha_{k+1} = 0$$

folgt. Es gilt

$$\begin{aligned} \mathbf{0} &= A \cdot \mathbf{0} \stackrel{(2.34)}{=} A \cdot (\alpha_1 v_1 + \dots + \alpha_{k+1} v_{k+1}) \\ &= \alpha_1 \cdot A v_1 + \dots + \alpha_{k+1} \cdot A v_{k+1} \\ &\stackrel{(2.33)}{=} \alpha_1 \cdot \lambda_1 v_1 + \dots + \alpha_{k+1} \cdot \lambda_{k+1} v_{k+1}. \end{aligned} \quad (2.35)$$

Ferner erhalten wir aus Gleichung (2.34) durch Multiplikation mit λ_{k+1} , dass

$$\lambda_{k+1} \cdot \alpha_1 v_1 + \dots + \lambda_{k+1} \cdot \alpha_{k+1} v_{k+1} = \mathbf{0}. \quad (2.36)$$

Nun liefert Subtraktion der Gleichung (2.36) von (2.35)

$$\mathbf{0} = \alpha_1 \cdot (\lambda_1 - \lambda_{k+1}) \cdot v_1 + \dots + \alpha_k \cdot (\lambda_k - \lambda_{k+1}) \cdot v_k, \quad (2.37)$$

wobei der letzte Summand $\alpha_{k+1} \cdot (\lambda_{k+1} - \lambda_{k+1}) \cdot v_{k+1}$ gerade weggefallen ist.

Da nach Induktionsvoraussetzung v_1, \dots, v_k linear unabhängig sind, verschwinden in der Linearkombination der Gleichung (2.37) alle Koeffizienten $\alpha_i \cdot (\lambda_i - \lambda_{k+1})$, d.h. es folgt

$$\alpha_1 \cdot \underbrace{(\lambda_1 - \lambda_{k+1})}_{\neq 0} = \dots = \alpha_k \cdot \underbrace{(\lambda_k - \lambda_{k+1})}_{\neq 0} = 0,$$

also $\alpha_1 = \dots = \alpha_k = 0$. Setzen wir diese Ergebnisse in (2.34) ein, so verbleibt $\alpha_{k+1} v_{k+1} = \mathbf{0}$, also ist $\alpha_{k+1} = 0$, da der Eigenvektor v_{k+1} nach Definition nicht der Nullvektor ist. \square

Folgerung 2.10.6

Da es im K^n höchstens n linear unabhängige Vektoren gibt, folgt aus dem Satz insbesondere, dass eine Matrix $A \in \text{Mat}(n \times n, K)$ höchstens n verschiedene Eigenwerte besitzt.

Somit hat die Matrix $A = S_\theta \in \text{Mat}(2 \times 2, \mathbb{R})$ genau die beiden Eigenwerte 1 und -1 .

Als Vorbereitung für die konkrete Berechnung der Eigenwerte einer Matrix machen wir folgende

Definition 2.10.7. Ist $A \in \text{Mat}(n \times n, K)$ und $\lambda \in K$, so heißt

$$\text{Eig}(A, \lambda) := \{v \in K^n \mid Av = \lambda v\}$$

Eigenraum von A zum Eigenwert λ .

Zum Stichwort Eigenräume bietet es sich an, folgende Liste nützlicher Zusammenhänge zu bemerken:

Bemerkung 2.10.8. $\text{Eig}(A, \lambda) \setminus \{0\} = \{\text{Eigenvektoren von } A \text{ zum Eigenwert } \lambda\}$

Bemerkung 2.10.9. λ ist genau dann Eigenwert, wenn $\text{Eig}(A, \lambda) \neq \{0\}$, d.h. wenn der Eigenraum nicht trivial ist.

Bemerkung 2.10.10. $\text{Eig}(A, \lambda) = \{v \in K^n \mid \underbrace{(A - \lambda \cdot \mathbb{1}_n)}_{\in \text{Mat}(n \times n, K)} \cdot v = 0\} = \text{Lös}(A - \lambda \cdot \mathbb{1}_n, 0)$

Bemerkung 2.10.11. $\text{Eig}(A, \lambda) \subset K^n$ ist ein Untervektorraum.

Bemerkung 2.10.12. Für $\lambda_1 \neq \lambda_2$ ist $\text{Eig}(A, \lambda_1) \cap \text{Eig}(A, \lambda_2) = \{0\}$, denn wäre $0 \neq v \in \text{Eig}(A, \lambda_1) \cap \text{Eig}(A, \lambda_2)$, so wären v, v insbesondere linear abhängig im Widerspruch zu Satz (2.10.5).

Wenden wir uns zwei konkreten Beispielen zu:

Beispiel 2.10.13. Für $A = S_\theta$ erhalten wir

$$\begin{aligned} \text{Eig}(S_\theta, 1) &= \{\alpha \cdot v_\theta \mid \alpha \in \mathbb{R}\} && \text{sowie} \\ \text{Eig}(S_\theta, -1) &= \{\alpha \cdot w_\theta \mid \alpha \in \mathbb{R}\}. \end{aligned}$$

Ansonsten, d.h. für $\lambda \notin \{1, -1\}$, gilt:

$$\text{Eig}(S_\theta, \lambda) = \{0\}.$$

Deuten wir die beiden nichttrivialen Eigenräume einmal geometrisch:

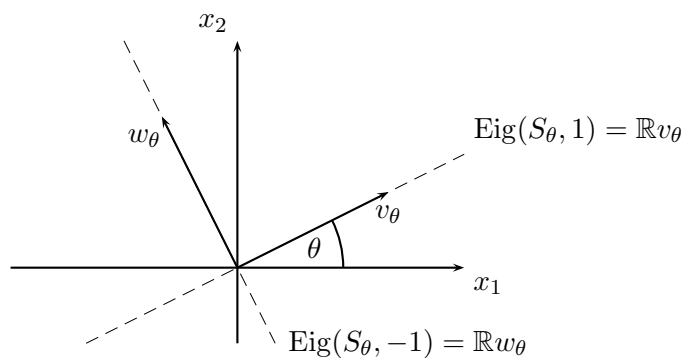


Abb. 65

2 Matrixrechnung

Beispiel 2.10.14. Die Matrix

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

könnte theoretisch bis zu drei Eigenwerte besitzen. Sofort sehen wir $\lambda_1 = 1$ und $\lambda_2 = 2$. Sollten das schon alle Eigenwerte sein? Wir bekommen

$$\text{Eig}(A, 1) = \left\{ \alpha \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \beta \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \mid \alpha, \beta \in \mathbb{R} \right\},$$

also einen 2-dimensionalen Eigenraum. Eine höhere Dimension tritt dagegen hier nicht auf, denn ein Eigenraum mit Dimension 3 ergäbe bereits den gesamten \mathbb{R}^3 und alle Vektoren des \mathbb{R}^3 würden somit auf sich abgebildet. Offenbar ist jedoch $A \neq \mathbb{1}_3$.

Wir erhalten weiter

$$\text{Eig}(A, 2) = \left\{ \alpha \cdot \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix} \mid \alpha \in \mathbb{R} \right\}.$$

Wie fasst man die Tatsache, dass der Eigenwert 1 doppelt gezählt wird?

Definition 2.10.15. $\dim(\text{Eig}(A, \lambda))$ heißt **geometrische Vielfachheit** oder **geometrische Multiplizität** des Eigenwerts λ .

Folgende Überlegung spendiert schließlich ein konkretes Verfahren zur konkreten Berechnung von Eigenwerten. Ist $A \in \text{Mat}(n \times n, K)$, so gilt:

$$\begin{aligned} \lambda \text{ ist Eigenwert von } A &\Leftrightarrow \text{Lös}(A - \lambda \cdot \mathbb{1}_n, \mathbf{0}) = \text{Eig}(A, \lambda) \neq \{\mathbf{0}\} \\ &\Leftrightarrow A - \lambda \cdot \mathbb{1}_n \text{ ist nicht invertierbar} \\ &\Leftrightarrow \det(A - \lambda \cdot \mathbb{1}_n) = 0. \end{aligned}$$

Definition 2.10.16. $\chi_A(\lambda) := \det(A - \lambda \cdot \mathbb{1}_n)$ heißt **charakteristisches Polynom** von A .

Wir sehen nun:

$$\lambda \text{ ist Eigenwert von } A \Leftrightarrow \chi_A(\lambda) = 0.$$

Beispiel 2.10.17. Überprüfen wir unsere Eigenwerte für $A = S_\theta$:

$$\begin{aligned} \chi_{S_\theta}(\lambda) &= \det(S_\theta - \lambda \cdot \mathbb{1}_n) \\ &= \det\left(\begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} - \lambda \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \\ &= \det\begin{pmatrix} \cos(2\theta) - \lambda & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) - \lambda \end{pmatrix} \\ &= (\cos(2\theta) - \lambda)(-\cos(2\theta) - \lambda) - \sin^2(2\theta) \\ &= \lambda^2 - \cos^2(2\theta) - \sin^2(2\theta) \\ &= \lambda^2 - (\cos^2(2\theta) + \sin^2(2\theta)) \\ &= \lambda^2 - 1 \\ &= (\lambda - 1)(\lambda + 1). \end{aligned}$$

Tatsächlich sind die Nullstellen von χ_{S_θ} die bekannten Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = -1$.

Beispiel 2.10.18. Für die Diagonalmatrix $A = \Delta(\lambda_1, \dots, \lambda_n)$ bekommen wir

$$\begin{aligned} \chi_{\Delta(\lambda_1, \dots, \lambda_n)} &= \det\left(\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} - \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}\right) \\ &= \det\begin{pmatrix} \lambda_1 - \lambda & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n - \lambda \end{pmatrix} \\ &= (\lambda_1 - \lambda) \cdot \dots \cdot (\lambda_n - \lambda). \end{aligned}$$

Also ist $\chi_{\Delta(\lambda_1, \dots, \lambda_n)}(\lambda) = 0$ genau dann, wenn ein $k \in \{1, \dots, n\}$ existiert, so dass $\lambda = \lambda_k$ gilt, d.h. $\Delta(\lambda_1, \dots, \lambda_n)$ hat genau die Eigenwerte $\lambda_1, \dots, \lambda_n$.

2 Matrixrechnung

Beispiel 2.10.19. Greifen wir noch einmal die in Beispiel 2.10.3 diskutierten Drehungen auf. Für $A = R_\theta$ ist zunächst

$$\begin{aligned}\chi_{R_\theta}(\lambda) &= \det \left(\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} - \lambda \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \\ &= \det \begin{pmatrix} \cos(\theta) - \lambda & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) - \lambda \end{pmatrix} \\ &= (\cos(\theta) - \lambda)^2 + \sin^2(\theta) \\ &= \lambda^2 - 2 \cos(\theta) \cdot \lambda + \cos^2(\theta) + \sin^2(\theta) \\ &= \lambda^2 - 2 \cos(\theta) \cdot \lambda + 1.\end{aligned}$$

Die Nullstellen von χ_{R_θ} sind nun gegeben durch

$$\lambda_{1,2} = \cos(\theta) \pm \sqrt{\cos^2(\theta) - 1},$$

jedoch ist für $0 < \theta < \pi$ der Ausdruck $\cos^2(\theta) - 1 < 0$ und somit sind λ_1 und λ_2 nicht reell. Drehmatrizen haben also keine reellen (wohl aber komplexe) Eigenwerte und Eigenvektoren.

Bemerkung 2.10.20. Nach dem Fundamentalsatz der Algebra besitzt für $K = \mathbb{C}$ jede $n \times n$ -Matrix Eigenwerte.

Auf dem Übungsblatt 13 behandelten wir ähnliche Matrizen, zur Erinnerung:

Für $A, B \in \text{Mat}(n \times n, K)$ heißt A genau dann **ähnlich** zu B , wenn es ein $T \in \text{GL}(n, K)$ gibt, so dass

$$B = T \cdot A \cdot T^{-1}$$

gilt. Setzt man B als Diagonalmatrix an, welche genau die Eigenwerte von A als Diagonalelemente besitzt, so offenbart sich ein interessanter Zusammenhang zwischen der Ähnlichkeit von Matrizen und dem Konzept von Eigenwerten und Eigenvektoren.

Satz 2.10.21

Sei $A \in \text{Mat}(n \times n, K)$. Dann sind äquivalent:

1. Es gibt eine Basis von K^n bestehend aus Eigenvektoren.
2. Es gibt ein $T \in \text{GL}(n, K)$, so dass $T \cdot A \cdot T^{-1} = \Delta(\lambda_1, \dots, \lambda_n)$ gilt, d.h. A ist ähnlich zu einer Diagonalmatrix.

Beweis.

„(1.) \Leftarrow (2.)“:

Sei $T \in \text{GL}(n, K)$, so dass $TAT^{-1} = \Delta(\lambda_1, \dots, \lambda_n)$ gilt. Durch Umstellen erhalten wir

$$A = T^{-1} \cdot \Delta(\lambda_1, \dots, \lambda_n) \cdot T.$$

Für $k \in \{1, \dots, n\}$ setzen wir nun $v_k := T^{-1}e_k$. Da die Standard-Basis-Vektoren e_1, \dots, e_n linear unabhängig sind und T^{-1} invertierbar ist, sind auch v_1, \dots, v_n linear unabhängig. Folglich bilden v_1, \dots, v_n eine Basis von K^n und es gilt

$$\begin{aligned} Av_k &= T^{-1} \cdot \Delta(\lambda_1, \dots, \lambda_n) \cdot T \cdot T^{-1} \cdot e_k \\ &= T^{-1} \cdot \Delta(\lambda_1, \dots, \lambda_n) \cdot e_k \\ &= T^{-1} \cdot \lambda_k \cdot e_k \\ &= \lambda_k \cdot T^{-1} \cdot e_k \\ &= \lambda_k \cdot v_k, \end{aligned}$$

womit v_k wie verlangt für $k \in \{1, \dots, n\}$ Eigenvektor ist.

„(1.) \Rightarrow (2.)“:

Sei nun $\{v_1, \dots, v_n\}$ eine Basis von K^n aus Eigenvektoren, d.h. gelte $A \cdot v_k = \lambda_k \cdot v_k$ für alle $k \in \{1, \dots, n\}$. Damit ist sicherlich $(v_1, \dots, v_n) \in \text{GL}(n, K)$, womit wir $T := (v_1, \dots, v_n)^{-1} \in \text{GL}(n, K)$ setzen können. Es bleibt zu zeigen, dass TAT^{-1} eine Diagonalmatrix ist. Hierfür prüfen wir nach, dass durch e_k Eigenvektoren gegeben sind. Es gilt

$$\begin{aligned} TAT^{-1} \cdot e_k &= T \cdot Av_k \\ &= T \cdot \lambda_k v_k \\ &= \lambda_k T \cdot v_k \\ &= \lambda_k T \cdot T^{-1} e_k \\ &= \lambda_k \cdot e_k, \end{aligned}$$

also

$$TAT^{-1} = \Delta(\lambda_1, \dots, \lambda_n).$$

□

Definition 2.10.22. Sind (1.) bzw. (2.) erfüllt, so heißt A **diagonalisierbar**.

Bemerkung 2.10.23. Der Beweis von Satz 2.10.21 zeigt folgendes: Kennen wir für eine diagonalisierbare Matrix A die Matrix $T \in \text{GL}(n, K)$ und die Diagonalmatrix $\Delta(\lambda_1, \dots, \lambda_n)$, so dass $TAT^{-1} = \Delta(\lambda_1, \dots, \lambda_n)$ gilt, so kennen wir die Eigenwerte von A , nämlich $\lambda_1, \dots, \lambda_n$ und die Eigenvektoren von A , nämlich die Spalten von T^{-1} .

2 Matrixrechnung

Beispiel 2.10.24. Für $K = \mathbb{R}$ und $A = S_\theta$ wurde auf Übungsblatt 13 nachgerechnet, dass $R_\theta S_0 R_{-\theta} = S_\theta$ für alle $\theta \in \mathbb{R}$ gilt. Umstellen nach der Spiegelung S_0 an der x -Achse liefert dann

$$\Delta(1, -1) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = S_0 = R_{-\theta} S_\theta R_{-\theta}^{-1},$$

d.h. S_θ ist diagonalisierbar mit $T = R_{-\theta}$.

Beispiel 2.10.25. Für reelle Drehmatrizen wissen wir, dass diese keine reellen Eigenwerte bzw. Eigenvektoren besitzen. Somit ist R_θ auch nicht (reell) diagonalisierbar.

Beispiel 2.10.26. Für $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in \text{Mat}(2 \times 2, \mathbb{R})$ ermitteln wir die Eigenvektoren, d.h. zunächst die Eigenwerte als Nullstellen des charakteristischen Polynoms:

$$\begin{aligned} \chi_A(\lambda) &= \det(A - \lambda \cdot \mathbf{1}_n) \\ &= \det\left(\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} - \lambda \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \\ &= \det\begin{pmatrix} 1-\lambda & 1 \\ 1 & 0-\lambda \end{pmatrix} \\ &= (1-\lambda) \cdot (-\lambda) - 1 \\ &= \lambda^2 - \lambda - 1. \end{aligned}$$

Die Eigenwerte von A sind also gegeben durch

$$\lambda_{1,2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 1},$$

d.h.

$$\lambda_1 = \frac{1 + \sqrt{5}}{2} \quad \text{und} \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}.$$

Da A eine 2×2 -Matrix ist und genau 2 verschiedene Eigenwerte hat, ist A diagonalisierbar. Offenbar sind $\text{Eig}(A, \lambda_1)$ und $\text{Eig}(A, \lambda_2)$ weder 0-dimensional noch 2-dimensional (da $A \neq \mathbf{1}_n$). Folglich kommt hier nur die geometrische Vielfachheit $\dim(\text{Eig}(A, \lambda_1)) = \dim(\text{Eig}(A, \lambda_2)) = 1$ in Frage. Ist nun $x \in \text{Eig}(A, \lambda_1)$, so gilt

$$\begin{aligned} \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= (A - \lambda_1 \cdot \mathbf{1}_2) \cdot x \\ &= \begin{pmatrix} 1 - \lambda_1 & 1 \\ 1 & -\lambda_1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} (1 - \lambda_1)x_1 + x_2 \\ x_1 - \lambda_1 x_2 \end{pmatrix}, \end{aligned}$$

⁷Die Zahl $\frac{1+\sqrt{5}}{2} \approx 1,618$ heißt **goldener Schnitt**.

womit wir insbesondere $x_1 - \lambda_1 x_2 = 0$ erhalten, d.h. $x_1 = \lambda_1 x_2$. Also ist x von der Gestalt

$$x = \begin{pmatrix} \lambda_1 x_2 \\ x_2 \end{pmatrix} = x_2 \cdot \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}$$

und damit muss der zu λ_1 gehörige Eigenraum

$$\text{Eig}(A, \lambda_1) \subset \left\{ x_2 \cdot \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} \mid x_2 \in \mathbb{R} \right\}$$

erfüllen. Wegen $\dim(\text{Eig}(A, \lambda_1)) = 1 = \dim \left\{ t \cdot \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} \mid t \in \mathbb{R} \right\}$ folgt

$$\text{Eig}(A, \lambda_1) = \left\{ t \cdot \begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix} \mid t \in \mathbb{R} \right\}.$$

Analog bekommen wir

$$\text{Eig}(A, \lambda_2) = \left\{ t \cdot \begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix} \mid t \in \mathbb{R} \right\}.$$

Damit bilden $\begin{pmatrix} \lambda_1 \\ 1 \end{pmatrix}$ und $\begin{pmatrix} \lambda_2 \\ 1 \end{pmatrix}$ eine Basis von \mathbb{R}^2 aus Eigenvektoren. Wir setzen

$$T^{-1} = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}.$$

und ermitteln

$$\begin{aligned} T &= \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix} \\ &= \frac{1}{\frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2}} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix}. \end{aligned}$$

Bemerkung 2.10.27 (Rezept zur Diagonalisierung). Ist $A \in \text{Mat}(n \times n, K)$ diagonalisierbar, so findet man die Diagonalmatrix sowie T und T^{-1} wie folgt:

1. Berechne das charakteristische Polynom χ_A von A .
2. Berechne die Eigenwerte $\lambda_1, \dots, \lambda_n$ von A (als Nullstellen von χ_A). Damit haben wir bereits die Diagonalmatrix $\Delta(\lambda_1, \dots, \lambda_n)$.
3. Berechne für jeden Eigenwert den Eigenraum $\text{Eig}(A, \lambda_k) = \text{Lös}(A - \lambda_k \cdot \mathbf{1}_n, \mathbf{0})$ (z.B. mittels Gauß-Algorithmus).
4. Wähle je eine Basis aus $\text{Eig}(A, \lambda_k)$ und setze diese zu einer Basis v_1, \dots, v_n von K^n zusammen. Dann ist $T^{-1} = (v_1, \dots, v_n)$.

2 Matrixrechnung

5. Invertiere T^{-1} und erhalte so T .

Bemerkung 2.10.28 (Potenzieren diagonalisierbarer Matrizen). Welche Dienste kann uns die Diagonalisierung einer Matrix A erweisen? Man kann damit z.B. effizient hohe Matrixpotenzen A^k berechnen. Dazu schreiben wir einmal

$$A = T^{-1} \cdot \Delta \cdot T,$$

wobei $\Delta = \Delta(\lambda_1, \dots, \lambda_n)$ ist, d.h. die Eigenwerte λ_k von A stehen auf der Diagonalen von Δ . Dann ist

$$\begin{aligned} A^k &= \underbrace{(T^{-1}\Delta T) \cdot (T^{-1}\Delta T) \cdot \dots \cdot (T^{-1}\Delta T)}_{k \text{ Faktoren}} \cdot (T^{-1}\Delta T) \\ &= T^{-1} \underbrace{\Delta T \cdot T^{-1} \Delta T}_{=\mathbb{1}_n} \cdot \dots \cdot T^{-1} \underbrace{\Delta T \cdot T^{-1} \Delta T}_{=\mathbb{1}_n} \\ &= T^{-1} \cdot \Delta^k \cdot T \\ &= T^{-1} \cdot (\Delta(\lambda_1, \dots, \lambda_n))^k \cdot T \\ &= T^{-1} \cdot \Delta(\lambda_1^k, \dots, \lambda_n^k) \cdot T. \end{aligned}$$

Beispiel 2.10.29. Im Beispiel (2.10.26) fanden wir für $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = T^{-1} \cdot \Delta \cdot T$ die Matrizen

$$T = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix}, \quad \Delta = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Demnach gilt

$$\begin{aligned} A^k &= T^{-1} \cdot \Delta^k \cdot T \\ &= \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^k \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{pmatrix} \begin{pmatrix} 1 & -\lambda_2 \\ -1 & \lambda_1 \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1^k & -\lambda_1^k \lambda_2 \\ -\lambda_2^k & \lambda_1 \lambda_2^k \end{pmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1^{k+1} - \lambda_2^{k+1} & -\lambda_1^{k+1} \lambda_2 + \lambda_1 \lambda_2^{k+1} \\ \lambda_1^k - \lambda_2^k & \lambda_1 \lambda_2^k - \lambda_1^k \lambda_2 \end{pmatrix}. \end{aligned}$$

Beispiel 2.10.30. Sind alle quadratischen Matrizen diagonalisierbar? Nein! Denn z.B. für

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \text{Mat}(2 \times 2, \mathbb{R})$$

bekommen wir

$$\begin{aligned}\chi_A(\lambda) &= \det(A - \lambda \cdot \mathbb{1}_2) \\ &= \det \begin{pmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{pmatrix} \\ &= (1 - \lambda)^2,\end{aligned}$$

also den Eigenwert $\lambda_1 = 1$ mit Vielfachheit 2. Wäre nun A diagonalisierbar, so müsste A ähnlich zu $\Delta(1, 1) = \mathbb{1}_2$ sein. Dies ist aber nicht möglich, da $\mathbb{1}_2$ nur zu sich selbst ähnlich ist, denn für jedes $T \in \text{GL}(2, \mathbb{R})$ ist $T\mathbb{1}_2T^{-1} = T\mathbb{1}_2T^{-1} = \mathbb{1}_2$. Also ist A nicht diagonalisierbar.

Beispiel 2.10.31. Fast jeder kennt das berühmte Kaninchen-Populationsmodell von Leonardo da Pisa, genannt Fibonacci, dem wohl berühmtesten Mathematiker des Mittelalters. In seinem Modell gelten folgende Regeln:

- Es gibt genau ein Paar geschlechtsreifer Kaninchen im ersten Monat.
- Jedes geschlechtsreife Paar wirft pro Monat ein weiteres Paar.
- Jedes neugeborene Paar wird zwei Monate später geschlechtsreif.
- Es gibt keine Todesfälle.

Die Mathematisierung dieser Regeln liefert die Fibonacci-Zahlen f_k , d.h. die Anzahl der geschlechtsreifen Paare im k -ten Monat, z.B. $f_1 = 1$, $f_2 = 1$, $f_3 = 2$, $f_4 = 3$, \dots , $f_{24} = 46.368$, \dots , $f_{60} = 1.548.008.755.920$, \dots

Das rekursive Bildungsgesetz lautet

$$f_{k+2} = f_{k+1} + f_k,$$

denn im $(k+2)$ -ten Monat kommen zu den bereits im Vormonat vorhandenen geschlechtsreifen Paaren (f_{k+1} viele) noch die neu geschlechtsreif gewordenen hinzu. Dies sind gerade so viele, wie zwei Monate vorher geschlechtsreif waren (f_k viele). Die so definierte Folge heißt dann **Fibonacci-Folge**. Um nun jedoch etwa die 24-te Fibonacci-Zahl f_{24} zu erhalten, rechnen wir eine ganze Weile.

Wir leiten jetzt eine geschlossene Formel für die rekursiv definierten Fibonacci-Zahlen her. Wie dies mittels der Matrixpotenz aus Beispiel (2.10.29) gelingt, sehen wir jetzt.

Sei $(f_k)_k$ die Fibonacci-Folge mit $f_0 = 0$ und $f_1 = f_2 = 1$ sowie $f_{k+2} = f_{k+1} + f_k$.



Abb. 66

2 Matrixrechnung

Wir setzen den Vektor $w_k = \begin{pmatrix} f_{k+1} \\ f_k \end{pmatrix} \in \mathbb{R}^2$ an und finden durch Einsetzen

$$w_0 = \begin{pmatrix} f_1 \\ f_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = e_1$$

sowie

$$\begin{aligned} w_{k+1} &= \begin{pmatrix} f_{k+2} \\ f_{k+1} \end{pmatrix} = \begin{pmatrix} f_k + f_{k+1} \\ f_{k+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f_{k+1} \\ f_k \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} w_k \\ &\stackrel{(2.10.29)}{=} Aw_k. \end{aligned}$$

Damit folgt

$$\begin{aligned} \begin{pmatrix} f_{k+1} \\ f_k \end{pmatrix} &= w_k \\ &= A \cdot w_{k-1} = A \cdot A \cdot w_{k-2} = \dots = A^k \cdot w_0 \\ &= A^k \cdot e_1 \\ &\stackrel{(2.10.29)}{=} \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1^{k+1} - \lambda_2^{k+1} & -\lambda_1^{k+1}\lambda_2 + \lambda_1\lambda_2^{k+1} \\ \lambda_1^k - \lambda_2^k & \lambda_1\lambda_2^k - \lambda_1^k\lambda_2 \end{pmatrix} \cdot e_1 \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \lambda_1^{k+1} - \lambda_2^{k+1} \\ \lambda_1^k - \lambda_2^k \end{pmatrix}. \end{aligned}$$

Die zweite Komponente liefert nun

$$f_k = \frac{\lambda_1^k - \lambda_2^k}{\sqrt{5}},$$

also

$$\boxed{f_k = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^k - \left(\frac{1-\sqrt{5}}{2} \right)^k \right)},$$

die Formel von Moivre-Binet. Erstaunlich ist die Tatsache, dass f_k für jedes $k \in \mathbb{N}$ ganzzahlig ist, obwohl λ_1 und λ_2 irrational sind. Wir wenden uns nachfolgend λ_1 noch einmal näher zu.

Der goldene Schnitt.



Abb. 67

Wir suchen Streckenlängen a und b derart, dass

$$\begin{aligned}\frac{a}{b} &= \frac{a+b}{a} \\ \Leftrightarrow \frac{a}{b} &= 1 + \frac{b}{a}\end{aligned}$$

gilt. Wir setzen $\lambda := \frac{a}{b}$ und erhalten $\lambda = 1 + \frac{1}{\lambda}$, also $\lambda^2 - \lambda - 1 = 0$. Die linke Seite ist genau χ_A aus Beispiel (2.10.26). Eine Lösung ist $\lambda = \lambda_1 = \frac{1+\sqrt{5}}{2} \approx 1.618$, die sogenannte **goldene Zahl**. Sie gilt als besonders ästhetisches Verhältnis, z.B. in der Kunst aber auch in der Anatomie.

So kann die goldene Zahl im Sinne des Verhältnisses $\frac{a}{b}$ auch am menschlichen Körper, d.h. bezüglich der Längenverhältnisse Ellebogen-Handwurzel zu Handwurzel-Mittelfingerspitze bzw. Ferse-Bauchnabel zu Bauchnabel-Haaransatz entdeckt werden.

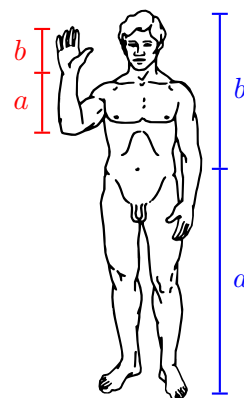


Abb. 68

Wir wissen bereits: Für Matrizen $A \in \text{Mat}(n \times n, K)$ ist $\lambda \in K$ genau dann Eigenwert von A , wenn $\chi_A(\lambda) = 0$ gilt. Die Dimension des zu λ gehörigen Eigenraums $\text{Eig}(A, \lambda)$ trägt den Namen geometrische Vielfachheit und wird manchmal auch mit $\mu_{\text{geo}}(\lambda)$ abgekürzt. Eine analoge Konvention gibt es für die Häufigkeit der Nullstellen λ von χ_A .

Definition 2.10.32. Die Nullstellenvielfachheit von λ als Nullstelle von χ_A heißt **algebraische Vielfachheit** bzw. **algebraische Multiplizität** von λ . Abkürzend schreiben wir für sie $\mu_{\text{alg}}(\lambda)$.

Die Vielfachheiten $\mu_{\text{geo}}(\lambda)$ und $\mu_{\text{alg}}(\lambda)$ brauchen nicht übereinzustimmen. Zum Beispiel wis-

2 Matrixrechnung

sen wir aus Beispiel (2.10.30), dass $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ nicht diagonalisierbar ist und als Eigenwert $\lambda = 1$ mit $\mu_{\text{alg}}(\lambda) = 2$ besitzt. Ferner ist $\mu_{\text{geo}}(\lambda) = 1$, wie wir ausführlich diskutierten. Eine allgemeine Feststellung hierzu (die wir an dieser Stelle nicht beweisen) ist Gegenstand von

Bemerkung 2.10.33. Für jeden Eigenwert λ einer $n \times n$ -Matrix A gilt

$$1 \leq \mu_{\text{geo}}(\lambda) \leq \mu_{\text{alg}}(\lambda).$$

Sind ferner $\lambda_1, \dots, \lambda_k$ die paarweise verschiedenen Eigenwerte von A , dann ist

$$\mu_{\text{alg}}(\lambda_1) + \dots + \mu_{\text{alg}}(\lambda_k) \leq n.$$

Ist $K = \mathbb{C}$, so gilt wegen dem Fundamentalsatz der Algebra sogar

$$\mu_{\text{alg}}(\lambda_1) + \dots + \mu_{\text{alg}}(\lambda_k) = n.$$

Gilt schließlich $\mu_{\text{alg}}(\lambda_j) = \mu_{\text{geo}}(\lambda_j)$ für alle $j \in \{1, \dots, k\}$, so ist A diagonalisierbar.

2.11 Orientierungen und Vektorprodukt

Seien $v_1, \dots, v_n \in \mathbb{R}^n$. Zuweilen ist es sinnvoll auf die Reihenfolge der Basisvektoren zu bestehen. Wir schreiben dann (v_1, \dots, v_n) statt $\{v_1, \dots, v_n\}$. Bekanntlich ist (v_1, \dots, v_n) genau dann eine Basis von \mathbb{R}^n , wenn v_1, \dots, v_n linear unabhängig sind, d.h. $\det(v_1, \dots, v_n) \neq 0$ gilt. Qualitativ unterscheiden sich hierbei genau zwei Klassen:

Definition 2.11.1. Eine Basis (v_1, \dots, v_n) von \mathbb{R}^n heißt genau dann **positiv orientiert**, wenn $\det(v_1, \dots, v_n) > 0$ und **negativ orientiert**, wenn $\det(v_1, \dots, v_n) < 0$ gilt.

Beispiel 2.11.2. Die Standard-Basis (e_1, \dots, e_n) ist positiv orientiert, denn es gilt

$$\det(e_1, \dots, e_n) = \det(\mathbf{1}_n) = 1 > 0.$$

Veranschaulichung. Verschaffen wir uns eine anschauliche Vorstellung von der Bedeutung des Konzepts der Orientierung im Falle kleiner Dimensionen n .

Für $n = 1$, also in \mathbb{R} , entspricht die Orientierung der Durchlauf-Richtung.

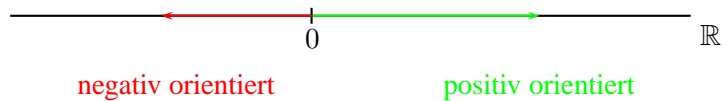


Abb. 69

Für $n = 2$, d.h. im \mathbb{R}^2 , betrachten wir hilfsweise die Drehung

$$J := R_{\frac{\pi}{2}} = \begin{pmatrix} \cos\left(\frac{\pi}{2}\right) & -\sin\left(\frac{\pi}{2}\right) \\ \sin\left(\frac{\pi}{2}\right) & \cos\left(\frac{\pi}{2}\right) \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

denn damit gilt für $v, w \in \mathbb{R}^2$:

$$\begin{aligned} \langle Jv, w \rangle &= \left\langle \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right\rangle \\ &= \left\langle \begin{pmatrix} -v_2 \\ v_1 \end{pmatrix}, \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \right\rangle \\ &= -v_2 w_1 + v_1 w_2 \\ &= \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \\ &= \det(v, w). \end{aligned}$$

Die geometrische Interpretation kann nun direkt aus der des Skalarprodukts abgeleitet werden:

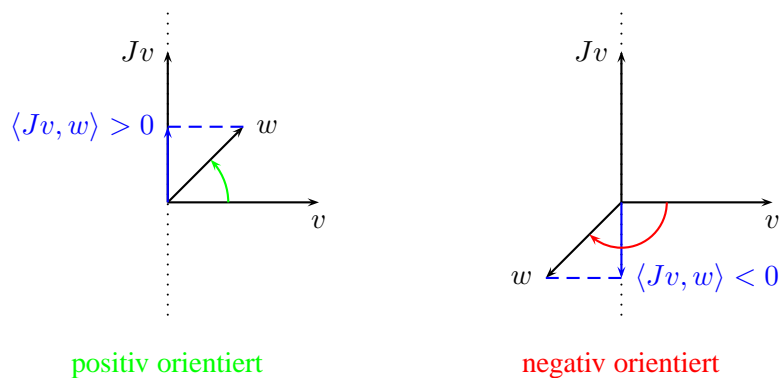


Abb. 70

Die Orientierung in Dimension 2 entspricht also dem Drehsinn.

Bevor wir uns noch dem Fall $n = 3$ widmen, sind hier einige wichtige Eigenschaften der Orientierung:

2 Matrixrechnung

Bemerkung 2.11.3. Die allgemeinen Eigenschaften der Determinante liefern folgende Aussagen:

1. Vertauschung zweier Basisvektoren kehrt die Orientierung um.
2. Ersetzung eines Basisvektors durch sein Negatives kehrt die Orientierung um.
3. Streckung eines Basisvektors um einen positiven Faktor erhält die Orientierung.

Bemerkung 2.11.4. Orientierungen sind sehr robust, d.h. bei kleinen Änderungen der Basis bleibt die Orientierung erhalten. Genauer: Sind $v_1, \dots, v_n : I \rightarrow \mathbb{R}$ stetige Abbildungen und $I \subset \mathbb{R}$ ein (Zeit-)Intervall, so dass $(v_1(t), \dots, v_n(t))$ für jedes $t \in I$ eine Basis ist, dann hat $(v_1(t), \dots, v_n(t))$ für alle $t \in I$ dieselbe Orientierung.

Beweis.

Die Funktion $I \rightarrow \mathbb{R}$ mit $t \mapsto \det(v_1(t), \dots, v_n(t))$ ist stetig und verschwindet auf ganz I nirgends, d.h. ist immer positiv oder immer negativ (vgl. Zwischenwertsatz aus der Analysis). \square

Für $n = 3$, d.h. im \mathbb{R}^3 , entspricht die Orientierung der Händigkeit.

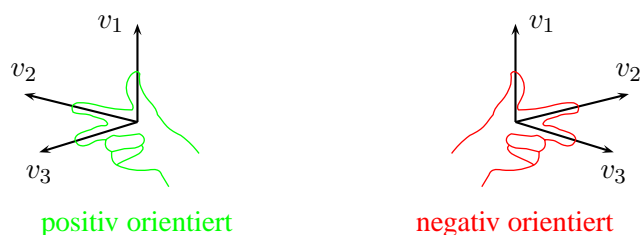


Abb. 71

Das Vektorprodukt. Im Allgemeinen ist für Vektoren des \mathbb{R}^n keine Multiplikation definiert, deren Ergebnis wieder ein Vektor wäre. Nur für Dimension $n = 3$ lässt sie sich wie folgt sinnvoll definieren:

Definition 2.11.5. Sind $v, w \in \mathbb{R}^3$, so heißt

$$v \times w := \begin{pmatrix} v_2 w_3 - v_3 w_2 \\ v_3 w_1 - v_1 w_3 \\ v_1 w_2 - v_2 w_1 \end{pmatrix}$$

das **Vektorprodukt** bzw. **Kreuzprodukt** von v und w .

Wie kann man sich diese scheinbar komplizierte Definition gut merken? Wir schreiben

$\begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \\ v_3 & w_3 \end{pmatrix}$, streichen die jeweils betrachtete Zeile und wenden das Schachbrettmuster an:

$$v \times w = \begin{pmatrix} + \det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} \\ - \det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} \\ + \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \end{pmatrix}.$$

Bevor wir nun wesentliche Eigenschaften des Vektorprodukts untersuchen, hier noch eine allgemeine

Bemerkung 2.11.6. Sind $x, y \in \mathbb{R}^n$, dann gilt

$$x = y \Leftrightarrow \forall z \in \mathbb{R}^n : \langle x, z \rangle = \langle y, z \rangle.$$

Beweis.

„ \Rightarrow “: Nichts zu zeigen.

„ \Leftarrow “: Speziell für $z = e_k$ gilt $x_k = \langle x, e_k \rangle = \langle y, e_k \rangle = y_k$ für alle k . Es folgt $x = y$. \square

Satz 2.11.7

Für alle $v, \tilde{v}, w, \tilde{w}, z \in \mathbb{R}^3$ und alle $\lambda \in \mathbb{R}$ gilt:

1. $\langle v \times w, z \rangle = \det(v, w, z)$
2. $v \times w = -(w \times v)$
3. $(v + \tilde{v}) \times w = (v \times w) + (\tilde{v} \times w)$
 $v \times (w + \tilde{w}) = (v \times w) + (v \times \tilde{w})$
 $(\lambda v) \times w = \lambda \cdot (v \times w) = v \times (\lambda w)$
4. $v \times w = \mathbf{0} \Leftrightarrow v, w$ sind linear abhängig
5. $v \times w \perp v, w$
6. $\|v \times w\| = \text{vol}_2(P)$, wobei P das von v und w aufgespannte Parallelogramm ist
7. Sind v und w linear unabhängig, so bildet $(v, w, v \times w)$ eine positiv orientierte Basis von \mathbb{R}^3 .

2 Matrixrechnung

Beweis.

Zu 1.)

Es gilt

$$\begin{aligned}\langle v \times w, z \rangle &= + \det \begin{pmatrix} v_2 & w_2 \\ v_3 & w_3 \end{pmatrix} \cdot z_1 - \det \begin{pmatrix} v_1 & w_1 \\ v_3 & w_3 \end{pmatrix} \cdot z_2 + \det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \cdot z_3 \\ &\stackrel{\text{Entw. nach}}{=} \det \begin{pmatrix} v_1 & w_1 & z_1 \\ v_2 & w_2 & z_2 \\ v_3 & w_3 & z_3 \end{pmatrix}. \\ &\quad \text{3. Spalte}\end{aligned}$$

Zu 2.)

Für alle $z \in \mathbb{R}^3$ gilt:

$$\begin{aligned}\langle v \times w, z \rangle &\stackrel{(1)}{=} \det(v, w, z) \\ &= -\det(w, v, z) \\ &\stackrel{(1)}{=} -\langle w \times v, z \rangle \\ &= \langle -(w \times v), z \rangle.\end{aligned}$$

Nach Bemerkung (2.11.6) folgt $v \times w = -(w \times v)$.

Zu 3.)

Für alle $z \in \mathbb{R}^3$ gilt:

$$\begin{aligned}\langle (v + \tilde{v}) \times w, z \rangle &\stackrel{(1)}{=} \det(v + \tilde{v}, w, z) \\ &= \det(v, w, z) + \det(\tilde{v}, w, z) \\ &\stackrel{(1)}{=} \langle v \times w, z \rangle + \langle \tilde{v} \times w, z \rangle \\ &= \langle (v \times w) + (\tilde{v} \times w), z \rangle.\end{aligned}$$

Nach (2.11.6) ist somit $(v + \tilde{v}) \times w = (v \times w) + (\tilde{v} \times w)$.

Die anderen beiden Regeln folgen analog.

Zu 4.)

Es gilt:

$$\begin{aligned}v, w \text{ sind linear abhängig} &\Leftrightarrow \text{rg}(v, w) \leq 1 \\ &\Leftrightarrow \text{rg}(v, w, z) \leq 2 \text{ für alle } z \in \mathbb{R}^3 \\ &\Leftrightarrow \det(v, w, z) = 0 \text{ für alle } z \in \mathbb{R}^3 \\ &\Leftrightarrow \langle v \times w, z \rangle = 0 \text{ für alle } z \in \mathbb{R}^3 \\ &\Leftrightarrow v \times w = \mathbf{0}.\end{aligned}$$

Zu 5.)

Wir berechnen

$$\langle v \times w, v \rangle \stackrel{(1)}{=} \det(v, w, v) = \mathbf{0}.$$

Hieraus folgt $v \times w \perp v$. Für $v \times w \perp w$ argumentiert man analog.

Zu 6.)

1. Fall:

Sind v, w linear abhängig, so folgt nach (4.) sofort $v \times w = \mathbf{0}$, also $\|v \times w\| = 0$.

Ferner ist P dann auch in einer Geraden des \mathbb{R}^2 enthalten, also folgt $\text{vol}_2(P) = 0$.

2. Fall:

Sind v, w linear unabhängig, so gilt

$$\begin{aligned} \|v \times w\|^2 &= \langle v \times w, v \times w \rangle \\ &\stackrel{(1.)}{=} \det(v, w, v \times w) \\ &= \text{vol}_3(\mathbf{Q}), \end{aligned} \tag{2.38}$$

wobei \mathbf{Q} das Parallelepiped ist, welches von $v, w, v \times w$ aufgespannt wird. Da nach (5.) $v \times w \perp v, w$ gilt, haben wir

$$\text{vol}_3(\mathbf{Q}) = \text{vol}_2(P) \cdot \|v \times w\|. \tag{2.39}$$

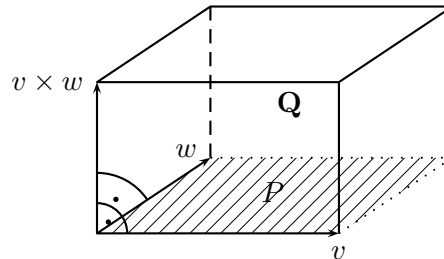


Abb. 72

Einsetzen von (2.39) in (2.38) liefert schließlich

$$\|v \times w\|^2 = \text{vol}_2(P) \cdot \underbrace{\|v \times w\|}_{\substack{\neq 0 \\ (4.)}},$$

also $\|v \times w\| = \text{vol}_2(P)$.

Zu 7.)

Zunächst einmal überlegen wir, dass $v, w, v \times w$ überhaupt linear unabhängig sind. Wir betrachten hierzu

$$\alpha v + \beta w + \gamma(v \times w) = \mathbf{0}. \tag{2.40}$$

und zeigen $\alpha = \beta = \gamma = 0$. Es gilt

$$\begin{aligned} 0 &= \langle \mathbf{0}, v \times w \rangle \stackrel{(2.40)}{=} \langle \alpha v + \beta w + \gamma(v \times w), v \times w \rangle \\ &= \alpha \cdot \langle v, v \times w \rangle + \beta \cdot \langle w, v \times w \rangle + \gamma \cdot \langle v \times w, v \times w \rangle \\ &\stackrel{(5.)}{=} \alpha \cdot 0 + \beta \cdot 0 + \gamma \cdot \langle v \times w, v \times w \rangle \\ &= \gamma \cdot \underbrace{\|v \times w\|^2}_{\substack{\neq 0 \\ (4.)}} \end{aligned}$$

und folglich ist $\gamma = 0$. Setzen wir dies in (2.40) ein, so kommt

$$\alpha v + \beta w = \mathbf{0}$$

2 Matrixrechnung

heraus und da v, w nach Voraussetzung linear unabhängig sind, folgt auch $\alpha = \beta = 0$. Also sind $v, w, v \times w$ linear unabhängig, d.h. $(v, w, v \times w)$ ist eine Basis von \mathbb{R}^3 . Die positive Orientiertheit von $(v, w, v \times w)$ folgt aus

$$\det(v, w, v \times w) \stackrel{(1.)}{=} \langle v \times w, v \times w \rangle = \|v \times w\|^2 > 0.$$

□

Fassen wir zusammen: $v \times w$ ist der eindeutige Vektor im \mathbb{R}^3 , der auf der Ebene, die von v und w aufgespannt wird, senkrecht steht, dessen Länge gleich dem Flächeninhalt des Parallelogramms ist, welches von v und w aufgespannt wird, so dass $(v, w, v \times w)$ rechtshändig ist (d.h. eine positive Basis des \mathbb{R}^3 bildet).

3 Vektorräume

Viele Menschen sind unglücklich, weil sie nicht abstrahieren können.

(Immanuel Kant,
Anthropologie in pragmatischer
Hinsicht)

Wir kennen uns bereits ziemlich gut im \mathbb{R}^n bzw. \mathbb{C}^n sowie mit einigen konkreten Untervektorräumen aus, so etwa für $n = 3$ mit den Ursprungsebenen im \mathbb{R}^3 . In diesem letzten Kapitel wenden wir uns einigen strukturmathematischen, eher abstrakten Sichtweisen zu.

3.1 Gruppen

Den Anfang soll das Konzept der Gruppe bilden. Gruppen formalisieren, kurz gesagt, Rechenoperationen (auf vorgegebenen Mengen). Jede Gruppe genügt bestimmten Minimalanforderungen, sogenannten Axiomen, welche von der Rechenvorschrift erfüllt werden muss, damit diese vernünftig auf ihrer Trägermenge arbeiten kann.

Definition 3.1.1. Eine **Gruppe** ist ein Paar (G, \circ) , wobei G eine Menge ist, auf der eine Verknüpfung

$$\begin{aligned} \circ : G \times G &\rightarrow G \\ (g_1, g_2) &\mapsto g_1 \circ g_2, \end{aligned}$$

so definiert ist, dass gilt:

1. $\forall g_1, g_2, g_3 \in G : g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$ (Assoziativität)
2. $\exists e \in G \forall g \in G : e \circ g = g \circ e = g$ (neutrales Element)
3. $\forall g \in G \exists g^{-1} \in G : g \circ g^{-1} = g^{-1} \circ g = e$ (inverses Element)

3 Vektorräume

Wir haben bereits wie selbstverständlich mit so einigen Gruppen und ihren Eigenschaften gearbeitet. Der Fundus an Kandidaten ist daher groß. Sehen wir uns einige Beispiel an und halten fest, bei welchen Paaren es sich um Gruppen handelt und bei welchen nicht.

Beispiel 3.1.2. Bei den folgenden Strukturen handelt es sich um Gruppen:

Gruppe	neutrales Element	inverses Element
$(\mathbb{Z}, +)$	$e = 0$	$g^{-1} = -g$
$(\mathbb{Q}, +)$	$e = 0$	$g^{-1} = -g$
$(\mathbb{Q} \setminus \{0\}, \cdot)$	$e = 1$	$g^{-1} = \frac{1}{g}$
(\mathbb{R}^+, \cdot)	$e = 1$	$g^{-1} = \frac{1}{g}$
$(\mathbb{C}, +)$	$e = 0$	$g^{-1} = -g$
$(\mathbb{C} \setminus \{0\}, \cdot)$	$e = 1$	$g^{-1} = -g$
$(\mathbb{R}, +)$	$e = 0$	$g^{-1} = -g$
$(\mathbb{R} \setminus \{0\}, \cdot)$	$e = 1$	$g^{-1} = -g$
$(GL(n, \mathbb{R}), \cdot)$	$e = \mathbf{1}_n$	$g^{-1} = \text{inverse Matrix}$
$(SO(2), \cdot) := (\{R_\theta \mid \theta \in \mathbb{R}\}, \cdot)$	$e = R_0 = \mathbf{1}_n$	$R_\theta^{-1} = R_{-\theta}$
$(O(2), \cdot) := (SO(2) \cup O(2)^-, \cdot)$	$e = R_0 = \mathbf{1}_n$	$R_\theta^{-1} = R_{-\theta}$ bzw. $S_\theta^{-1} = S_\theta$

Beispiel 3.1.3. Keine Gruppen sind dagegen:

Struktur	verletzt	Begründung
$(\mathbb{N}_0, +)$	Axiom 3	zu $g \neq 0$ existiert kein inverses Element
(\mathbb{Q}, \cdot)	Axiom 3	zu $g = 0$ existiert kein inverses Element
(\mathbb{R}^-, \cdot)	Definition	$\cdot : \mathbb{R}^- \times \mathbb{R}^- \rightarrow \mathbb{R}^+$ statt \mathbb{R}^-
$(Mat(n \times n, \mathbb{R}), \cdot)$	Axiom 3	nicht alle Matrizen sind invertierbar
$(O(2)^-, \cdot) := (\{S_\theta \mid \theta \in \mathbb{R}\}, \cdot)$	Definition	$\cdot : O(2)^- \times O(2)^- \rightarrow SO(2)$ statt $O(2)^-$

Eine Sonderform von Gruppen, bei der es nicht auf die Reihenfolge der Argumente ankommt, ist die folgende.

Definition 3.1.4. Eine Gruppe (G, \circ) heißt genau dann **abelsch** bzw. **kommutativ**, wenn gilt:

$$4. \forall g_1, g_2 \in G : g_1 \circ g_2 = g_2 \circ g_1 \quad (\text{Kommutativität})$$

Beispiel 3.1.5. Bis auf $GL(n, \mathbb{R})$ für $n \geq 2$ und $O(2)$ sind alle Gruppen aus Beispiel (3.1.2) abelsch.

Wir führen zwei Mini-Beweise:

Bemerkung 3.1.6. In jeder Gruppe G gibt es *genau ein* neutrales Element $e \in G$, denn sind $e, e' \in G$ zwei neutrale Elemente, so gilt

$$e \stackrel{e' \text{ neutral}}{=} e \circ e' \stackrel{e \text{ neutral}}{=} e'.$$

Bemerkung 3.1.7. In jeder Gruppe G gibt es zu jedem Element $g \in G$ *genau ein* inverses Element, denn sind h, h' invers zu G , so gilt

$$h' = h' \circ e = h' \circ (g \circ h) \stackrel{(1.)}{=} (h' \circ g) \circ h = e \circ h = h.$$

Wir haben nun das Konzept der Gruppen durchschaut. Dieser Grundbegriff allein ist natürlich noch nicht besonders erhehend. Der Nutzen wird erst richtig erkennbar, wenn sich Funktionen zwischen Gruppen finden lassen, welche die gutartige Struktur dieser Gruppen erhalten, d.h. wenn bestimmte Abbildungen die Rechenregeln respektieren.

Definition 3.1.8. Seien (G, \circ) und $(H, *)$ Gruppen. Dann heißt eine Abbildung $f : G \rightarrow H$ genau dann **Homomorphismus**, wenn für alle $g_1, g_2 \in G$ gilt:

$$f(g_1 \circ g_2) = f(g_1) * f(g_2).$$

Der Begriff Homomorphismus stammt aus dem Griechischen:

homós: gleich; *morphé*: Form.

Sehen wir uns hierzu einmal vier Beispiele an:

Beispiel 3.1.9. Sei $(G, \circ) = (\mathbb{Z}, +) = (H, *)$. Für $f : \mathbb{Z} \rightarrow \mathbb{Z}$ mit $f(x) = 2 \cdot x$ gilt dann

$$f(x + y) = 2 \cdot (x + y) = 2 \cdot x + 2 \cdot y = f(x) + f(y).$$

Beispiel 3.1.10. Seien $(G, \circ) = (\mathbb{R}, +)$ und $(H, *) = (\mathbb{R}^+, \cdot)$. Für $f(x) = e^x$ gilt

$$f(x + y) = e^{x+y} = e^x \cdot e^y = f(x) \cdot f(y).$$

3 Vektorräume

Beispiel 3.1.11. Seien $(G, \circ) = (\text{GL}(n, \mathbb{R}), \cdot)$ und $(H, *) = (\mathbb{R} \setminus \{0\}, \cdot)$. Dann gilt für $f : \text{GL}(n, \mathbb{R}) \rightarrow \mathbb{R} \setminus \{0\}$ mit $f(A) = \det(A)$

$$f(A \cdot B) = \det(A \cdot B) = \det(A) \cdot \det(B) = f(A) \cdot f(B).$$

Beispiel 3.1.12. Seien $(G, \circ) = (\mathbb{R}, +)$ und $(H, *) = (\text{SO}(2), \cdot)$. Für $f : \mathbb{R} \rightarrow \text{SO}(2)$ mit $f(\theta) = R_\theta$ gilt dann

$$f(\theta + \theta') = R_{\theta + \theta'} = R_\theta \cdot R_{\theta'} = f(\theta) \cdot f(\theta').$$

Die folgenden drei allgemeinen Eigenschaften von (Gruppen-)Homomorphismen entlocken obigen vier Beispielen zahlreiche wichtige Resultate, die wir ohne das Wissen über Gruppen zum Teil mühsam explizit nachrechnen mussten.

Bemerkung 3.1.13. Sind (G, \circ) und $(H, *)$ Gruppen sowie $f : G \rightarrow H$ ein Homomorphismus, so gilt:

1. $f(e_G) = e_H$, d.h. f bildet das neutrale Element von G auf das von H ab.
2. $\forall g \in G : f(g^{-1}) = (f(g))^{-1}$.
3. Ist f bijektiv, dann ist auch $f^{-1} : H \rightarrow G$ ein Homomorphismus.

Beweis.

Zu 1.)

Zunächst einmal gilt

$$f(e_G) = f(e_G \circ e_G) = f(e_G) * f(e_G).$$

Nun multiplizieren wir die Gleichung von links mit dem inversen Element $(f(e_G))^{-1}$ von $f(e_G)$ und erhalten

$$\begin{aligned} e_H &= (f(e_G))^{-1} * f(e_G) = (f(e_G))^{-1} * (f(e_G) * f(e_G)) \\ &= ((f(e_G))^{-1} * f(e_G)) * f(e_G) = e_H * f(e_G) = f(e_G). \end{aligned}$$

Zu 2.)

Wir berechnen

$$f(g^{-1}) * f(g) = f(g^{-1} \circ g) = f(e_G) \stackrel{(1.)}{=} e_H, \quad (3.1)$$

„multiplizieren“ von rechts mit $(f(g))^{-1}$ durch und bekommen

$$\begin{aligned} (f(g))^{-1} &= e_H * (f(g))^{-1} \stackrel{(3.1)}{=} (f(g^{-1}) * f(g)) * (f(g))^{-1} \\ &= f(g^{-1}) * (f(g) * (f(g))^{-1}) = f(g^{-1}) * e_H = f(g^{-1}) \end{aligned}$$

Zu 3.)

Sind $h, h' \in H$, so sind $f^{-1}(h), f^{-1}(h') \in G$. Da f ein Homomorphismus ist, gilt dann

$$f(f^{-1}(h) \circ f^{-1}(h')) = f(f^{-1}(h)) * f(f^{-1}(h')) = h * h'.$$

Wenden wir f^{-1} auf diese Gleichung an, so erhalten wir genau

$$\begin{aligned} f^{-1}(h) \circ f^{-1}(h') &= f^{-1} \left(\underline{f(f^{-1}(h) \circ f^{-1}(h'))} \right) \\ &= f^{-1} \left(\underline{f(f^{-1}(h)) * f(f^{-1}(h'))} \right) \\ &= f^{-1} (h * h'). \end{aligned}$$

□

Für die vorangegangenen Beispiele liefern die Aussagen (1.) bis (3.):

	(3.1.9)	(3.1.10)	(3.1.11)	(3.1.12)
(1.) $f(0) =$	$2 \cdot 0 = 0$	$e^0 = 1$	$\det(\mathbb{1}_n) = 1$	$R_0 = \mathbb{1}_2$
(2.) $f(-x) =$	$2 \cdot (-x) = -2 \cdot x$	$e^{-x} = \frac{1}{e^x}$	$\det(A^{-1}) = \frac{1}{\det(A)}$	$R_{-x} = R_x^{-1}$
(3.)	f nicht surjektiv	✓ (s.u.)	f nicht injektiv	f nicht injektiv

Im Fall von Beispiel (3.1.10) ist der Gruppenhomomorphismus $f : \mathbb{R} \rightarrow \mathbb{R}^+$ mit $f(x) = e^x$ bijektiv und besitzt die Umkehrabbildung

$$f^{-1} : \mathbb{R}^+ \rightarrow \mathbb{R} \quad \text{mit} \quad f^{-1}(y) = \ln(y),$$

welche dann ebenfalls ein Gruppenhomomorphismus ist, d.h. für $y, y' \in \mathbb{R}^+$ gilt die als Logarithmengesetz bekannte Gleichung

$$\ln(y \cdot y') = \ln(y) + \ln(y').$$

Ein bijektiver Gruppenhomomorphismus heißt auch **Gruppenisomorphismus**. Gibt es einen Gruppenisomorphismus, so nennt man die betreffenden Gruppen *isomorph*. Sind zwei Gruppen isomorph, so kann jede Rechnung in der einen Gruppe mit Hilfe des Isomorphismus in eine entsprechende Rechnung in der anderen Gruppe übersetzt werden.

Die Gruppen $(\mathbb{R}, +)$ und (\mathbb{R}^+, \cdot) sind also isomorph. Somit kann die Multiplikation positiver reeller Zahlen auf die Addition reeller Zahlen zurückgeführt werden. Das hat man sich vor dem Siegeszug der Taschenrechner auf Rechenschiebern zunutze gemacht. Sie haben eine logarithmische Skala, mit deren Hilfe man Multiplikationen dadurch durchführen kann, indem man Strecken aneinander legt (also eigentlich eine Addition vornimmt).



Abb. 73

3.2 Körper

Definiert man auf einer Menge zwei Rechenarten mit „guten“ Eigenschaften und sorgt mit einer weiteren Forderung für deren Verträglichkeit, so erhält man einen Körper. Dies ist aber nicht etwa ein räumliches geometrisches Objekt, sondern:

Definition 3.2.1. Ein (Zahl-)Körper ist ein Tripel $(K, +, \cdot)$, wobei K eine Menge ist und $+$ sowie \cdot Verknüpfungen

$$\begin{aligned} + : K \times K &\rightarrow K \\ (x, y) &\mapsto x + y \end{aligned}$$

sowie

$$\begin{aligned} \cdot : K \times K &\rightarrow K \\ (x, y) &\mapsto x \cdot y, \end{aligned}$$

sind, so dass gilt:

1. $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0.
2. Sind $x, y \in K^* := K \setminus \{0\}$, so ist auch $x \cdot y \in K^*$. (Nullteilerfreiheit)
3. (K^*, \cdot) ist eine abelsche Gruppe mit neutralem Element 1.
4. $\forall x, y, z \in K : (x + y) \cdot z = x \cdot z + y \cdot z$. (Distributivität)

Körper zu sein, stellt eine echte Verschärfung da, wie die nun doch ziemlich eingeschränkte Auswahl von einfach zu findenden Beispielen erkennen lässt:

Beispiel 3.2.2. $(\mathbb{Z}, +, \cdot)$ ist kein Körper. Hier ist Axiom 3 verletzt, d.h. für $x \neq \pm 1$ existieren keine multiplikativen Inversen.

Beispiel 3.2.3. Bei $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$ und $(\mathbb{C}, +, \cdot)$ handelt es sich um Körper.

Beispiel 3.2.4. $(\text{Mat}(n \times n, \mathbb{R}), +, \cdot)$ bildet für $n \geq 2$ keinen Körper, da die Nullteilerfreiheit verletzt ist, etwa gilt

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

Bemerkung 3.2.5. Ist K ein Körper, so gilt für alle $x \in K$:

1. $0 \cdot x = 0$, denn

$$\begin{aligned} 0 \cdot x &= (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x \\ \Rightarrow -(0 \cdot x) + 0 \cdot x &= \underbrace{-(0 \cdot x) + 0 \cdot x + 0 \cdot x}_{=0} \\ \Rightarrow 0 &= 0 \cdot x. \end{aligned}$$

2. $(-1) \cdot x = -x$, denn

$$\begin{aligned} (-1) \cdot x + x &= (-1 + 1) \cdot x = 0 \cdot x \stackrel{(1.)}{=} 0 \\ \Rightarrow (-1) \cdot x &= -x. \end{aligned}$$

Beispiel 3.2.6. Wir betrachten nun den einfachsten Körper, den es gibt. Zwei Elemente muss ein Körper auf jeden Fall haben, nämlich 0 und 1. Tatsächlich kommt man damit bereits aus und kann den Körper $(\mathbb{F}_2, +, \cdot)$ bilden, wobei

$$\mathbb{F}_2 := \{0, 1\} \quad \text{sowie} \quad \begin{array}{c|cc} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array} \quad \text{und} \quad \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 1 \end{array}$$

ist. Es ist nicht schwer zu überprüfen, dass die Körperaxiome gelten. Dieser Körper ist auch wichtig für Anwendungen, beschreibt er doch die binären Rechenregeln, die Computer intern benutzen. In diesem Körper gilt z.B. $-1 = 1$.

Bemerkung 3.2.7. Alle Erkenntnisse aus dem Kapitel über Matrixrechnung behalten für Matrizen mit Einträgen in einem beliebigen Körper ihre Gültigkeit. So gilt beispielsweise für $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in \text{Mat}(2 \times 2, \mathbb{F}_2)$, dass

$$\begin{aligned} A \cdot B &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot 1 + 1 \cdot 1 & 1 \cdot 1 + 1 \cdot 0 \\ 0 \cdot 1 + 0 \cdot 1 & 0 \cdot 1 + 0 \cdot 0 \end{pmatrix} \\ &= \begin{pmatrix} 1+1 & 1+0 \\ 0+0 & 0+0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

3.3 Vektorräume

Körper abstrahieren das Rechnen mit Zahlen. Nun abstrahieren wir noch das Rechnen mit Vektoren. Dabei muss man zunächst festlegen, mit welchen Zahlen, d.h. mit welchem Körper man arbeiten will.

Definition 3.3.1. Sei $(K, +, \cdot)$ ein Körper. Ein **K-Vektorraum** bzw. **Vektorraum über K** ist ein Tripel $(V, +, \cdot)$, wobei V eine Menge ist und $+$ sowie \cdot Verknüpfungen

$$\begin{aligned} + : V \times V &\rightarrow K \\ (v, w) &\mapsto v + w \end{aligned}$$

sowie

$$\begin{aligned} \cdot : K \times V &\rightarrow V \\ (\lambda, v) &\mapsto \lambda \cdot v, \end{aligned}$$

sind, so dass gilt:

1. $(V, +)$ ist eine abelsche Gruppe.

2. $\forall \lambda, \mu \in K \forall v, w \in V :$

a) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$

b) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$

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

d) $1 \cdot v = v$

Beispiel 3.3.2. Wir kennen bereits Koordinaten-Vektorräume $V = K^n$ sowie Untervektorräume $V \subset K^n$ jeweils über einem Körper K . Insbesondere ist somit $V = \{\mathbf{0}\}$ ein Vektorraum.

Beispiel 3.3.3. Sei $K = \mathbb{R}$. Dann ist

$$\begin{aligned} V &:= \mathbb{R}[x] \\ &:= \{\text{Polynome mit reellen Koeffizienten}\} \\ &= \{a_n x^n + \dots + a_1 x + a_0 \mid n \in \mathbb{N}_0 \text{ \& } a_j \in \mathbb{R}\} \end{aligned}$$

ein \mathbb{R} -Vektorraum.

Definition 3.3.4. Seien V, W Vektorräume über demselben Körper K . Dann heißt eine Abbildung $f : V \rightarrow W$ genau dann **linear** bzw. **Vektorraum-Homomorphismus**, wenn gilt:

$$1. \forall v, v' \in V : f(v + v') = f(v) + f(v') \quad (\text{Additivität})$$

$$2. \forall \lambda \in K \forall v \in V : f(\lambda \cdot v) = \lambda \cdot f(v) \quad (\text{Homogenität})$$

Die erste Bedingung sagt nichts anderes, als dass f ein Gruppenhomomorphismus von $(V, +)$ nach $(W, +)$ ist. Damit die zweite Bedingung sinnvoll ist, mussten wir sicherstellen, dass beide Vektorräume über demselben Körper definiert sind.

Beispiel 3.3.5. Die uns vertraute Multiplikation mit Matrizen stellt in diesem Sinne gerade den Prototyp für lineare Abbildungen bereit. Ist $V = K^n$ und $W = K^m$ sowie $A \in \text{Mat}(m \times n, K)$, so ist $f : V \rightarrow W$ mit

$$v \mapsto A \cdot v$$

linear.

Beispiel 3.3.6. Für $V = W = \mathbb{R}[x]$ ist $f : V \rightarrow W$ mit

$$p \mapsto \frac{d}{dx} p(x)$$

linear.

Beispiel 3.3.7. Für $V = \mathbb{R}[x]$ und $W = \mathbb{R}$ ist $f : V \rightarrow W$ mit

$$p \mapsto \int_0^1 p(x) dx$$

linear.

Abbildungsnachweise

Nr	Titel	Quelle
	Titelblatt	http://www.sxc.hu
1	Euklid	http://de.wikipedia.org/wiki/Euklid
6	Hilberts Hotel	http://de.wikipedia.org/wiki/Hilberts_Hotel
7	Äpfel und Bananen	http://www.sxc.hu
8	Niels Henrik Abel	http://de.wikipedia.org/wiki/Niels_Henrik_Abel
42	Bundesländer	http://de.wikipedia.org/wiki/Land_(Deutschland)
43	Königsberger Brückenproblem	http://de.wikipedia.org/wiki/Brückenproblem
66	Kaninchen	http://www.sxc.hu
68	Mensch	http://de.wikipedia.org/wiki/Mensch
71	rechte und linke Hand	http://de.wikipedia.org/wiki/Drei-Finger-Regel
73	Rechenschieber	http://www.sxc.hu

Index

- Abbildung, 10
- Abel, Niels Henrik, 101
- Abelpreis, 101
- abelsch, 168
- Addition von Matrizen, 25
- Adjazenzmatrix, 104
- ähnlich, 152
- affin, 133
- affiner Unterraum, 134
- algebraische Multiplizität, 159
- algebraische Vielfachheit, 159
- All-Quantor, 5
- allgemeine lineare Gruppe, 126
- Argument, 85
- aufeinander senkrecht, 75
- Aufpunkt, 56

- Basis, 40
- Basisergänzungssatz, 45
- bijektiv, 13
- Bild, 10
- Bogenmaß, 74

- Cardanische Gleichungen, 98
- Cauchy-Schwarz-Ungleichung, 60
- charakteristisches Polynom, 150
- Cramer'sche Regel, 129

- Definitionsbereich, 10
- Descartes, René, 56
- Determinanten-Funktion, 109
- Determinanten-Multiplikations-Satz, 123
- Diagonalensatz, 68
- diagonalisierbar, 153
- Diagonalmatrix, 129
- Differenz, 8

- Dimension, 43
- Dimension eines affinen Unterraums, 136
- Diskriminante einer kubischen Gleichung, 98
- Diskriminante einer quadratischen Gleichung, 96
- Drehmatrix, 131
- Dreieck, 69
 - nicht entartet, 69
- Dreiecksungleichung, 61, 63
- Durchschnitt, 8

- Eigenraum, 148
- Eigenvektor, 146
- Eigenwert, 146
- Einheitsmatrix, 120
- Einheitswurzel, 95
- elementare Zeilenumformungen, 47
- erweiterte Koeffizientenmatrix, 47
- Euklid, 4
- euklidischer Abstand, 62
- Euler'sche Darstellung, 90
- Euler'sche Formel, 92
- Euler, Leonhard, 90
- Existenz-Quantor, 5

- Fibonacci-Folge, 157
- Fundamentalsatz der Algebra, 92

- Gauß'sche Zahlenebene, 78
- Gauß, Carl Friedrich, 46
- Gauß-Algorithmus, 46
- geometrische Multiplizität, 150
- geometrische Vielfachheit, 150
- geordnetes n -Tupel, 9
- Gerade, 56

Index

- gleichmächtig, 15
- goldene Zahl, 159
- goldener Schnitt, 154
- Gradmaß, 74
- Graph, 103
- Gruppe, 167
- Gruppenisomorphismus, 171

- Hilbert, David, 16
- Hilberts Hotel, 16
- Homomorphismus, 169

- identische Abbildung, 10
- imaginäre Achse, 81
- Imaginärteil, 81
- Induktionsanfang, 2
- Induktionsbehauptung, 2
- Induktionsschritt, 2
- Induktionsvoraussetzung, 2
- injektiv, 13
- Innenwinkel, 73
- inverse Matrix, 126
- invertierbar, 126
- isomorphe Gruppen, 171

- K -Vektorraum, 174
- Königsberger Brückenproblem, 108
- Körper, 172
- Kaninchen, 157
- kartesisches Produkt, 9
- Koeffizienten, 24
- Koeffizientenmatrix, 30
- kommutativ, 168
- komplexe Konjugation, 82
- komplexe Zahl, 78
- Komposition, 12
- Kreisteilungspolynom, 95
- Kreuzprodukt, 162

- Lösungsmenge, 30
- linear, 175
- linear abhängig, 39
- linear unabhängig, 39
- lineare Gleichungssysteme, 21
- lineare Hülle, 34

- lineares Gleichungssystem, 24
 - homogenes, 24
 - inhomogenes, 24
- Linearfaktorzerlegung, 92
- Linearkombination, 34

- $m \times n$ -Matrix, 25
- Matrix, 25
- Menge, 7
- Mittelpunkt, 65
- Mittelpunkts-Lemma, 63
- Moivre-Binet
 - Formel von, 158
- monoton, 60
- Multiplikation von Matrizen, 26

- n -dimensionales Parallelotop, 138
- n -dimensionales Volumen, 137
- n -te Einheitswurzeln, 95
- negativ orientiert, 160
- Norm, 59

- Obermenge, 7
- orthogonal, 75

- p - q -Formel, 96
- parallel, 65
- Parallelepipid, 138
- Parallelogramm, 67
 - nicht entartet, 67
- Parallelotop, 138
- Polardarstellung, 88
- positiv orientiert, 160
- Potenzmenge, 8
- Primzahl, 3
- Pythagoras, Satz des, 76

- Rang, 51
- Raute, 78
- Realteil, 81
- reelle Achse, 81
- Rhombensatz, 76
- Rhombus, 78
- Richtungsvektor, 56

- Salpetersäure, 22

Satz des Pythagoras, 76
Satz von Vieta, 96
Schwerpunkt, 71
Schwerpunktsatz, 70
Seitenhalbierende, 70
Skalare, 26
Skalarprodukt, 59
Spaltenentwicklungssatz von Laplace, 116
Spaltenvektor, 25, 51
spezielle Zeilenstufenform, 48
spezielle Zeilenumformungen, 114
Spiegelungsmatrix, 131
Streckungsfaktor, 56
Streichungsmatrix, 116
surjektiv, 13

Teilmenge, 7
Toluol, 22
Transponierte, 119
Trinitrotoluol, 22

Unbekannte, 24
Untervektorraum, 32
Urbild, 11

Vektorprodukt, 162
Vektorraum über K , 174
Vektorraum-Homomorphismus, 175
Vereinigung, 8
Verkettung, 12
Vielfachheit, 93
Vierfarbenproblem, 108
vollständige Induktion, 2

Wasser, 22
Weg der Länge m , 104
Wertebereich, 10

Zeilenentwicklungssatz von Laplace, 120
Zeilenstufenform, 48
Zeilenvektor, 25