Lineare Algebra I und II

Prof. Richard Pink

Zusammenfassung
Herbstsemester 2014
Frühjahrssemester 2015
ETH Zürich

Korrigierte vollständige Version

6. Oktober 2016

Inhalt

1 Vorbemerkungen

1.1 Mathematik zwischen Formalismus und Anschauung

Ich möchte mit einigen allgemeinen Bemerkungen zur Sprache der Mathematik beginnen und auf Aspekte aufmerksam machen, die ich als besonders wichtig ansehe. Mathematik zu treiben bedeutet generell, mathematische Objekte wie Zahlen, Funktionen, Gleichungen, Ungleichungen, Mengen, Räume, Gruppen und viele weitere zu untersuchen, also wichtige Fragen über diese Objekte zu beantworten und interessante Sätze über sie zu beweisen, aber auch neue Theorien zu entwickeln für vorher noch nicht untersuchte Objekte, ... sicher müsste man diese Liste fortsetzen. All dies spielt sich immer gleichzeitig auf zwei Ebenen ab: der formalen Ebene und der Ebene der abstrakten Anschauung.

Auf der formalen Ebene drücken wir mathematische Sachverhalte aus, indem wir logische und mathematische Symbole zu Formeln verbinden, und führen Beweise, indem wir aus Aussagen, die wir bereits als wahr erkannt haben, durch Anwendung strenger logischer Gesetze neue wahre Aussagen herleiten. Diese Ebene entspricht derjenigen eines Computerprogramms. Ein Programm, das ein Computer ausführen kann, muss strengen syntaktischen und semantischen Regeln gehorchen. Der Computer kann nicht erraten, was der Autor meint, wenn ein wesentliches Detail fehlt, und kann es nicht korrigieren, wenn der Autor etwas Falsches geschrieben hat; und Einwände der Art, es sei natürlich so oder so gemeint gewesen, lässt er nicht gelten. Die formale Sprache der Mathematik ist genauso stringent. Von Mathematikern wird verlangt, dass sie alles, was sie sagen, korrekt und vollständig und unmissverständlich mit allen notwendigen Details auf der formalen Ebene ausdrücken und begründen können. Dies zu lernen und zu üben wird vor allem jetzt im Basisjahr von Ihnen erwartet.

Gleichzeitig wollen wir als Menschen die Welt, also auch die Welt der Mathematik, verstehen und unsere Erkenntnisse darüber anderen Menschen mitteilen. Dazu brauchen wir eine abstrakte Anschauung für die mathematischen Objekte, mit denen wir arbeiten. Wir brauchen ein Verständnis jenseits der formalen Ebene, um erkennen zu können, was wichtig oder unwichtig, nützlich oder unnütz, interessant oder langweilig ist, also um die irrsinnige Menge an Formeln überhaupt zu überblicken und uns beim Umgang damit in die richtige Richtung leiten zu lassen. Wir brauchen dieses Verständnis auch, um zu wissen, welche mathematischen Theorien wir in welcher Situation ausserhalb der Mathematik anwenden können.

Fast jede Kommunikation in der Mathematik erfolgt gleichzeitig auf beiden Ebenen. Wenn ich zum Beispiel in der Geometrie sage: „Jede Gerade enthält mindestens zwei verschiedene Punkte“, so ist das einerseits das sprachliche Äquivalent der formalen Aussage „$\forall g \in \mathcal{G}\ \exists P \in g\ \exists Q \in g : P \neq Q$“, wenn $\mathcal{G}$ die Menge aller Geraden bezeichnet. Andererseits haben wir dabei eine Vorstellung davon, was ein Punkt und eine Gerade sei und was es bedeutet, ob zwei Punkte gleich oder verschieden sind, und dadurch bekommt die Aussage für uns eine anschauliche Bedeutung. Für den Formalismus ist diese Bedeutung irrelevant, für uns Menschen aber nicht. Wir Menschen können Mathematik nur betreiben, indem wir uns gleichzeitig auf beiden Ebenen be-

wegen. Wir können weder rein auf der formalen Ebene operieren, weil wir dann nichts verstünden, noch allein auf der Anschauungsebene, weil wir dann nichts beweisen und somit als wahr oder falsch erkennen könnten.

Nebenbei gesagt ist es grundsätzlich immer möglich und erlaubt, einen mathematischen Formalismus mit einer anderen als der ursprünglich vorgesehenen Anschauung zu versehen. Wenn wir zum Beispiel die Menge aller Schuhschachteln $\mathcal{G}$ nennen und jedes Element $g \in \mathcal{G}$ eine Menge von Schuhen ist, dann bedeutet die oben genannte formale Aussage, dass jede Schuhschachtel mindestens zwei verschiedene Schuhe enthält. Das ist völlig in Ordnung, soweit auch die weiteren postulierten Eigenschaften der Theorie in der neuen Interpretation gelten (was in diesem Beispiel allerdings zweifelhaft erscheint).

Um uns frei auf beiden Ebenen bewegen zu können, müssen wir dazu fähig sein, nach Belieben von der einen auf die andere zu wechseln. Das heisst: Wir müssen alles in die jeweils andere Ebene übersetzen können. Was immer wir auf deutsch, englisch, oder sonst einer natürlichen Sprache sagen, müssen wir auch in mathematischen Formeln ausdrücken können. Umgekehrt sollten wir uns selbst genauso wie anderen stets erklären können, was eine gegebene Formel anschaulich bedeutet. Beides müssen wir ständig praktizieren, während wir mathematische Gedanken entwickeln, aber insbesondere auch auf Nachfrage dann, wenn wir jemand anderem unsere Gedanken mündlich oder schriftlich mitteilen wollen und uns der Adressat um die Übersetzung bittet, damit er genau versteht, was wir meinen. Das ist eine der wichtigsten Grundfähigkeiten, die Sie erwerben sollen. Mängel an dieser Stelle sind oft ein Haupthindernis gegen ein erfolgreiches Mathematikstudium, und darum sollten Sie diese Fähigkeit vor allem im Basisjahr ausführlich einüben.

Sie werden erleben, dass wir zwar zu Beginn sehr auf formale Genauigkeit und Vollständigkeit pochen, dies aber bald reduzieren und uns viele explizite und implizite Abkürzungen erlauben. Das liegt daran, dass die Welt der Mathematik so reichhaltig und komplex ist, dass wir gar nicht dazu in der Lage sind, stets alles so vollständig auszudrücken, wie es die anerkannten Regeln eigentlich erfordern würden. Unsere formalen Aussagen sind also leider oft nicht ganz vollständig, und sogar den besten Mathematikern unterlaufen gelegentliche Irrtümer und sie schreiben formal etwas anderes, als sie tatsächlich meinen. Auch Dozenten von Anfängervorlesungen können keine Vollständigkeit erreichen, schon allein, weil wir gar nicht die Zeit haben, alles auf der formalen Ebene vollständig auszuschreiben. Das wird Ihnen das Verständnis des Stoffs erschweren, bis Sie etwas mehr mathematische Reife entwickelt haben.

Dazu baut der meiste Stoff in der Mathematik auf anderem, davor besprochenem Stoff auf. Dieser wird als bekannt und selbstverständlich vorausgesetzt und nicht weiter diskutiert. Mathematik auf hohem Niveau zu verstehen ist vergleichbar damit, die Funktionsweise eines komplexen elektronischen Geräts wie z.B. eines Handys zu verstehen. Dieses besteht aus verschiedenen Einzelteilen, darunter einem Mikroprozessor, der aus Modulen mit verschiedenen Aufgaben besteht, welche auf bestimmte Weise miteinander interagieren und wiederum in kleinere Einheiten unterteilt sind, bis hin zu Flip-Flops, welche selbst wieder aus einzelnen Transistoren und Widerständen zusammengesetzt sind; daneben läuft auf diesem Mikroprozessor ein Programm, wel-

ches ebenfalls eine geschachtelte Struktur hat, usw. Die Grundlage der elektronischen Bauteile und deren Verdrahtung kann man als Analogon der formalen Ebene der Mathematik, das Interagieren der Module auf den höheren Ebenen als Analogon der Anschauungsebene interpretieren. Genauso entsprechen die einzelnen Befehle des Programms in Maschinensprache der formalen Ebene, seine Gesamtstruktur dagegen der Anschauungsebene der Mathematik. Die Funktionsweise des Geräts als Ganzes kann man nur vollständig erfassen, wenn man auf jeder dieser Komplexitätsebenen Bescheid weiss. Nicht mehr und nicht weniger als das Entsprechende in der Mathematik wird von Ihnen erwartet.

Fast jede mathematische Mitteilung ist also sowohl formal unvollständig als auch darin, dass sie andere Begriffe als bekannt voraussetzt. Die Lösung dieses Problems besteht darin, dass man von Mathematikern erwartet, dass sie alles, was sie sagen, auf Nachfrage hin präziser und vollständiger ausdrücken und erklären und die verwendeten Grundlagen ergänzen können. Auch von Ihnen wird erwartet, dass Sie alles, was Sie lernen, auf der formalen und der Anschauungsebene vollständig erfassen. Solange Sie nicht sicher sind, dass Sie das tun, müssen Sie eben nachfragen, und sollen das auch. Hier kommt es wesentlich auf Ihre eigene Initiative an. Auf Ihre Nachfrage hin erhalten Sie von uns entweder die gewünschten Details, oder Hinweise dazu, wie Sie sich diese Details selbst erarbeiten können.

In jedem Fall ist mathematische Erkenntnis ein aktiver Prozess. Wie beim Fahrradfahren-Lernen geht es darum, gewisse Abläufe so einzuüben, dass sie mit der Zeit automatisch werden. Dabei muss das Gehirn an bestimmten Stellen umverdrahtet werden, und das geschieht nur durch viel Übung und ständige aktive Mitarbeit. Dieser Prozess ist mit Teilerfolgen, aber auch Rückschlägen und Schmerzen verbunden, denn wie man beim Fahrradfahren-Lernen gelegentlich hinfällt, gehört es zur Mathematik, dass man manchmal Fehler begeht und sich dies von anderen sagen lassen muss. Daran muss man sich gewöhnen und es nicht überbewerten. Versuchen Sie zu erreichen, was die folgende Maxime besagt: Mathematiker zeichnen sich den meisten anderen Menschen gegenüber dadurch aus, dass sie sich darüber freuen, wenn man ihnen einen Denkfehler nachweist, weil sie den Erkenntnisgewinn höher bewerten als die damit verbundene Peinlichkeit. Dazu sollten Sie natürlich auch lernen, Ihre eigenen Aussagen ständig auf Korrektheit und Klarheit und Vollständigkeit zu überprüfen. Insgesamt erfordert mathematische Erkenntnis also eine erhebliche psychische Anstrengung, mit der man aber auch einen enormen Gewinn erzielen kann. Das ähnelt der physischen Anstrengung, die man braucht, um auf einen sehr hohen Berg steigen und dort die einzigartige Aussicht geniessen und sich als einer von relativ wenigen Menschen sagen zu können: Das habe ich aus eigener Kraft getan.

1.2 Die Grobstruktur mathematischer Kommunikation

Mathematische Kommunikation, ob mündlich oder schriftlich, folgt gewissen Regeln, im Grossen wie im Kleinen. Im Grossen geschieht Folgendes: Man führt mathematische Objekte ein, trifft Annahmen über sie und zieht schrittweise Folgerungen daraus, beweist also Sätze über sie. Die folgenden Aspekte sind dabei besonders wichtig:

Zuerst müssen die verwendeten Begriffe und Symbole erklärt werden, und zwar bevor man sie benutzt, sofern sie nicht zum gemeinsamen Standardrepertoire aller Beteiligten gehören. Genauso muss gesagt werden, aus welcher Menge die gebrauchten Variablen gewählt werden sollen. Ohne diese Erklärungen kann der Adressat nicht wissen, was gemeint ist. Es ist sehr lästig für diesen, nachfragen zu müssen, was denn damit und damit gemeint sei. Eine vernünftige mathematische Mitteilung beginnt daher oft mit einer Einleitung der Art: „Sei ... ein ... mit der Eigenschaft ..., und sei ...“

Die Variablendeklaration funktioniert genau wie in einer höheren Programmiersprache. Dort hat man die Möglichkeit, Variablen mit einem wohldefinierten Gültigkeitsbereich zu deklarieren, zum Beispiel in einem Unterprogramm. Wie in der Mathematik muss man dort zuerst sagen, welcher Art die Variable sein soll, bevor man sie verwenden darf. (Gewisse ältere Programmiersprachen, die implizite Konventionen für Variablennamen erlauben, verleiten zu Fehlern und sollten heute nicht mehr benutzt werden.) In der Mathematik definiert jeder Quantor einen Gültigkeitsbereich für die quantifizierte Variable, ausserhalb dessen die Variable entweder keinen Sinn ergibt oder, falls sie vorher schon eine Bedeutung hatte, so nimmt sie diese nach Ende des Gültigkeitsbereichs des Quantors wieder an, unabhängig von allem, was während der Gültigkeit des Quantors geschah. Ein Beweis ist wie ein Unterprogramm; was innerhalb eines Beweises eingeführt worden ist, ergibt ausserhalb keinen Sinn.

Wichtig ist, dass Anfang und Ende von Unterstrukturen deutlich markiert werden. Wie für Klammerungen in mathematischen Formeln, insbesondere für Quantoren, muss klar sein, wo eine Definition beginnt und endet, wo ein Satz beginnt und endet, wo ein Beweis beginnt und endet. Der Standard ist auch, dass man das Ende eines Beweises, an dem man also erklärt, dass eine vorher angegebene Behauptung nun bewiesen sei, besonders markiert. Das kann man mit Worten tun wie z.B. „was zu beweisen war“, oder entsprechend lateinisch „quod erat demonstrandum“, meist abgekürzt zu „q.e.d.“, oder ähnlichem. Üblich ist auch das Symbol „□“, der Übersichtlichkeit halber oft am rechten Rand des Textes, wie am Ende des vorliegenden Absatzes. Wenn man dieses Symbol verwendet, so sollte man es aber korrekt tun und nicht etwa als Zeichen für „irgendetwas endet hier“: □

Kleine, mittlere, und grosse Sätze unterscheidet man wie folgt: Einen zentralen Satz einer Theorie nennt man oft Hauptsatz. Das Wort Theorem entspricht dem englischen und französischen Wort für Satz; im deutschen wird es oft für einen grossen Satz verwendet. Einen mittleren oder kleinen Satz, dessen Aussage man im Rahmen der zu entwickelnden Theorie für relevant hält und später weiter zu benutzen gedenkt, nennt man oft Proposition. Das ist lateinisch für Behauptung, das deutsche Wort Behauptung benutzt man jedoch in der Regel nur für eine Zwischenbehauptung innerhalb eines Beweises, für welche man sich ausserhalb des Beweises nicht mehr interessiert. Ein Korollar oder eine Folge ist ein Satz, der mit wenig Aufwand aus einem grösseren Satz folgt. Ein Lemma ist ein kleiner Satz, der dazu dient, eine Proposition oder einen Satz zu beweisen, der aber nicht selbst von Interesse ist. Ein Lemma, das nur dazu dient, ein anderes Lemma zu beweisen, heisst Sublemma. Ein Lemma kann für sich alleine oder innerhalb eines Beweises auftreten. Oft enthält es eine mehr oder weniger hässliche formale Aussage im Rahmen einer Rechnung, wogegen der dazugehörige grössere Satz

eine griffigere strukturelle Eigenschaft ausdrückt. Wenn man einen mathematischen Text überfliegt, orientiert man sich vor allem an den Definitionen und grösseren Sätzen und lässt Beweise und Lemmata beim ersten Durchgang oft ganz beiseite. Gelegentlich hat man im Nachhinein festgestellt, dass ein ursprünglich als Lemma ausgedrückter Sachverhalt doch eine fundamentale Bedeutung hat. Als man darauf das Lemma zu einem Satz befördern wollte, hatte sich die Bezeichnung Lemma schon eingebürgert, und darum behielt man sie bei. Ein typisches Beispiel dafür ist das Zornsche Lemma.

Formal gesehen ist ein Beweis eine endliche Folge von Aussagen, von denen jede entweder ein Axiom oder eine früher bewiesene Aussage ist oder mittels einer Schlussregel aus im Beweis davor stehenden Aussagen folgt, so dass die letzte Aussage des Beweises die zu beweisende Aussage ist. Zum vollständigen Aufschrieb eines Beweises gehört, für jede Aussage zu erklären, wieso sie gilt. Wenn zum Beispiel eine Aussage direkt aus der unmittelbar vorhergehenden folgt, so kann man dies mit einem Doppelpfeil $\Rightarrow$ anzeigen. Wenn dabei noch eine bestimmte Grundeigenschaft oder ein Satz der Theorie benutzt wird, so muss man dies ebenfalls erwähnen. Wenn eine Aussage aus einer früheren Aussage oder einer Kombination von solchen folgt, so muss man auch dies irgendwie aufzeigen. Ein verständlicher Beweis ist also nicht eine blosse Folge von Aussagen oder Formeln, sondern erfordert für jeden Schritt eine Begründung. Sonst ist er lückenhaft und wird zu Recht nicht akzeptiert. Genauso ist eine Berechnung, zum Beispiel eines Integrals, nicht bloss eine Folge von Umformungen einer mathematischen Formel, sondern eine Folge von Umformungen mit jeweiligen Begründungen.

1.3 Vorbilder

Zu Beginn versuchen wir in den Vorlesungen, jeden elementaren Rechen- und Beweisschritt auszuschreiben und zu begründen. Das Gleiche verlangen wir von Ihnen in den Übungsaufgaben. Das verlangen wir auch dann noch, wenn wir in der Vorlesung schon längst aus Zeitgründen damit begonnen haben, bestimmte Details zu überschlagen. Denn als Mathematiker/innen im Werden müssen Sie diese Grundlagen noch lange üben, bis sie für Sie selbstverständlich geworden sind. Sie müssen erst noch das richtige Gefühl dafür entwickeln, welche Details wichtig und welche weniger wichtig sind. Darum sind in den Übungsaufgaben, und natürlich genauso in den Klausuren, grundsätzlich alle Details aufzuschreiben und alle Begründungen explizit anzugeben.

Gute Vorbilder dafür finden Sie in vielen Skripten und Lehrbüchern, aber nicht notwendigerweise in allen. In meinen eigenen Vorlesungen bestehe ich darauf, dass die Musterlösungen der Übungsaufgaben mustergültig sind, aber auch das ist nicht überall so. Am besten ist es, Sie bilden sich selbst eine fundierte Meinung darüber, was für Sie mustergültig ist und was nicht. Vielleicht wählen Sie sich ein Vorbild und versuchen ihm nahezukommen, denken aber auch später einmal daran zu überprüfen, ob Sie nicht noch andere Vorbilder finden. Auf höherem Niveau sind die renommiertesten Fachzeitschriften in der Regel gute Vorbilder.

Trotz unvermeidbarer Lücken hoffe ich, dass auch der Inhalt meiner Vorlesung diesen Anforderungen entspricht und Vorbildcharakter hat. Jedoch wird mein Tafelanschrieb

allein keineswegs vorbildlich sein. Denn vieles, was ich mündlich dazu sage und was auch wichtig ist, werde ich aus Zeitgründen nicht anschreiben. Insbesondere werde ich viele Begründungen nur mündlich angeben. Darum wird es auch nicht ausreichen, dass Sie nur den Tafelanschrieb kopieren, sondern Sie sollten sich die wichtigsten mündlichen Nebenbemerkungen ebenfalls merken.

1.4 Die Sprache der Prädikatenlogik

Mathematische Formeln bestehen im Innern aus Konstanten, Variablen, Funktionen und Relationen. Mit diesen Symbolen kann man Aussagen wie $1 + 1 = 2$ und $(a + b)^2 = a^2 + 2ab + b^2$ ausdrücken. Solche einzelnen Aussagen werden verbunden mit den logischen Symbolen 'und' $\wedge$, 'oder' $\vee$, 'nicht' $\neg$, 'impliziert' $\to$, 'dann und nur dann, wenn' oder 'genau dann, wenn' $\leftrightarrow$, sowie den Quantoren 'für alle' $\forall$ und 'es existiert' $\exists$. Der etwas weniger allgemeingebräuchliche Quantor $\exists!$ bedeutet 'es existiert ein und nur ein' oder 'es existiert genau ein'. Klammern der Form $(\ldots)$ oder $[\ldots]$ dienen dazu, Teilformeln zusammenzufassen und eindeutig festzulegen, wie sie als Teil einer grösseren Formel aufzufassen sind. Gelegentlich haben sie wie Mengenklammern $\{\ldots\}$ auch spezielle Bedeutungen.

Ein Quantor bezieht sich immer auf eine nachfolgende Aussage; zum Beispiel steht „$\exists x\ C(x)$“ für „Es existiert ein x mit der Eigenschaft $C(x)$“. Meist schreibt man „$\forall x \in X : C(x)$“ als Abkürzung für „$\forall x : (x \in X) \to C(x)$“ und „$\exists x \in X : C(x)$“ als Abkürzung für „$\exists x : (x \in X) \wedge C(x)$“ und analog für $\exists!$. Der Doppelpunkt dient hier nur dazu, verschiedene Formelteile besser auseinanderhalten zu können, und hat keine weitere mathematische Bedeutung. Einen Doppelpunkt oder senkrechten Strich benutzt man auch, um durch eine Eigenschaft eine Teilmenge einer Menge zu spezifizieren in der Form $\{x \in X : C(x)\} = \{x \in X \mid C(x)\}$. Für Zuweisungen verwendet man das Symbol $:=$, wobei der Doppelpunkt stets auf der Seite des Symbols steht, dem die andere Seite als Wert zugewiesen wird, wie in $x := 2$.

Die Implikation $A \to B$ als Teil einer mathematischen Formel ist zu unterscheiden von der Folgerung „Wenn A gilt, dann gilt B“ oder der Begründung „Da A gilt, folgt B“ als Teil der natürlichen menschlichen Sprache. In letzteren beiden Fällen behauptet man die Folgerung aufgrund eines inneren Zusammenhangs zwischen den Aussagen A und B. Solche Folgerungen ziehen wir oft im normalen Leben und ständig, wenn wir Mathematik betreiben. Sie sind nicht als Teil von Formeln anzusehen, sondern erklären die Beziehung zwischen verschiedenen Formeln.

Bei der Implikation $A \to B$ als Teil einer mathematischen oder logischen Formel wird dagegen kein innerer Zusammenhang zwischen A und B verlangt, schon weil man gar nicht formal spezifizieren kann, was das genau bedeuten soll. Stattdessen sieht man die Implikation einfach dann als wahr an, wenn B wahr ist oder wenn A falsch ist, und als falsch nur dann, wenn B falsch ist und A wahr. Insbesondere kann die Implikation den Wahrheitswert „falsch“ annehmen und beliebig als Teil einer grösseren Formel auftreten.

Entsprechendes gilt für die Beziehung zwischen der Äquivalenz $A \leftrightarrow B$ als Teil einer

Formel und Aussagen der Form „A gilt dann und nur dann, wenn B gilt“ in der natürlichen Sprache.

Natürlich führt das zu einer gewissen Verwirrung. Experten der mathematischen Logik empfehlen, Implikation und Äquivalenz als Teile von Formeln mit einfachen Pfeilen zu bezeichnen, solche ausserhalb dagegen mit den Doppelpfeilen $\Rightarrow$ and $\Leftrightarrow$ abzukürzen. Die Unterscheidung hat sich unter Mathematikern jedoch noch nicht durchgesetzt. Ich versuche dennoch in dieser Vorlesung, mich an die beschriebene Konvention zu halten.

Die vorigen Bemerkungen sind auch ein Beispiel dafür, dass man die Symbole der Prädikatenlogik nicht einfach als Abkürzung für Teile der natürlichen Sprache ansehen sollte. Ein weiteres Beispiel sind Quantoren. In einem natürlichen Satz ist es z.B. in Ordnung zu sagen, „dass $x-x=0$ ist für alle reellen Zahlen x“. In diesem Fall steht die Quantifizierung „für alle x“ nach der quantifizierten Aussage „$x-x=0$“; die Regeln der natürlichen Sprache sorgen dennoch dafür, dass wir in diesem wie in den meisten Fällen eindeutig verstehen, was gemeint ist. Im mathematischen Formalismus dagegen gelten strengere Regeln, aus guten Gründen. Wie für Computerprogramme hat man dort festgelegt, dass jeder Quantor vor der durch ihn quantifizierten Aussage stehen muss. Dadurch wird die Gültigkeit des Quantors eindeutig festgelegt für den Bereich von dem Quantor bis zum Ende der Formel oder der nächsten schliessenden Klammer, die zu einer vor dem Quantor stehenden öffnenden Klammer gehört. Ohne Beachtung dieser Regel käme man bald zu Formeln der Art „$\exists x\, D(x,y)\, \forall y$“, bei denen keine eindeutige logische Präzedenz der Quantoren mehr auszumachen ist. Die Beachtung der Regel macht dagegen den entscheidenden Unterschied zwischen Formeln deutlich, bei denen Quantoren vertauscht wurden, wie z.B. bei

$$\forall z \in \mathbb{C}\ \exists w \in \mathbb{C} : w^2 = z \quad \text{gegenüber} \quad \exists w \in \mathbb{C}\ \forall z \in \mathbb{C} : w^2 = z.$$

Beim Umgang mit dem Allquantor $\forall$ ist weiter zu bemerken — und das ist etwas gewöhnungsbedürftig — dass im Fall der leeren Menge $X = \varnothing$ jede Aussage der Form „$\forall x \in X : C(x)$“ wahr ist. Das liegt nicht etwa daran, dass die leere Menge doch irgendwelche Elemente hätte, sondern daran, dass die Aussage ja bedeutet „$\forall x : (x \in X) \to C(x)$“ und die Implikation „$(x \in X) \to C(x)$“ immer wahr ist, wenn die Voraussetzung „$x \in X$“ falsch ist. Dies gilt auch dann, wenn $C(x)$ eine von vorneherein absurde Aussage über x ist. Zum Beispiel ist die Aussage „Jedes Element der leeren Menge ist ein grosser grüner Bär, der sein eigener Vater ist,“ wahr.

Missverständnisse mit der leeren Menge können auch dadurch entstehen, dass man ihre Eigenschaften mit den Eigenschaften ihrer Elemente verwechselt. Zum Beispiel gibt es bekanntlich keine reelle Zahl x mit der Eigenschaft $x+1<x$. Dagegen gibt es durchaus eine Menge reeller Zahlen X mit der Eigenschaft $\forall x \in X : x+1<x$, nämlich die leere Menge $X=\varnothing$. Um solche Verwechslungen zu vermeiden, sollte man sich stets genau überlegen, worauf sich ein Quantor bezieht und worauf nicht.

Man muss sich dessen bewusst sein, dass, wenn eine Menge X leer oder möglicherweise leer ist, uns dies nicht verbietet, über Elemente von X zu sprechen. Dass wir dies können und dürfen, ist sogar essentiell wichtig. Wenn wir zum Beispiel für alle Elemente x einer Menge X eine Aussage $C(x)$ beweisen wollen, so können wir dies oft

durch einen Widerspruchsbeweis erreichen, indem wir annehmen, es gäbe ein Gegenbeispiel, das heisst, ein Element der Menge $X' := \{x \in X \mid \neg C(x)\}$. In diesem Fall versuchen wir zu zeigen, dass die Menge X' leer ist, indem wir die Eigenschaften eines hypothetischen Elements x erkunden und schliesslich zu einem Widerspruch führen. Wir sprechen also absichtlich über ein Element einer Menge, von der wir in Wirklichkeit hoffen, dass sie leer ist. Die Schlussfolgerungen, die wir dabei benutzen, werden ja nicht falsch, wenn die Anfangsvoraussetzung, es gäbe ein Element $x \in X'$, falsch ist; im Gegenteil, bei einer falschen Voraussetzung A wird eine Implikation $A \to B$, wie oben besprochen, ja erst recht richtig. Für mathematische Beweise, insbesondere für Widerspruchsbeweise, müssen wir also argumentieren können, ohne zu wissen, ob die Voraussetzungen überhaupt erfüllbar sind. Die Regeln darüber, wann zusammengesetzte Aussagen der Form $A \to B$ u.a. richtig sind, ermöglichen uns genau das.

Eine gute Übung für den Umgang mit dem mathematischen Formalismus besteht darin, beliebige Sätze der natürlichen Sprache in die Sprache der Prädikatenlogik zu übersetzen und umgekehrt. Wenn zum Beispiel X die Menge aller Menschen bezeichnet, so kann man die Aussage „Jeder Mensch hat einen Vater“ formal ausdrücken durch „$\forall x \in X\ \exists y \in X : (y \text{ ist Vater von } x)$“. Dass dieser Vater dann auch noch eindeutig ist, kann man ausdrücken durch „$\forall x \in X\ \exists!\, y \in X : (y \text{ ist Vater von } x)$“, oder ausgeschrieben durch

$$\forall x \in X\ \exists y \in X : (y \text{ ist Vater von } x) \wedge \forall z \in X : (z \text{ ist Vater von } x) \to z = y.$$

Üben Sie insbesondere den Umgang mit Implikationen und Quantoren. Ein weiteres Beispiel: Wenn ich sage: „Jedesmal, wenn ich Stöckelschuhe trage, fühle ich mich unsicher auf den Beinen“, so ist diese Aussage in meinem Fall nicht etwa deshalb richtig, weil ich zwar selten, aber eben doch manchmal Stöckelschuhe tragen und mich dann aus Mangel an Übung unsicher fühlen würde, sondern deshalb, weil ich es nie tue. Aus demselben Grund ist die Aussage „Jedesmal wenn ich Stöckelschuhe trage, habe ich einen Sechser im Lotto“ in meinem Fall wahr; trotzdem hatte ich leider noch nie einen Hauptgewinn.

Die Übersetzung zwischen natürlicher Sprache und Prädikatenlogik ist nie ganz eindeutig, und es ist auch eine gute Übung, verschiedene äquivalente Übersetzungen zu finden. Zum Beispiel ist die Implikation $A \to B$ äquivalent zu ihrem Kontrapositiv $\neg B \to \neg A$ sowie zu $\neg A \vee B$, die Aussage $\neg\exists x : C(x)$ ist äquivalent zu $\forall x : \neg C(x)$, die Aussage $\neg\forall x : C(x)$ ist äquivalent zu $\exists x : \neg C(x)$, und anderes mehr.

1.5 Widersprüche

Ein berühmter Widerspruchsbeweis ist Russells Paradoxon. Ein Paradoxon ist ein vermeintlicher Widerspruch in einer Theorie. Eine Grundkonstruktion der Mengenlehre besagt, dass man zu einer beliebigen formalen Eigenschaft $C(x)$ für Elemente x einer beliebigen Menge X eine Teilmenge $Y := \{x \in X \mid C(x)\}$ spezifizieren kann mit der Eigenschaft: $\forall x \in X : x \in Y \leftrightarrow C(x)$. Wenn wir dies akzeptieren, und es eine „Menge aller Mengen“ gäbe, so gäbe es auch eine Teilmenge davon, deren Elemente genau

diejenigen Mengen sind, welche sich nicht selbst enthalten. Russells Paradoxon zeigt aber, dass es eine solche Menge nicht geben kann:

Satz: Es gibt keine Menge S mit der Eigenschaft, dass S genau diejenigen Mengen enthält, welche sich nicht selbst enthalten. In Symbolen, wenn M die Kollektion aller Mengen bezeichnet: $\neg\exists S \in M\ \forall X \in M : X \in S \leftrightarrow X \notin X$.

Beweis: Sei doch S eine Menge mit der genannten Eigenschaft. Dann können wir diese Eigenschaft insbesondere auf die Menge $X := S$ anwenden und erhalten die Folgerung $S \in S \leftrightarrow S \notin S$. Das bedeutet, dass die Aussage $S \in S$ wahr ist genau dann, wenn sie falsch ist. In jedem Fall ist diese Äquivalenz ein Widerspruch. Somit muss die Annahme, dass es eine solche Menge S gibt, falsch sein. **q.e.d.**

Russells Paradoxon hat zur Folge, dass man neue Mengen nicht allein durch Eigenschaften definieren kann wie in $\{x \mid C(x)\}$, sondern dass man bei der Konstruktion neuer Mengen immer irgendwie von bereits bekannten Mengen ausgehen muss wie z.B. in $\{x \in X \mid C(x)\}$ oder der Bildung von Potenzmengen. Die heute allgemein akzeptierten Axiome der Mengenlehre nach Zermelo und Fränkel leisten genau dies. Ob diese aber auf irgendeinem anderen Weg zu einem Widerspruch führen, wissen wir nicht. Aus einem fundamentalen Satz von Gödel folgt sogar, dass es gar nicht möglich ist, ihre Widerspruchsfreiheit zu beweisen, wenn sie denn widerspruchsfrei sind. Eine vollständige Klärung der logischen Grundlagen unserer Disziplin bleibt uns also endgültig verwehrt.

Wenn wir bei unserer mathematischen Tätigkeit auf einen Widerspruch stossen, ist es theoretisch immer möglich, dass dies ein Widerspruch ist, der das gesamte auf der Mengenlehre errichtete Gebäude der Mathematik zum Einsturz bringt. In der Praxis hat sich aber bisher jeder solche vermeintliche Widerspruch als Folge eines Irrtums herausgestellt.

1.6 Irrtümer

Auch in der Mathematik liegt der Teufel im Detail, und die Möglichkeiten, Fehler zu begehen, sind grenzenlos. Ich will hier nur einige Fallen erwähnen, in die auch erfahrene Mathematiker immer wieder tappen.

Ein verbreiteter Irrtum besteht darin, mathematische Objekte ohne nachzudenken als verschieden anzusehen, wenn sie verschiedene Namen tragen. Dabei dürfen natürlich verschiedene Symbole dasselbe Objekt bezeichnen, genauso wie verschiedene Variablen denselben Wert annehmen dürfen. Wenn wir zum Beispiel sagen „Seien x und y reelle Zahlen“, so erlaubt dies selbstverständlich, dass sie gleich sind. Genauso schliesst die Aussage der Geometrie „Seien P und Q Punkte einer Geraden g“ die Möglichkeit $P = Q$ mit ein. Auch in Aussagen der Form „Für je zwei Elemente von ... gilt“ werden die Elemente nicht automatisch als verschieden vorausgesetzt, zum Beispiel wenn man sagt: „Für je zwei Elemente a und b einer Gruppe G existiert ein eindeutiges Element x von G mit $ax = b$“. Was man als verschieden voraussetzen will, muss man also zu Beginn klar benennen, um Mehrdeutigkeiten und Missverständnisse zu vermeiden. So-

bald man eine Aussage in Formeln ausdrückt, wie in „$\forall g \in \mathcal{G}\ \exists P \in g\ \exists Q \in g\colon P \neq Q$“, wird meist klar, was gemeint ist.

Noch ein Beispiel dazu: Nach Definition hat jeder unitäre Ring R ein Nullelement 0_R und ein Einselement 1_R, und wenn man nicht aufpasst, nimmt man unbewusst an, dass diese verschieden seien. Dabei können sie durchaus gleich sein, nämlich für den Nullring. Wo man dies verbieten will, z.B. in den Axiomen für Körper, muss man es extra dazu sagen. Genauso neigt man generell dazu, sich jegliche mathematische Objekte als „nicht-trivial“ vorzustellen, also Mengen und Räume als nichtleer, Gruppen und Vektorräume und Ringe als aus mehr als einem Element bestehend, und so weiter.

Eine gute Vorkehrung gegen Irrtümer besteht daher darin, alle Aussagen anhand von Extrembeispielen zu testen. Zum Beispiel wende man eine Aussage über Mengen auf die leere Menge an, eine Aussage über Vektorräume auf den Nullraum, eine Aussage über Zahlen auf die Zahl 0, eine Aussage über Gruppen auf die Gruppe mit einem Element, usw., oder man betrachte eben den Fall, dass gegebene Objekte trotz verschiedener Bezeichnungen gleich sind.

Ein ähnliches Problem entsteht beim unbedachten Verwenden der Alternative „entweder ... oder“. Ausserhalb der Mathematik benutzt man diese oft, ohne sich zu überlegen, ob man das gleichzeitige Zutreffen beider Möglichkeiten wirklich ausschliessen möchte oder nicht. Innerhalb der Mathematik, bzw. der Logik, bedeutet „entweder ... oder“ aber das ausschliessende „oder“, wohingegen das einschliessende logische „oder“ in Wirklichkeit viel häufiger richtig ist. Man kann dieses Problem vermeiden, indem man sich bewusst dazu erzieht, das Wort „entweder“ gar nicht erst zu benutzen, es also aus dem aktiven Wortschatz streicht, bis man seine Sprechgewohnheiten entsprechend umgestellt hat.

Ein weiterer häufiger Denkfehler liegt darin, ein Objekt durch gewisse Eigenschaften zu charakterisieren und dann von „dem“ Objekt zu sprechen, ohne geklärt zu haben, ob es denn existiert und durch diese Eigenschaften eindeutig bestimmt ist. Vor allem die Eindeutigkeit wird leicht vergessen. Wenn wir zum Beispiel einen Vektorraum haben, so dürfen wir zwar eine Basis wählen und danach von „der“ (gewählten) Basis sprechen. Wir dürfen aber nicht einfach so von „der“ Basis sprechen, weil es im allgemeinen verschiedene gibt. Die korrekte Definition der Dimension eines Vektorraums lautet daher „die Kardinalität *einer* Basis“, und damit das wohldefiniert ist, muss man beweisen, dass eine Basis existiert und dass jede Basis dieselbe Kardinalität besitzt. Danach ist es in Ordnung, von „der“ Dimension des Vektorraums zu sprechen. Genauso sprechen wir erst dann von „dem“ Einselement einer Gruppe oder eines Körpers, wenn wir bewiesen haben, dass dieses eindeutig bestimmt ist; bis dahin müssen wir uns mit der Formulierung „ein Einselement“ begnügen.

Nach einer verbreiteten Konvention meint man mit „der“ Quadratwurzel einer nichtnegativen reellen Zahl x stets die eindeutige nichtnegative reelle Zahl y mit der Eigenschaft $y^2 = x$. Die Bezeichnung $\sqrt{x}$ für diese ist durch ihre Eindeutigkeit gerechtfertigt. Für Quadratwurzeln von komplexen Zahlen existiert dagegen keine vernünftige Vorzeichenregel; darum darf man dort stets nur von „einer“ Quadratwurzel sprechen und muss zuerst eine geeignete wählen, bevor man mit dieser weiter arbeiten kann.

Noch ein Beispiel aus der ebenen Geometrie: Ein *Kreis* K ist definiert als die Menge aller Punkte, die von einem gegebenen Punkt O einen gegebenen positiven Abstand r haben. Den Punkt O nennt man dann *Mittelpunkt* und die Zahl r *Radius von* K. Diese Definition alleine schliesst aber nicht aus, dass eine andere Wahl von O und r dieselbe Punktmenge K liefern kann. Bevor man geklärt hat, ob dies in einer gegebenen Geometrie der Fall ist, darf man daher nicht von „dem“ Mittelpunkt und „dem“ Radius eines Kreises sprechen, ohne solche extra gewählt zu haben.

Generell sollte man also stets auf eine saubere Formulierung achten und aufpassen, wo man den bestimmten Artikel und wo den unbestimmten Artikel verwendet.

Missverständnisse können entstehen, wo einander widersprechende Konventionen gebräuchlich sind. Allgemein akzeptiert ist, dass Zahlen $x > 0$ positiv und Zahlen $x \geqslant 0$ nichtnegativ heissen, also insbesondere, dass die Zahl 0 weder positiv noch negativ ist. Die Mehrheit der Mathematiker folgt der Konvention, dass die Bezeichnung $X \subset Y$ für eine Inklusion von Mengen auch die Gleichheit erlaubt, und schreiben $X \subsetneq Y$ oder $X \subsetneqq Y$ für eine echte Inklusion. Andere schreiben für ersteres sicherheitshalber $X \subseteq Y$; was sie dann mit $X \subset Y$ meinen, ist nicht immer klar. Völlig durcheinander geht der Gebrauch des Symbols $\mathbb{N}$ für die Menge der natürlichen Zahlen: Für manche schliesst es die Zahl 0 mit ein, für andere nicht. Ich empfehle daher, das Symbol $\mathbb{N}$ gar nicht erst zu verwenden, sondern stattdessen klarer $\mathbb{Z}^{\geqslant 0}$ bzw. $\mathbb{Z}^{>0}$ zu schreiben. Meine persönliche Meinung ist, dass die natürlichen Zahlen die möglichen Kardinalitäten endlicher Mengen sind und daher mit der Kardinalität der leeren Menge, also mit 0 beginnen. In der mathematischen Logik ist das allgemein akzeptiert.

Viele Irrtümer beginnen mit Worten der Art „Offensichtlich gilt ...“ oder verstecken sich in Formulierungen wie „Das ist trivial“ oder „Man zeigt leicht, dass ...“. Viel zu oft übertünchen diese lediglich die Tatsache, dass es einem zu mühsam war, sich die Details zu überlegen. In Wirklichkeit weiss man also genau, dass man eine erhebliche Lücke lässt, will sich selbst und anderen aber aus Bequemlichkeit das Gegenteil weismachen und begeht dabei leicht ernsthafte Fehler.

Natürlich muss man andere nicht mit langweiligen Rechnungen quälen. Sich selbst aber schon, wo man nicht die gleiche Rechnung schon hundertmal gemacht hat. Oft stellt man dabei fest, dass die Sache viel weniger trivial war, als man gedacht hatte. Man sollte sich also wirklich stets alle Details überlegen und zumindest für sich dokumentieren. Wenn man dann in der Kommunikation Details weglässt, kann man dem Adressaten genau sagen, worauf es dabei ankommt, und muss sich nicht mit Wischiwaschi-Formulierungen herausreden. Dann kann man zum Beispiel sagen „Durch direkte Anwendung von ... und ... folgt ...”. Das ist viel hilfreicher, und kaum länger, als „Jetzt zeigt man leicht ...“.

Ausserdem hängt es stets vom Zusammenhang ab, was man als leicht und offensichtlich oder als schwer und undurchsichtig empfindet. Mathematik hat die faszinierende Eigenschaft, dass einem etwas für eine Weile als völlig unverständlich erscheinen kann, bis sich die Ideen im Kopf so angeordnet haben, dass es einem danach als die natürlichste Sache der Welt vorkommt. Dann vergisst man nur zu leicht, wie sehr man um die neue Erkenntnis ringen musste. Vor dem Moment des Aha-Erlebnisses war die Sache dann eben weder trivial noch offensichtlich.

Diesen Denkprozess beim Adressaten sollte man bei jeder mathematischen Mitteilung berücksichtigen, schon aus Respekt vor dem anderen Menschen. Ohnehin hört dieser bei Worten wie „trivial“ und „offensichtlich“ schnell den unterschwelligen Vorwurf „Du bist dumm, wenn du das nicht verstehst“. Solche Worte werden schnell zu Waffen, die den Adressaten beleidigen und zurückstossen, auch wenn sie nicht so gemeint sind. Man sollte sie daher nur sehr zurückhaltend verwenden. Zum Respekt vor anderen gehört auch, klar zu kommunizieren und Mehrdeutigkeiten zu vermeiden und sie nicht raten zu lassen, was man denn eigentlich meint.

1.7 Die axiomatische Methode

Wie oben erklärt, besteht das Ziel der Mathematik darin, Sätze über die untersuchten mathematischen Objekte zu beweisen, also wahre Aussagen darüber zu etablieren. Dies tun wir, indem wir neue Aussagen mittels logischer Schlussregeln aus anderen, bereits davor als wahr erkannten Aussagen herleiten. Vielleicht haben wir jene auf noch frühere wahre Aussagen zurückgeführt, und so weiter, aber wir können diesen Regress nicht unendlich oft durchführen. Wir müssen uns daher auf geeignete Anfangsaussagen einigen, welche wir als wahr postulieren und nicht länger zu beweisen versuchen. Diese Anfangsaussagen nennt man Axiome.

Die sogenannte *axiomatische Methode* besteht darin, geeignete Axiomensysteme herauszusuchen und alle weiteren Sätze nur aus diesen herzuleiten. Da die Axiome innerhalb des formalen Systems nicht zu beweisen sind, müssen wir uns damit begnügen, solche Axiome zu wählen, die uns aus unserer menschlichen Anschauung heraus als besonders plausibel erscheinen. Dabei können wir meist nicht beweisen, dass wir nicht einem fundamentalen Irrtum erliegen. Zum Beispiel kann man nicht beweisen, dass die üblichen Axiome der natürlichen Zahlen oder der Mengenlehre widerspruchsfrei sind. Wir können nur versuchen, Probleme zu minimieren, indem wir möglichst wenige und möglichst einfache Axiome wählen.

Oft hat man bei der Wahl der Axiome eine gewisse Freiheit, und verschiedene Axiomensysteme für dieselben Objekte stellen sich als äquivalent heraus. Die etablierten Axiomensysteme für die gängigsten mathematischen Begriffe haben sich jedoch als günstig bewährt.

Axiome für die gesamte Mathematik umfassen solche für die Prädikatenlogik und die Mengenlehre. Diejenigen, welche man heute verwendet, gehen auf das erste Drittel des 20. Jahrhunderts zurück und werden von fast allen Mathematikern akzeptiert. Dazu gibt es Axiome für die natürlichen Zahlen und die reellen Zahlen, welche man aber auf die üblichen Axiome der Mengenlehre zurückführen kann, das heisst: Im Rahmen der Mengenlehre kann man beweisen, dass die Axiomensysteme für die natürlichen Zahlen und die reellen Zahlen erfüllbar sind.

Axiomensysteme benutzt man ausserdem, um spezielle mathematische Objekte wie Körper, Gruppen, Ringe, Vektorräume, topologische Räume, usw. zu definieren. Solche Objekte bestehen in der Regel aus einer oder mehreren Mengen sowie Funktionen und Relationen auf diesen, welche gemeinsam gewisse Eigenschaften, eben Axiome,

erfüllen müssen. Das Studium dieser Objekte besteht dann einerseits darin, Folgerungen aus diesen Axiomen herzuleiten, und andererseits darin, die Beziehungen zwischen verschiedenen Objekten, welche jedes für sich die Axiome erfüllen, zu klären.

1.8 Euklids „Elemente“

Die axiomatische Methode wurde im antiken Griechenland entwickelt und vor rund einem Jahrhundert endgültig klar formuliert und zur Grundlage der Mathematik erklärt. Dieser Methode folgt das Werk „Elemente“ aus dem 4. Jahrhundert vor unserer Zeitrechnung, in dem Euklid das im Mittelmeerraum verfügbare geometrische Wissen seiner Zeit zusammengefasst hat.

Er beginnt mit einigen Erläuterungen wie „*Punkt* ist, was ohne Teil ist“ und „*Linie* ist Länge ohne Breite“. Er benennt also zuerst die Begriffe, mit denen er im folgenden arbeiten will. Seine Definitionen haben allerdings keinen formalen mathematischen Sinn, weil nicht bereits vorher erklärt worden ist, was „Teil“ und „Länge“ und „Breite“ bedeuten. Sie sind eher wie Definitionen aus einem Wörterbuch, in dem die Bedeutung jedes Worts mittels anderer Wörter erklärt wird und das nur sinnvoll benutzen kann, wer bereits einen nicht näher bestimmten Basiswortschatz besitzt. Der einzige sinnvolle mathematische Inhalt der zitierten Stellen liegt in der Ankündigung „Wir werden im folgenden von Punkten und Linien sprechen, deren Eigenschaften wir noch näher angeben werden“. Oder in die heutige Sprache der Mengen übersetzt: „Gegeben sei eine Menge $\mathcal{E}$, deren Elemente wir Punkte nennen wollen“ und „Gegeben sei eine Menge $\mathcal{L}$, deren Elemente Teilmengen von $\mathcal{E}$ sind, welche wir Linien nennen“.

Beim Umgang mit Euklids Bezeichnungen ist aber Vorsicht geboten. Zum Beispiel meint er mit „Gleichheit“ von Strecken oder Winkeln in Wirklichkeit Kongruenz. Und mit Linien meint er Kurven und muss es folglich extra dazu sagen, wenn er von einer geraden Linie sprechen will. Er lässt auch unklar, welche Gebilde von Punkten er genau als Linien ansehen will, und kann nur Beispiele wie z.B. Kreise angeben, aber keine vollständige mathematische Definition. Da er von Winkeln spricht, wo sich zwei Linien schneiden, schwebt ihm sicher etwas der Art vor, was man heute regulär eingebettete differenzierbare Kurven nennt; insbesondere dürfen Linien keine Ecken oder Selbstüberschneidungen haben. Sie dürfen durch Anfangs- und Endpunkte begrenzt sein oder sich bis ins Unendliche erstrecken.

Einen Kreis definiert er nach heutigen Massstäben vollständig präzise als die Menge aller Punkte, die von einem gegebenen Punkt, genannt Mittelpunkt, einen gegebenen positiven Abstand, genannt Radius, haben. Er ist sich dessen bewusst, dass diese Definition alleine keinesfalls impliziert, dass der Mittelpunkt oder der Radius durch den Kreis eindeutig bestimmt sind, sondern dass dies erst später aus anderen Aussagen hergeleitet werden muss und kann.

Sodann gibt Euklid einige Postulate und Axiome an, welche beide als Axiome im heutigen Sinn gemeint sind. Zum Beispiel besagt sein Axiom „Was demselben gleich ist, ist auch untereinander gleich“, dass die Kongruenzrelation transitiv ist. Moderner als das geht es nicht. Andere seiner Axiome bedürfen aus heutiger Perspektive jedoch

Präzisierungen und Ergänzungen. Sein Axiom „Das Ganze ist grösser als der Teil“ ist eher eine Definition des Begriffs „grösser“ denn ein Axiom, weil dieser Begriff vorher noch nicht eingeführt worden war.

Der Hauptteil von Euklids Werk enthält Propositionen, also Lehrsätze, die er aus seinen Postulaten und Axiomen herleitet. Seine Behandlung ist aus heutiger Sicht zwar nicht ganz vollständig, aber insgesamt ein eindrücklicher Beweis für die Stärke der axiomatischen Methode.

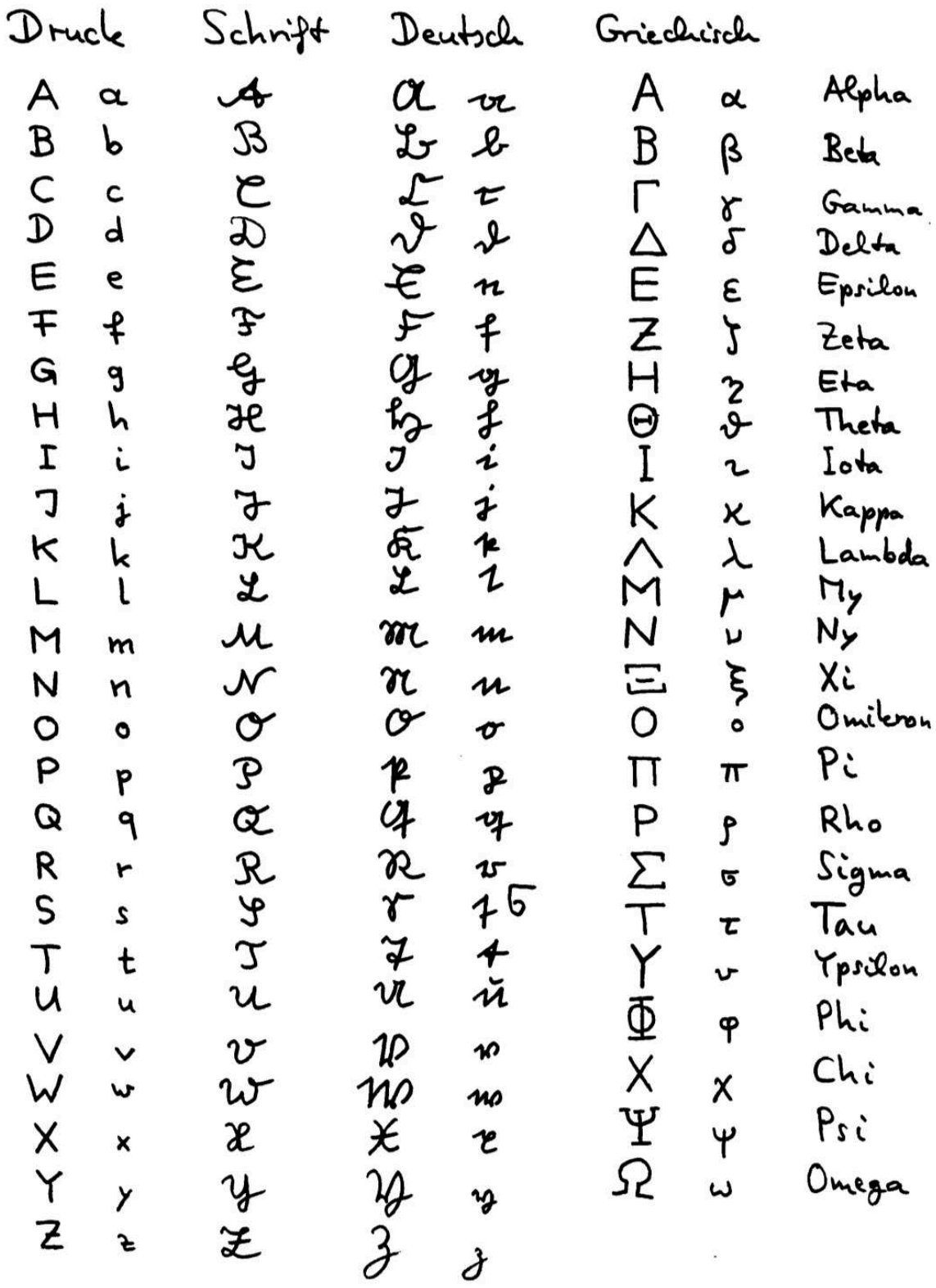
2 Grundlagen

2.1 Alphabete

Häufige Schriftsätze zur Bezeichnung mathematischer Objekte sind:

Blackboard	Fett		Druck		Schrift		Fraktur		Griechisch		
$\mathbb{A}$	$\mathbf{A}$	$\mathbf{a}$	A	a	$\mathcal{A}$	$\mathscr{A}$	$\mathfrak{A}$	$\mathfrak{a}$	A	α	Alpha
$\mathbb{B}$	$\mathbf{B}$	$\mathbf{b}$	B	b	$\mathcal{B}$	$\mathscr{B}$	$\mathfrak{B}$	$\mathfrak{b}$	B	β	Beta
$\mathbb{C}$	$\mathbf{C}$	$\mathbf{c}$	C	c	$\mathcal{C}$	$\mathscr{C}$	$\mathfrak{C}$	$\mathfrak{c}$	Γ	γ	Gamma
$\mathbb{D}$	$\mathbf{D}$	$\mathbf{d}$	D	d	$\mathcal{D}$	$\mathscr{D}$	$\mathfrak{D}$	$\mathfrak{d}$	Δ	δ	Delta
$\mathbb{E}$	$\mathbf{E}$	$\mathbf{e}$	E	e	$\mathcal{E}$	$\mathscr{E}$	$\mathfrak{E}$	$\mathfrak{e}$	E	ε	Epsilon
$\mathbb{F}$	$\mathbf{F}$	$\mathbf{f}$	F	f	$\mathcal{F}$	$\mathscr{F}$	$\mathfrak{F}$	$\mathfrak{f}$	Z	ζ	Zeta
$\mathbb{G}$	$\mathbf{G}$	$\mathbf{g}$	G	g	$\mathcal{G}$	$\mathscr{G}$	$\mathfrak{G}$	$\mathfrak{g}$	H	η	Eta
$\mathbb{H}$	$\mathbf{H}$	$\mathbf{h}$	H	h	$\mathcal{H}$	$\mathscr{H}$	$\mathfrak{H}$	$\mathfrak{h}$	Θ	ϑ	Theta
$\mathbb{I}$	$\mathbf{I}$	$\mathbf{i}$	I	i	$\mathcal{I}$	$\mathscr{I}$	$\mathfrak{I}$	$\mathfrak{i}$	I	ι	Iota
$\mathbb{J}$	$\mathbf{J}$	$\mathbf{j}$	J	j	$\mathcal{J}$	$\mathscr{J}$	$\mathfrak{J}$	$\mathfrak{j}$	K	κ	Kappa
$\mathbb{K}$	$\mathbf{K}$	$\mathbf{k}$	K	k	$\mathcal{K}$	$\mathscr{K}$	$\mathfrak{K}$	$\mathfrak{k}$	Λ	λ	Lambda
$\mathbb{L}$	$\mathbf{L}$	$\mathbf{l}$	L	l, ℓ	$\mathcal{L}$	$\mathscr{L}$	$\mathfrak{L}$	$\mathfrak{l}$	M	μ	My
$\mathbb{M}$	$\mathbf{M}$	$\mathbf{m}$	M	m	$\mathcal{M}$	$\mathscr{M}$	$\mathfrak{M}$	$\mathfrak{m}$	N	ν	Ny
$\mathbb{N}$	$\mathbf{N}$	$\mathbf{n}$	N	n	$\mathcal{N}$	$\mathscr{N}$	$\mathfrak{N}$	$\mathfrak{n}$	Ξ	ξ	Xi
$\mathbb{O}$	$\mathbf{O}$	$\mathbf{o}$	O	o	$\mathcal{O}$	$\mathscr{O}$	$\mathfrak{O}$	$\mathfrak{o}$	O	o	Omikron
$\mathbb{P}$	$\mathbf{P}$	$\mathbf{p}$	P	p	$\mathcal{P}$	$\mathscr{P}$	$\mathfrak{P}$	$\mathfrak{p}$	Π	π	Pi
$\mathbb{Q}$	$\mathbf{Q}$	$\mathbf{q}$	Q	q	$\mathcal{Q}$	$\mathscr{Q}$	$\mathfrak{Q}$	$\mathfrak{q}$	P	ρ	Rho
$\mathbb{R}$	$\mathbf{R}$	$\mathbf{r}$	R	r	$\mathcal{R}$	$\mathscr{R}$	$\mathfrak{R}$	$\mathfrak{r}$	Σ	σ	Sigma
$\mathbb{S}$	$\mathbf{S}$	$\mathbf{s}$	S	s	$\mathcal{S}$	$\mathscr{S}$	$\mathfrak{S}$	$\mathfrak{s}$	T	τ	Tau
$\mathbb{T}$	$\mathbf{T}$	$\mathbf{t}$	T	t	$\mathcal{T}$	$\mathscr{T}$	$\mathfrak{T}$	$\mathfrak{t}$	Y	υ	Ypsilon
$\mathbb{U}$	$\mathbf{U}$	$\mathbf{u}$	U	u	$\mathcal{U}$	$\mathscr{U}$	$\mathfrak{U}$	$\mathfrak{u}$	Φ	φ	Phi
$\mathbb{V}$	$\mathbf{V}$	$\mathbf{v}$	V	v	$\mathcal{V}$	$\mathscr{V}$	$\mathfrak{V}$	$\mathfrak{v}$	X	χ	Chi
$\mathbb{W}$	$\mathbf{W}$	$\mathbf{w}$	W	w	$\mathcal{W}$	$\mathscr{W}$	$\mathfrak{W}$	$\mathfrak{w}$	Ψ	ψ	Psi
$\mathbb{X}$	$\mathbf{X}$	$\mathbf{x}$	X	x	$\mathcal{X}$	$\mathscr{X}$	$\mathfrak{X}$	$\mathfrak{x}$	Ω	ω	Omega
$\mathbb{Y}$	$\mathbf{Y}$	$\mathbf{y}$	Y	y	$\mathcal{Y}$	$\mathscr{Y}$	$\mathfrak{Y}$	$\mathfrak{y}$			
$\mathbb{Z}$	$\mathbf{Z}$	$\mathbf{z}$	Z	z	$\mathcal{Z}$	$\mathscr{Z}$	$\mathfrak{Z}$	$\mathfrak{z}$			

Handgeschriebene Versionen der wichtigsten Schriftsätze sind:



2.2 Aussagenlogik

Definition: Eine *Aussage* ist eine wohlgeformte mathematische Behauptung, die entweder *wahr* oder *falsch* ist.

Definition: Mit *wahr* bezeichnen wir eine Aussage, die immer wahr ist; mit *falsch* eine, die immer falsch ist.

Definition: Gegebene Aussagen A und B kann man wie folgt zusammensetzen:
$A \wedge B$, gesprochen „A *und* B“, ist genau dann wahr, wenn A und B beide wahr sind.
$A \vee B$, gesprochen „A *oder* B“, ist genau dann wahr, wenn A oder B oder beide wahr sind.
$\neg A$, gesprochen „*nicht* A“, ist genau dann wahr, wenn A falsch ist.

Vorsicht: Die Aussage „*entweder* A *oder* B“, die genau dann wahr ist, wenn A oder B aber nicht beide wahr sind, muss man extra konstruieren durch $(A \vee B) \wedge \neg(A \wedge B)$.

Definition: Die Aussage $A \to B$, gesprochen „A *impliziert* B“ oder „*wenn* A *dann* B“, ist definiert als $(\neg A) \vee B$ und wahr genau dann, wenn A falsch oder B wahr ist.

Definition: Die Aussage $A \leftrightarrow B$, gesprochen „A *ist äquivalent zu* B“ oder „A *genau dann, wenn* B“ ist definiert als $(A \to B) \wedge (B \to A)$. Sie ist also wahr genau dann, wenn A und B denselben Wahrheitswert haben.

Vorsicht: Bei Implikation und Äquivalenz wird kein innerer Zusammenhang zwischen A und B gefordert.

Variante: Oft schreibt man dasselbe mit Doppelpfeilen $A \Rightarrow B$ bzw. $A \Leftrightarrow B$ anstatt einfacher Pfeile. Genaugenommen sollte man die Doppelpfeile aber für die metamathematische Implikation bzw. Äquivalenz verwenden, also wenn man tatsächlich inhaltliche Folgerungen durchführt in einem Beweis.

Zusammengesetzte Formeln: Die Negation $\neg$ hat höhere Bindungskraft als die Operatoren $\wedge$ und $\vee$, welche ihrerseits höhere Bindungskraft als die Operatoren $\to$ und $\leftrightarrow$ haben. (Letzteres wird aber nicht ganz einheitlich gehandhabt.) Um eine eindeutige Lesung der Formel zu erreichen sollte man Klammern setzen.

Grundregeln: Für alle Aussagen A, B, C gilt:

$$
\begin{array}{rcll}
A \wedge (B \wedge C) & \iff & (A \wedge B) \wedge C & \text{(Assoziativität)} \\
A \vee (B \vee C) & \iff & (A \vee B) \vee C & \text{(Assoziativität)} \\
A \wedge B & \iff & B \wedge A & \text{(Kommutativität)} \\
A \vee B & \iff & B \vee A & \text{(Kommutativität)} \\
A \wedge (B \vee C) & \iff & (A \wedge B) \vee (A \wedge C) & \text{(Distributivität)} \\
A \vee (B \wedge C) & \iff & (A \vee B) \wedge (A \vee C) & \text{(Distributivität)} \\
A \wedge A & \iff & A & \text{(Idempotenz)} \\
A \vee A & \iff & A & \text{(Idempotenz)} \\
A \wedge (A \vee B) & \iff & A & \text{(Absorption)} \\
A \vee (A \wedge B) & \iff & A & \text{(Absorption)} \\
A \wedge \text{wahr} & \iff & A & \\
A \wedge \text{falsch} & \iff & \text{falsch} & \\
A \vee \text{falsch} & \iff & A & \\
A \vee \text{wahr} & \iff & \text{wahr} & \\
\neg(A \wedge B) & \iff & \neg A \vee \neg B & \text{(de Morgan)} \\
\neg(A \vee B) & \iff & \neg A \wedge \neg B & \text{(de Morgan)} \\
\neg(\neg A) & \iff & A & \\
A \vee \neg A & \iff & \text{wahr} & \\
A \wedge \neg A & \iff & \text{falsch} &
\end{array}
$$

Aus diesen Grundregeln kann man alle anderen universell gültigen Äquivalenzen herleiten. Sie bilden somit ein *Axiomensystem der Aussagenlogik*. Manche dieser Regeln folgen schon aus anderen und könnten weggelassen werden, aber letztlich sind dies die Regeln, die man immer benutzt.

Weitere nützliche Regeln: Für alle Aussagen A, B gilt:

$$
\begin{array}{rcll}
A \vee (\neg A \wedge B) & \iff & A \vee B & \\
A \wedge (\neg A \vee B) & \iff & A \wedge B & \\
(A \to B) \wedge (\neg A \to B) & \iff & B & \text{(Fallunterscheidung)} \\
A \to B & \iff & \neg B \to \neg A & \text{(Kontrapositiv)} \\
A \leftrightarrow B & \iff & \neg B \leftrightarrow \neg A & \text{(Kontrapositiv)} \\
A \leftrightarrow B & \iff & B \leftrightarrow A &
\end{array}
$$

Proposition: Für jede natürliche Zahl $n \geqslant 1$ und beliebige Aussagen $A_1, \ldots, A_n$ gilt: Bei jeder möglichen Klammerung der (so noch nicht wohldefinierten) Formel $A_1 \wedge \ldots \wedge A_n$ ist der Wahrheitswert derselbe. Das Gleiche gilt für die Formel $A_1 \vee \ldots \vee A_n$. In diesen Fällen dürfen wir also doch auf Klammern verzichten.

2.3 Prädikatenlogik

In der Prädikatenlogik können Aussagen noch von einer oder mehreren Variablen abhängen. Sie erhalten dann erst einen definitiven Wahrheitswert, wenn konkrete Werte in die Variablen eingesetzt werden oder wenn die Variablen durch Quantoren gebunden werden.

Beispiel: Bezeichnet $A(x,y)$ die Relation „$x \geqslant y$“, so entsteht durch Einsetzen die variablenfreie Aussage $A(2,3) =$ „$2 \geqslant 3$“, die in diesem Fall falsch ist.

Definition: Das Symbol $\forall$ heisst *Allquantor*, das Symbol $\exists$ heisst *Existenzquantor*.
— Die Aussage $\forall x\ A(x)$ bedeutet „*für alle x gilt $A(x)$*“, oder „*für jedes x gilt $A(x)$*“, oder „*für beliebiges x gilt $A(x)$*“.
— Die Aussage $\exists x\ A(x)$ bedeutet „*es gibt ein x mit $A(x)$*“, oder „*es gibt mindestens ein x mit $A(x)$*“, oder „*es existiert x mit $A(x)$*“.
— Die Aussage $\exists! x\ A(x)$ bedeutet „*es gibt genau ein x mit $A(x)$*“.

Bemerkung: $\exists! x\ A(x)$ ist gleichbedeutend mit $\exists x \big(A(x) \wedge \forall y\colon (A(y) \to y = x)\big)$.

Bemerkung: Der Geltungsbereich jedes Quantors erstreckt sich von dem Quantor bis zum Ende der Formel oder zur nächsten schliessenden Klammer, welche zu einer vor dem Quantor stehenden öffnenden Klammer gehört. Der Quantor muss also immer *vor* der quantifizierten Aussage stehen.

Bemerkung: Einen Quantor kann man ansehen wie eine Variablendeklaration in einem Computerprogramm. Dort muss jede Variable deklariert werden bevor sie benutzt wird, und der Geltungsbereich erstreckt sich bis zum Ende des Unterprogramms, in dem die Deklaration steht.

Bemerkung: Meist quantifiziert man über Elemente einer gegebenen Menge X und kürzt dies wie folgt ab:

$$\begin{array}{rcl} \forall x \in X\colon A(x) & \text{bedeutet} & \forall x\colon (x \in X) \to A(x). \\ \exists x \in X\colon A(x) & \text{bedeutet} & \exists x\colon (x \in X) \wedge A(x). \\ \exists! x \in X\colon A(x) & \text{bedeutet} & \exists x \in X\colon \big(A(x) \wedge \forall y \in X\colon (A(y) \to y = x)\big). \end{array}$$

Vorsicht: Diese zusammengesetzten Aussagen sind auch dann sinnvoll, wenn X die leere Menge ist. Man muss dies sogar explizit zulassen, um Widerspruchsbeweise führen zu können. Für $X = \varnothing$ ist jede Aussage der Form „$\forall x \in X : A(x)$“ wahr, auch wenn $A(x)$ selbst absurd ist, und jede Aussage der Form „$\exists x \in X : A(x)$“ falsch, auch wenn $A(x)$ tautologisch richtig ist.

Bemerkung: Die Reihenfolge der Quantoren macht oft einen wichtigen Unterschied. So darf in der Aussage $\forall x\, \exists y\, A(x,y)$ das y, welches für gegebenenes x die Formel $A(x,y)$ erfüllt, von x abhängen, in der Aussage $\exists y\, \forall x\, A(x,y)$ dagegen nicht.

Beispiel: Die Aussage $\forall x \in \mathbb{Z}\ \exists y \in \mathbb{Z}\colon y > x$ drückt aus, dass es zu jeder ganzen Zahl x eine ganze Zahl y gibt, welche echt grösser als x ist. Diese Aussage ist wahr. Die Aussage $\exists y \in \mathbb{Z}\ \forall x \in \mathbb{Z}\colon y > x$ dagegen drückt aus, dass es eine ganze Zahl y gibt, welche echt grösser als jede ganze Zahl x ist. Diese Aussage ist falsch.

Bemerkung: In natürlichen Sprachen steht die Quantifizierung oft erst nach der quantifizierten Aussage, zum Beispiel wenn man sagt, dass „$x^2 \geqslant x$ ist für jede reelle Zahl $x \geqslant 1$“. Dort sorgen aber die (meist unbewussten) grammatikalischen Regeln für eine hoffentlich eindeutige Interpretation der Gesamtaussage. In der formalen Sprache der Prädikatenlogik gibt es einfachheitshalber nur die eine Regel, dass der Quantor *davor* stehen muss.

Vorsicht: Ignoriert man dies, so begeht man früher oder später Fehler, indem man Quantoren vergisst oder mehrdeutig formuliert wie zum Beispiel $\exists y\, A(x,y)\, \forall x$, wo nicht mehr klar ist, welche Reihenfolge der Quantoren gemeint ist.

Grundregeln: Einige Grundregeln zum Umformen zusammengesetzter Aussagen sind:

$$\begin{aligned}
\forall x\, \forall y\colon A(x,y) &\iff \forall y\, \forall x\colon A(x,y)\\
\exists x\, \exists y\colon A(x,y) &\iff \exists y\, \exists x\colon A(x,y)\\
\neg(\forall x\, A(x)) &\iff \exists x\colon \neg A(x)\\
\neg(\exists x\, A(x)) &\iff \forall x\colon \neg A(x)\\
\forall x\colon (A(x) \wedge B(x)) &\iff (\forall x\, A(x)) \wedge (\forall x\, B(x))\\
\exists x\colon (A(x) \vee B(x)) &\iff (\exists x\, A(x)) \vee (\exists x\, B(x))
\end{aligned}$$

Dagegen sind folgende Aussagen im allgemeinen nicht äquivalent:

$$\begin{aligned}
\forall x\, \exists y\, A(x,y) &\not\Longleftrightarrow \exists y\, \forall x\, A(x,y)\\
\forall x\colon (A(x) \vee B(x)) &\not\Longleftrightarrow (\forall x\, A(x)) \vee (\forall x\, B(x))\\
\exists x\colon (A(x) \wedge B(x)) &\not\Longleftrightarrow (\exists x\, A(x)) \wedge (\exists x\, B(x))
\end{aligned}$$

Abkü: Tritt derselbe Quantor mehrmals in Folge auf, so kürzt man oft ab wie z.B.:

$$\forall x_1, \ldots, x_n \in X\colon A(x_1, \ldots, x_n) \quad\iff\quad \forall x_1 \in X\ \ldots\ \forall x_n \in X\colon A(x_1, \ldots, x_n)$$

Bemerkung: Viele Aussagen aus dem realen Leben können in Prädikatenlogik übersetzt werden. Wenn zum Beispiel X die Menge aller Menschen bezeichnet, so bedeutet „Jeder Mensch hat einen Vater“ übersetzt „$\forall x \in X\ \exists y \in X$: (y ist Vater von x)“. Eine mögliche Übersetzung von „You can fool some people some time, but you can't fool all the people all the time“ lautet:

$$\big(\exists x \in X\ \exists t\colon (\text{you can fool } x \text{ at time } t)\big) \wedge \neg\big(\forall x \in X\ \forall t\colon (\text{you can fool } x \text{ at time } t)\big).$$

2.4 Mengen

Definition: Eine *Menge* ist eine Ansammlung von *Elementen*. Falls a ein Element von X ist, schreiben wir $a \in X$, andernfalls $a \notin X$.

Extensionalitätsaxiom: Zwei Mengen X und Y sind gleich genau dann, wenn sie dieselben Elemente besitzen, das heisst:

$$X = Y \iff \forall x\colon (x \in X) \leftrightarrow (x \in Y).$$

Konstruktion: Eine Menge kann man angeben durch Auflisten ihrer Elemente, wie zum Beispiel {Hans, Valeria, Bülent} oder $\{x_1, \ldots, x_n\}$. Dabei sind sowohl die Reihenfolge als auch etwaige Wiederholungen irrelevant; so gilt zum Beispiel

$$\{1, 5, 7, 2, 4, 7, 8, 1\} \;=\; \{1, 2, 4, 5, 7, 8\}.$$

Beispiel: Die *leere Menge* $\varnothing = \{\,\}$ ist die Menge ohne Elemente.

Aussonderungsaxiom: Eine Menge kann man angeben als die Menge derjenigen Elemente einer bereits bekannten Menge X, welche eine bestimmte Eigenschaft P besitzen, geschrieben:

$$\{x \in X \mid P(x)\} \quad \text{oder} \quad \{x \in X : P(x)\}.$$

Beispiel: Die Menge aller positiven reellen Zahlen $\mathbb{R}^{>0} = \{n \in \mathbb{R} \mid n > 0\}$.

Definition: Ein Ausdruck der Form $(x_1, \ldots, x_n)$ heisst *n-Tupel*. Zwei gegebene n-Tupel $(x_1, \ldots, x_n)$ und $(y_1, \ldots, y_n)$ sind gleich genau dann, wenn für alle $1 \leqslant i \leqslant n$ gilt $x_i = y_i$. Ein 2-Tupel heisst auch ein *Paar*, ein 3-Tupel ein *Tripel*, ein 4-Tupel ein *Quadrupel*, und so weiter.

Vorsicht: Bei Tupeln kommt es auf die Reihenfolge an; also gilt $(0, 1) \neq (1, 0)$, aber $\{0, 1\} = \{1, 0\}$.

Definition: Das *kartesische Produkt* endlich vieler Mengen $X_1, \ldots, X_n$ ist die Menge aller n-Tupel

$$X_1 \times \ldots \times X_n \;:=\; \bigtimes_{i=1}^{n} X_i \;:=\; \big\{(x_1, \ldots, x_n) \;\big|\; \forall 1 \leqslant i \leqslant n\colon x_i \in X_i\big\}.$$

Das kartesische Produkt von n Kopien derselben Menge wird abgekürzt mit

$$X^n \;:=\; X \times \ldots \times X.$$

Konstruktion: Für gegebene Mengen X und Y ist $\mathrm{Abb}(Y, X)$ oder X^Y die Menge aller Abbildungen $Y \to X$. Durch Zusatzbedingungen kann man auch die Menge aller Abbildungen mit gewissen gewünschten Eigenschaften bilden.

Konstruktion: Das *kartesische Produkt* einer beliebigen Kollektion von Mengen X_i für $i \in I$ ist die Menge

$$\bigtimes_{i\in I} X_i \;:=\; \big\{(x_i)_{i\in I} \in \mathrm{Abb}(I, \textstyle\bigcup_{i\in I} X_i) \;\big|\; \forall i \in I\colon x_i \in X_i\big\}.$$

Definition: Für gegebene Mengen X und Y sind

$$\begin{aligned} X \cup Y &:= \{x \mid x \in X \vee x \in Y\} && \text{die } \textit{Vereinigung} \text{ von } X \text{ und } Y, \\ X \cap Y &:= \{x \mid x \in X \wedge x \in Y\} && \text{der } \textit{Durchschnitt} \text{ von } X \text{ und } Y, \\ X \smallsetminus Y &:= \{x \mid x \in X \wedge x \notin Y\} && \text{die } \textit{Differenzmenge} \text{ von } X \text{ und } Y. \end{aligned}$$

Sicherheitshalber vereinbaren wir, dass diese drei binären Operatoren dieselbe Bindungskraft haben, also in zusammengesetzten Formeln überall Klammern erfordern.

Elementare Grundregeln: Die Operationen $\cup$ und $\cap$ erfüllen dieselben formalen Regeln wie die logischen Operationen $\vee$ und $\wedge$. Die leere Menge $\varnothing$ spielt dabei die Rolle der immer falschen Aussage *falsch*. Pendants zur logischen Verneinung und zur immer wahren Aussage gibt es allerdings nicht. Für alle Mengen X, Y, Z gilt also:

$$\begin{aligned} X \cap (Y \cap Z) &= (X \cap Y) \cap Z && \text{(Assoziativität)} \\ X \cup (Y \cup Z) &= (X \cup Y) \cup Z && \text{(Assoziativität)} \\ X \cap Y &= Y \cap X && \text{(Kommutativität)} \\ X \cup Y &= Y \cup X && \text{(Kommutativität)} \\ X \cap (Y \cup Z) &= (X \cap Y) \cup (X \cap Z) && \text{(Distributivität)} \\ X \cup (Y \cap Z) &= (X \cup Y) \cap (X \cup Z) && \text{(Distributivität)} \\ X \cap X &= X && \text{(Idempotenz)} \\ X \cup X &= X && \text{(Idempotenz)} \\ X \cap (X \cup Y) &= X && \text{(Absorption)} \\ X \cup (X \cap Y) &= X && \text{(Absorption)} \\ X \cap \varnothing &= \varnothing \\ X \cup \varnothing &= X \end{aligned}$$

Diese Grundregeln wie auch den folgenden Satz leitet man am einfachsten aus ihren Pendants in der Aussagenlogik her:

Proposition: Für jede natürliche Zahl $n \geqslant 1$ und beliebige Mengen $X_1, \ldots, X_n$ gilt: Jede mögliche Klammerung des (so noch nicht wohldefinierten) Ausdrucks $X_1 \cap \ldots \cap X_n$ definiert dieselbe Menge. Das Gleiche gilt für den Ausdruck $X_1 \cup \ldots \cup X_n$. In diesen Fällen dürfen wir also doch auf Klammern verzichten.

Definition: Eine Menge Y heisst *Teilmenge* einer Menge X, und wir schreiben $Y \subset X$ oder $X \supset Y$, wenn jedes Element von Y auch ein Element von X ist, das heisst, wenn gilt:

$$\forall x\colon x \in Y \to x \in X.$$

Falls zusätzlich $Y \neq X$ ist, so heisst Y eine *echte Teilmenge* von X und wir schreiben $Y \subsetneqq X$ oder $X \supsetneqq Y$.

Vorsicht: Manche Mathematiker benutzen das Symbol $\subset$ nur für echte Teilmengen und schreiben $\subseteq$, wenn sie die mögliche Gleichheit auch erlauben wollen. In dieser Vorlesung bleiben wir aber immer bei der oben beschriebenen Konvention.

Definition: Eine Menge X heisst *endlich der Kardinalität* $n \in \mathbb{Z}^{\geqslant 0}$ wenn es eine bijektive Abbildung $\{1, \ldots, n\} \to X$ gibt, das heisst, wenn man $X = \{x_1, \ldots, x_n\}$ schreiben kann mit paarweise verschiedenen $x_1, \ldots, x_n$. Andernfalls heisst X *unendlich* oder *der Kardinalität* ∞. Die Kardinalität von X ist eindeutig bestimmt und wird bezeichnet mit $|X|$ oder $\#(X)$ oder $\operatorname{card}(X)$.

Beispiel: Die leere Menge hat die Kardinalität $|\varnothing| = 0$.

Definition: Die *natürlichen Zahlen* sind die möglichen Kardinalitäten endlicher Mengen, also die Zahlen $0, 1, 2, \ldots$.

Vorsicht: Leider sind sich die Mathematiker uneinig: Für einige beginnen die natürlichen Zahlen mit 1. Um Missverständnisse zu vermeiden, benutze ich keine spezielle Notation für die Menge der natürlichen Zahlen, sondern schreibe jeweils präzise $\mathbb{Z}^{\geqslant 0}$ oder $\mathbb{Z}^{>0}$.

Vorsicht: Verschiedene Bezeichnungen oder Formeln garantieren nicht, dass Elemente einer Menge verschieden sind. Zum Beispiel ist

$$\big|\{\text{Heinz, Harun, Hrvoje}\}\big| \;=\; 1,$$

da es sich hier um verschiedene Vornamen derselben Person handelt, und

$$\big|\{n^2 - 6n + 10 \;\big|\; n = 1, 2, \ldots, 6\}\big| \;=\; \big|\{5, 2, 1, 2, 5, 10\}\big| \;=\; \big|\{1, 2, 5, 10\}\big| \;=\; 4.$$

Weitere Axiome: Ein praktikables *Axiomensystem für die Mengenlehre* erfordert noch eine Reihe weiterer Axiome, auf die wir hier aber nicht eingehen. Nur mit dem *Auswahlaxiom* oder dem dazu äquivalenten *Lemma von Zorn* werden wir zu tun haben. Ansonsten stellen wir uns auf den Standpunkt der *naiven Mengenlehre*, wie man sie im wesentlichen aus der Schule kennt.

Vorsicht: Man darf eine Menge nicht einfach nur durch Angabe einer Eigenschaft, also ohne eine umgebende Menge wie im Aussonderungsaxiom, oder ohne eine besondere Konstruktion wie die oben genannten, definieren. Denn erlaubte man dies, so könnte man mit dem *Russell-schen Paradoxon* einen Widerspruch herleiten, und die Theorie würde bedeutungslos.

2.5 Relationen

Definition: Eine *binäre Relation R auf einer Menge X* ist eine Teilmenge des kartesischen Produkts $X \times X$. Für $(x, y) \in R$ schreibt man meist kürzer $x\,R\,y$. Als Relationssymbole verwendet man oft abstrakte Symbole anstatt Buchstaben.

Definition: Eine binäre Relation R auf X heisst

$$\begin{aligned}
\textit{reflexiv}\text{, wenn gilt } & \forall x \in X\colon\ x\,R\,x.\\
\textit{symmetrisch}\text{, wenn gilt } & \forall x, y \in X\colon\ x\,R\,y \ \to\ y\,R\,x.\\
\textit{antisymmetrisch}\text{, wenn gilt } & \forall x, y \in X\colon\ (x\,R\,y \wedge y\,R\,x) \ \to\ x = y.\\
\textit{total}\text{, wenn gilt } & \forall x, y \in X\colon\ x\,R\,y \vee y\,R\,x.\\
\textit{transitiv}\text{, wenn gilt } & \forall x, y, z \in X\colon\ (x\,R\,y \wedge y\,R\,z) \ \to\ x\,R\,z.
\end{aligned}$$

Definition: Eine reflexive antisymmetrische transitive binäre Relation auf X heisst eine *Teilordnung* oder *Partialordnung auf X*. Eine totale Teilordnung heisst *Totalordnung*. Eine Menge zusammen mit einer Teil- bzw. Totalordnung heisst eine *teil-* bzw. *totalgeordnete Menge*.

Beispiel: Die Relation $\leqslant$ auf der Menge der reellen Zahlen $\mathbb{R}$ ist eine Totalordnung.

Beispiel: Die Relation $\subset$ auf einer beliebigen Menge von Mengen ist eine Teilordnung.

Definition: Sei $(X, \preceq)$ eine teilgeordnete Menge. Ein Element $x \in X$ mit

(a) $\forall y \in X\colon y \preceq x$ heisst *grösstes Element von X*.

(b) $\forall y \in X\colon x \preceq y \ \to\ x = y$ heisst ein *maximales Element von X*.

(c) $\forall y \in X\colon x \preceq y$ heisst *kleinstes Element von X*.

(d) $\forall y \in X\colon y \preceq x \ \to\ y = x$ heisst ein *minimales Element von X*.

Proposition: (a) Besitzt $(X, \preceq)$ ein grösstes (bzw. kleinstes) Element, so ist dieses eindeutig bestimmt und gleichzeitig maximal (bzw. minimal).

(b) In einer Totalordnung ist jedes maximale (bzw. minimale) Element ein grösstes (bzw. kleinstes) Element.

Vorsicht: Es kann sein, dass keine oder mehrere maximale (bzw. minimale) Elemente existieren. Zum Beispiel besitzt $(\mathbb{R}, \leqslant)$ kein maximales oder minimales Element. Gegenbeispiele zur Eindeutigkeit findet man unter anderen im folgenden Spezialfall:

Proposition: Jede nichtleere endliche teilgeordnete Menge besitzt ein maximales und ein minimales Element.

2.6 Äquivalenzrelationen

Definition: Eine reflexive symmetrische transitive binäre Relation heisst *Äquivalenzrelation.*

Sei $\sim$ eine Äquivalenzrelation auf X.

Definition: Die *Äquivalenzklasse* eines Elements $x \in X$ ist die Teilmenge

$$[x] \;:=\; \{y \in X : y \sim x\},$$

und jedes Element $y \in [x]$ heisst ein *Repräsentant von* $[x]$. Die Menge

$$X/\!\sim \;:=\; \{[x] : x \in X\}$$

aller Äquivalenzklassen heisst der *Quotient von X nach $\sim$.*

Proposition: Für je zwei Elemente $x, y \in X$ gilt

$$[x] \cap [y] \neq \varnothing \iff [x] = [y] \iff y \in [x] \iff x \sim y.$$

Beispiel: Die Gleichheit $=$ ist eine Äquivalenzrelation auf X mit den Äquivalenzklassen $[x] = \{x\}$ für alle $x \in X$.

Beispiel: Für beliebige ganze Zahlen a, b, n sagen wir *„a ist kongruent zu b modulo n“* und schreiben $a \equiv b \bmod (n)$, wenn $a - b$ ein Vielfaches von n ist. Fixieren wir $n > 0$, so definiert dies eine Äquivalenzrelation auf $\mathbb{Z}$. Ihre Äquivalenzklassen sind die Teilmengen $\{a + kn : k \in \mathbb{Z}\}$ für alle $a \in \mathbb{Z}$. Jede Äquivalenzklasse besitzt einen eindeutigen Repräsentanten a mit $0 \leqslant a < n$.

2.7 Gruppen

Definition: Eine *Gruppe* ist ein Tripel $(G, \circ, e)$ bestehend aus einer Menge G mit einer Abbildung

$$\circ: \quad G \times G \to G, \quad (a, b) \mapsto a \circ b$$

und einem ausgezeichneten Element $e \in G$, so dass gilt:

$$\begin{aligned} \forall a, b, c \in G: &\quad a \circ (b \circ c) = (a \circ b) \circ c && \text{(Assoziativität)} \\ \forall a \in G: &\quad e \circ a = a && \text{(Linksneutrales Element)} \\ \forall a \in G\ \exists a' \in G: &\quad a' \circ a = e && \text{(Linksinverses Element)} \end{aligned}$$

Die Gruppe heisst *kommutativ* oder *abelsch*, wenn zusätzlich gilt:

$$\forall a, b \in G: \quad a \circ b = b \circ a \qquad \text{(Kommutativität)}$$

Proposition: In jeder Gruppe $(G, \circ, e)$ gilt:

(a) Jedes linksneutrale Element e ist auch rechtsneutral, das heisst, es gilt $\forall a \in G$: $a \circ e = a$. Wir nennen e darum kurz *neutrales Element von* G.

(b) Jedes zu $a \in G$ linksinverse Element $a' \in G$ ist auch rechtsinvers, das heisst, es gilt $a \circ a' = e$. Wir nennen a' darum kurz *inverses Element zu* a.

(c) Das neutrale Element von G ist eindeutig bestimmt.

(d) Zu jedem $a \in G$ ist das inverse Element eindeutig bestimmt. Wir bezeichnen es mit a^{-1}.

(e) Für alle $a \in G$ gilt $(a^{-1})^{-1} = a$.

(f) Für alle $a, b \in G$ gilt $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$.

(g) Für alle $a, b \in G$ existiert ein eindeutiges $x \in G$ mit $a \circ x = b$.

(h) Für alle $a, b \in G$ existiert ein eindeutiges $y \in G$ mit $y \circ a = b$.

(i) Für alle $a, b, c \in G$ gilt $b = c \iff a \circ b = a \circ c$.

(j) Für alle $a, b, c \in G$ gilt $b = c \iff b \circ a = c \circ a$.

Proposition: Für jede natürliche Zahl $n \geqslant 1$ und für beliebige $a_1, \ldots, a_n \in G$ gilt: Bei jeder möglichen Klammerung der (a priori nicht wohldefinierten) Formel $a_1 \circ \ldots \circ a_n$ ist das Resultat gleich. Wir dürfen hier also doch auf Klammern verzichten.

Konvention: Bei allen Objekten der Mathematik, die aus einer Menge X mit Zusatzstrukturen bestehen, schreibt man meist nur kurz X für das ganze Objekt und sieht die Zusatzdaten als implizit mitgegeben an. Je nach Zusammenhang steht das Symbol X dann für die Menge X oder für das Tupel $(X, \ldots)$.

Konvention: Das neutrale Element einer multiplikativ geschriebenen Gruppe G bezeichnet man meist mit 1_G oder kurz 1, wobei sich jeweils aus dem Zusammenhang ergeben sollte, welches Einselement man meint.

Eine abelsche Gruppe schreibt man oft additiv, das heisst mit dem Operator $+$, dem neutralen Element 0_G oder 0, und dem inversen Element $-g$ zu g. Für $g+(-h)$ schreibt man dann auch kürzer $g - h$.

2.8 Körper, Ringe

Definition: Ein *Körper* ist ein Tupel $(K, +, \cdot, 0, 1)$ bestehend aus einer Menge K mit zwei Abbildungen

$$+ : \quad K \times K \to K, \quad (x, y) \mapsto x + y$$
$$\cdot : \quad K \times K \to K, \quad (x, y) \mapsto x \cdot y$$

und ausgezeichneten Elementen $0,\ 1 \in K$, so dass die folgenden *Körperaxiome* gelten:

(I)	$\forall x, y, z \in K\colon\ x + (y + z) = (x + y) + z$	(Assoziativität der Addition)
(II)	$\forall x, y \in K\colon\ x + y = y + x$	(Kommutativität der Addition)
(III)	$\forall x \in K\colon\ 0 + x = x$	(Neutrales Element der Addition)
(IV)	$\forall x \in K\ \exists x' \in K\colon\ x + x' = 0$	(Inverses Element der Addition)
(V)	$\forall x, y, z \in K\colon\ x \cdot (y \cdot z) = (x \cdot y) \cdot z$	(Assoziativität der Multiplikation)
(VI)	$\forall x \in K\colon\ 1 \cdot x = x$	(Neutrales Element der Multiplikation)
(VII)	$\forall x \in K \smallsetminus \{0\}\ \exists x' \in K\colon\ x' \cdot x = 1$	(Inverses Element der Multiplikation)
(VIII)	$\forall x, y, z \in K\colon \left\{ \begin{array}{l} x \cdot (y + z) = x \cdot y + x \cdot z \\ (y + z) \cdot x = y \cdot x + z \cdot x \end{array} \right\}$	(Distributivität)
(IX)	$1 \neq 0$	(Nichttrivialität)
(X)	$\forall x, y \in K\colon\ x \cdot y = y \cdot x$	(Kommutativität der Multiplikation)

Die Axiome (I) bis (IV) besagen, dass $(K, +, 0)$ eine abelsche Gruppe ist, genannt die *additive Gruppe von* K. Insbesondere ist das inverse Element $-x$ von x bezüglich der Addition eindeutig bestimmt.

Das inverse Element der Multiplikation zu $x \neq 0$ ist ebenfalls eindeutig bestimmt und wird mit $\frac{1}{x}$ bezeichnet. Für $x \cdot \frac{1}{y}$ schreibt man auch $\frac{x}{y}$. Für $x \cdot y$ schreibt man oft xy. Sei $K^\times := K \smallsetminus \{0\}$ die Menge aller von Null verschiedenen Elemente von K. Die Axiome (V) bis (VII) und (X) implizieren, dass $(K^\times, \cdot, 1)$ eine abelsche Gruppe ist, genannt die *multiplikative Gruppe von* K.

Beispiel: Die rationalen Zahlen $\mathbb{Q}$, die reellen Zahlen $\mathbb{R}$, die komplexen Zahlen $\mathbb{C}$, jeweils mit den üblichen Rechenoperationen und den üblichen neutralen Elementen.

Definition: Lässt man das Axiom (X) weg, so erhält man den Begriff des *Schiefkörpers* oder einer *Divisionsalgebra*. Ein kommutativer Schiefkörper ist also ein Körper.

Definition: Verlangt man nur die Axiome (I) bis (V) und (VIII), so erhält man den Begriff des *Rings*. Gilt zusätzlich (VI), so heisst der Ring *unitär* oder *Ring mit Eins*. Gilt zusätzlich (X), so heisst der Ring *kommutativ*.

Beispiel: Der Ring $\mathbb{Z}$ der ganzen Zahlen mit den üblichen $+$, $\cdot$, 0, 1.

Beispiel: Der Ring Z_n der ganzen Zahlen modulo n für jede ganze Zahl $n \geqslant 1$.

Beispiel: Der Körper $\mathbb{F}_p$ der ganzen Zahlen modulo p für jede Primzahl p. Insbesondere für $p = 2$ der Körper der binären Zahlen $\mathbb{F}_2 = \{0, 1\}$ mit den Operationen:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Definition: Für jedes Element x eines Rings R und jedes $n \in \mathbb{Z}$ definieren wir

$$n \cdot x := \begin{cases} x + \ldots + x \text{ mit } n \text{ Summanden} & \text{falls } n > 0, \\ 0 & \text{falls } n = 0, \\ -(x + \ldots + x) \text{ mit } |n| \text{ Summanden} & \text{falls } n < 0. \end{cases}$$

Proposition: Für alle $x, y \in R$ und alle $m, n \in \mathbb{Z}$ gilt:

$$\begin{aligned} (-n) \cdot x &= -(n \cdot x) \\ (m+n) \cdot x &= m \cdot x + n \cdot x \\ (m \cdot n) \cdot x &= m \cdot (n \cdot x) \\ m \cdot (x+y) &= m \cdot x + m \cdot y \\ m \cdot (x \cdot y) &= (m \cdot x) \cdot y \end{aligned}$$

Definition: Für jedes Element x eines Rings R und jedes $n \in \mathbb{Z}$ definieren wir

$$x^n := \begin{cases} x \cdots x \text{ mit } n \text{ Faktoren} & \text{falls } n > 0, \\ 1 & \text{falls } n = 0, \\ x^{-1} \cdots x^{-1} \text{ mit } |n| \text{ Faktoren} & \text{falls } n < 0 \text{ ist und } x^{-1} \text{ existiert.} \end{cases}$$

Proposition: Für alle $x, y \in K$ und alle $m, n \in \mathbb{Z}$ gilt, soweit definiert:

$$\begin{aligned} x^{m+n} &= x^m \cdot x^n \\ (x \cdot y)^m &= x^m \cdot y^m \\ x^{m \cdot n} &= (x^m)^n \end{aligned}$$

2.9 Vollständige Induktion

Beweis durch Induktion:

Sei $n_0 \in \mathbb{Z}$, und sei $A(n)$ eine für jede natürliche Zahl $n \geqslant n_0$ wohldefinierte Aussage.

Ziel: Beweise die Aussage $\forall n \geqslant n_0 \colon A(n)$ durch Induktion.

Methode 1: (Grundform)
Induktionsverankerung: Beweise $A(n_0)$ direkt.
Sodann nimm ein beliebiges $n \geqslant n_0$.
Induktionsvoraussetzung: Es gelte $A(n)$.
Induktionsschritt: Beweise $A(n+1)$ unter Benützung der Induktionsvoraussetzung.

Methode 2: (stärkere Induktionsvoraussetzung, integrierter Induktionsanfang)
Nimm ein beliebiges $n \geqslant n_0$.
Induktionsvoraussetzung: Für alle n' mit $n_0 \leqslant n' < n$ gelte $A(n')$.
Induktionsschritt: Beweise $A(n)$ unter Benützung der Induktionsvoraussetzung.

Methode 3: (Widerspruchsbeweis)
Nimm an, die Aussage $\forall n \geqslant n_0 \colon A(n)$ gelte nicht. Dann existiert $n \geqslant n_0$ mit $\neg A(n)$. Insbesondere existiert dann ein kleinstes solches n, das heisst, ein $n \geqslant n_0$ mit $\neg A(n)$ und $\forall n' \colon n_0 \leqslant n' < n \longrightarrow A(n')$. Aus diesen Aussagen leite einen Widerspruch her.

Proposition: Jede nichtleere Teilmenge von $\{n \in \mathbb{Z} \mid n \geqslant n_0\}$ besitzt ein eindeutiges kleinstes Element.

„Proposition“ (nicht ganz ernst gemeint): Jede natürliche Zahl ist interessant.

Varianten: Analog beweist man eine Aussage $A(n)$ für alle ganzen Zahlen $n \leqslant n_0$ durch *absteigende Induktion*, was äquivalent dazu ist, die Aussage $A(-m)$ für alle $m \geqslant -n_0$ durch aufsteigende Induktion zu beweisen. Oft ist eine Aussage auch nur für alle ganzen Zahlen in einem endlichen Intervall $n_0 \leqslant n \leqslant n_1$ zu beweisen, wofür je nach Situation aufsteigende oder absteigende Induktion günstiger sein kann.

Definition durch Induktion (oder Rekursion):

Ziel: Definiere ein mathematisches Objekt oder eine Eigenschaft B_n für alle $n \geqslant n_0$ durch Induktion.

Methode 1:
Induktionsverankerung: Definiere B_{n_0} direkt.
Sodann nimm ein beliebiges $n \geqslant n_0$.
Induktionsvoraussetzung: Sei B_n bereits definiert.
Induktionsschritt: Definiere B_{n+1} unter Benützung von B_n.

Methode 2:
Nimm ein beliebiges $n \geqslant n_0$.
Induktionsvoraussetzung: Für alle n' mit $n_0 \leqslant n' < n$ sei $B_{n'}$ bereits definiert.
Induktionsschritt: Definiere B_n unter Benützung von $B_{n'}$ für alle $n_0 \leqslant n' < n$.

Beispiel: Für jede natürliche Zahl n ist $n!$, sprich n *Fakultät*, definiert durch

$$n! \;:=\; \begin{cases} 1 & \text{für } n = 0, \\ (n-1)! \cdot n & \text{für alle } n > 0. \end{cases}$$

Beispiel: Die *Fibonacci-Zahlen* F_n sind für alle $n \geqslant 0$ definiert durch

$$F_n \;:=\; \begin{cases} 0 & \text{falls } n = 0, \\ 1 & \text{falls } n = 1, \\ F_{n-1} + F_{n-2} & \text{für alle } n \geqslant 2. \end{cases}$$

Beispiel: Sei $*$ eine binäre Operation, welche das Assoziativgesetz $\forall a, b, c\colon (a * b) * c = a * (b * c)$ erfüllt. Für alle $n \geqslant 1$ und alle geeigneten $a_1, \ldots, a_n$ wählen wir eine spezifische Klammerung der Formel $a_1 * \ldots * a_n$ durch die induktive Definition

$$a_1 * \ldots * a_n \;:=\; \begin{cases} a_1 & \text{für } n = 1, \\ (a_1 * \ldots * a_{n-1}) * a_n & \text{für alle } n > 1. \end{cases}$$

Dann zeigen wir ebenfalls durch Induktion:

Proposition: Für alle $1 \leqslant m < n$ und $a_1, \ldots, a_n$ gilt

$$a_1 * \ldots * a_n \;=\; (a_1 * \ldots * a_m) * (a_{m+1} * \ldots * a_n).$$

Folge: Für jede natürliche Zahl $n \geqslant 1$ und für beliebige $a_1, \ldots, a_n$ ist dann bei jeder möglichen Klammerung der (a priori nicht wohldefinierten) Formel $a_1 * \ldots * a_n$ das Resultat gleich. Wir dürfen hier also doch auf Klammern verzichten.

Pünktchen:

Jede Aufzählung oder Formel mit Pünktchen ergibt nur einen Sinn, wenn das Bildungsprinzip aus dem Zusammenhang eindeutig ersichtlich ist. Dabei sollen die Leser nie raten müssen, was gemeint ist.

Meist hat die Aufzählung eine natürliche Laufvariable i, welche alle ganzen Zahlen in einem gegebenen Intervall $m \leqslant i \leqslant n$ durchläuft, und das i-te Glied ist durch eine pünktchenfreie Formel gegeben. Oft ist es dann günstig, diese Formel für das i-te Glied mit anzugeben, wie zum Beispiel in

$$a_1, \ldots, a_i, \ldots, a_n$$

$$1, 4, 9, \ldots, i^2, \ldots, n^2.$$

$$0, 1, 3, 6, 10, \ldots, \tfrac{i(i+1)}{2}, \ldots$$

Am besten werden Missverständnisse aber dadurch vermieden, dass man Pünktchen ganz durch Quantifizierung und/oder Induktion ersetzt.

Beispiele zur Übersetzung:

Mit Pünktchen	Ohne Pünktchen
Seien $x_1, \ldots, x_n \in X$.	Seien $x_i \in X$ für alle $1 \leqslant i \leqslant n$.
Das Tupel $(a_1, \ldots, a_n)$	Das Tupel $(a_i)_{i=1}^n$ oder $(a_i)_{1 \leqslant i \leqslant n}$
$(\ldots((a_1 * a_2) * a_3) * \ldots * a_{n-1}) * a_n$	$B_1 := a_1$ und $B_m := B_{m-1} * a_m$ für alle $2 \leqslant m \leqslant n$
$n! := 1 \cdot 2 \cdots n$	$0! := 1$ und $n! := (n-1)! \cdot n$ für alle $n > 0$.
$x^n := x \cdots x$ mit n Faktoren	$x^1 := x$ und $x^n := x \cdot x^{n-1}$ für alle $n \geqslant 2$.

2.10 Summen

Sei K ein Körper.

Definition: Gegeben seien $p, q \in \mathbb{Z}$ und $a_i \in K$ für alle $i \in \mathbb{Z}$ mit $p \leqslant k \leqslant q$. Die *Summe* dieser Elemente kürzen wir ab wie folgt:

$$\sum_{i=p}^{q} a_i \;:=\; \begin{cases} a_p + a_{p+1} + \ldots + a_q & \text{falls } p \leqslant q, \\ 0 & \text{sonst.} \end{cases}$$

Ohne Pünktchen geschrieben bedeutet das:

$$\sum_{i=p}^{q} a_i \;:=\; \begin{cases} 0 & \text{falls } p > q, \\ a_p & \text{falls } p = q, \\ \left(\sum_{i=p}^{q-1} a_i\right) + a_q & \text{falls } p < q. \end{cases}$$

Bemerkung: Der Wert der leeren Summe ist also das neutrale Element 0 der Addition. Dies garantiert, dass die folgende erste Grundregel über das Aufspalten einer Summe in zwei Teilsummen auch für leere Teilsummen gilt.

Proposition: Für alle $p, q, r \in \mathbb{Z}$ und $a_i, b_i, a \in K$ gilt:

$$\begin{aligned} \sum_{i=p}^{r} a_i &= \sum_{i=p}^{q} a_i + \sum_{i=q+1}^{r} a_i \qquad \text{falls } p-1 \leqslant q \leqslant r, \\ \sum_{i=p}^{q} a_i &= \sum_{j=p}^{q} a_j \qquad \text{(Umbenennen der Laufvariablen)}, \\ \sum_{i=p}^{q} (a_i + b_i) &= \sum_{i=p}^{q} a_i + \sum_{i=p}^{q} b_i, \\ a \cdot \sum_{i=p}^{q} b_i &= \sum_{i=p}^{q} a b_i. \end{aligned}$$

Variante: Eine Summe über eine endliche Menge ganzer Zahlen I schreiben wir in der Form:

$$\sum_{i \in I} a_i$$

Alternativ schreiben wir nur die Bedingungen unter das Summenzeichen.

Proposition:

$$\sum_{i \in I_1 \cup I_2} a_i \;=\; \sum_{i \in I_1} a_i + \sum_{i \in I_2} a_i - \sum_{i \in I_1 \cap I_2} a_i.$$

Beispiel: Es gilt

$$\Big(\sum_{p\leqslant i\leqslant q} a_i\Big)^2 \;=\; \sum_{p\leqslant i\leqslant q} a_i^2 \;+\; 2\cdot\sum_{p\leqslant i<j\leqslant q} a_i a_j.$$

Satz: Für alle $n \in \mathbb{Z}^{\geqslant 0}$ und alle $a, b \in K$ gilt die allgemeine *binomische Formel*

$$(a+b)^n \;=\; \sum_{k=0}^{n} \binom{n}{k} \cdot a^{n-k} \cdot b^k.$$

Variante: Gegeben seien eine beliebige Menge I und ein Element $a_i \in K$ für jedes $i \in I$. Wir fordern $a_i = 0$ für fast alle $i \in I$, das heisst, es existiert eine endliche Teilmenge $I_0 \subset I$ mit $a_i = 0$ für alle $i \in I \smallsetminus I_0$. Dann definieren wir

$$\sum_{i\in I}{}' a_i \;:=\; \sum_{i\in I_0} a_i.$$

Dies ist unabhängig von der gewählten Teilmenge I_0 und daher wohldefiniert. Der Strich $'$ am Summenzeichen bezeichnet die Forderung $a_i = 0$ für fast alle $i \in I$ und dient zur Unterscheidung von einer unendlichen Reihe im Kontext der Analysis.

Die Summe in diesem verallgemeinerten Sinn erfüllt dieselben Grundregeln wie oben.

2.11 Produkte

Definition: Das *Produkt* endlich vieler $a_i \in K$ kürzen wir ab wie folgt:

$$\prod_{i=p}^{q} a_i \;:=\; \prod_{p\leqslant i\leqslant q} a_i \;:=\; \begin{cases} a_p a_{p+1} \cdots a_q & \text{für } p \leqslant q, \\ 1 & \text{sonst.} \end{cases}$$

Ohne Pünktchen geschrieben bedeutet das:

$$\prod_{i=p}^{q} a_i \;:=\; \begin{cases} 1 & \text{falls } p > q, \\ a_p & \text{falls } p = q, \\ \left(\prod_{i=p}^{q-1} a_i\right) \cdot a_q & \text{falls } p < q. \end{cases}$$

Bemerkung: Der Wert des leeren Produkts ist also das neutrale Element 1 der Multiplikation. Dies ermöglicht wieder das Aufspalten eines Produkts wie folgt.

Proposition: Für alle $p, q, r \in \mathbb{Z}$ und $a_i, b_i \in K$ und $n \in \mathbb{Z}^{\geqslant 0}$ gilt:

$$\begin{aligned} \prod_{i=p}^{r} a_i &= \prod_{i=p}^{q} a_i \cdot \prod_{i=q+1}^{r} a_i \qquad \text{falls } p-1 \leqslant q \leqslant r, \\ \prod_{i=p}^{q} a_i &= \prod_{j=p}^{q} a_j \qquad \text{(Umbenennen der Laufvariablen)}, \\ \prod_{i=p}^{q} a_i b_i &= \prod_{i=p}^{q} a_i \cdot \prod_{i=p}^{q} b_i \\ \Big(\prod_{i=p}^{q} a_i\Big)^n &= \prod_{i=p}^{q} a_i^n. \end{aligned}$$

Variante: In Analogie zu Summen schreiben wir ein Produkt über eine endliche Menge I in der Form:

$$\prod_{i\in I} a_i$$

oder schreiben nur die Bedingungen unter das Summenzeichen. Für eine beliebige Menge I und Elemente $a_i \in K$ mit $a_i = 1$ für fast alle $i \in I$ definieren wir

$$\prod_{i\in I}{}' a_i \;:=\; \prod_{i\in I_0} a_i$$

für jede endliche Teilmenge $I_0 \subset I$ mit $a_i = 1$ für alle $i \in I \smallsetminus I_0$. Dies ist unabhängig von der gewählten Teilmenge I_0 und daher wohldefiniert.

Das Produkt in diesem verallgemeinerten Sinn erfüllt dieselben Grundregeln wie oben.

3 Matrizen und Lineare Gleichungssysteme

3.1 Definitionen

Sei K ein Körper, und seien m, n, ℓ natürliche Zahlen.

Definition: Eine *Matrix mit m Zeilen und n Spalten*, oder kurz eine $m \times n$*-Matrix, über* K ist ein Ausdruck $A = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ mit Koeffizienten $a_{ij} \in K$, ausführlich geschrieben als rechteckiges Schema

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

Der erste Index i in $(\dots)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ bezeichnet stets die Zeile, der zweite j die Spalte.

Definition: Eine $m \times m$-Matrix heisst *quadratisch*. Eine $m \times 1$-Matrix heisst ein *Spaltenvektor*, eine $1 \times n$-Matrix ein *Zeilenvektor*.

Beispiel: Die *Nullmatrix* $O_{m,n} := \big(0\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ mit allen Einträgen gleich 0.

Beispiel: Die (quadratische) *Einheitsmatrix* $I_m := \big(\delta_{ij}\big)_{1\leqslant i,j\leqslant m}$ mit der *Kronecker-Deltafunktion*

$$\delta_{ij} := \begin{cases} 1 & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$$

das heisst, mit allen Diagonaleinträgen gleich 1 und allen übrigen Einträgen gleich 0.

Definition: Das *Produkt* einer $m\times n$-Matrix $A = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ mit einem Element $\lambda \in K$ ist die $m \times n$-Matrix

$$\lambda \cdot A \;:=\; A \cdot \lambda \;:=\; \big(\lambda a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}.$$

Definition: Die *Summe* zweier $m \times n$-Matrizen $A = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ und $B = \big(b_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ ist die $m \times n$-Matrix

$$A + B \;:=\; \big(a_{ij} + b_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}.$$

Wir immer bei Addition kürzen wir ab $-B := (-1) \cdot B$ und $A - B := A + (-1) \cdot B$.

Definition: Das *Produkt* einer $m \times n$-Matrix $A = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ mit einer $n \times \ell$-Matrix $B = \big(b_{jk}\big)_{\substack{1\leqslant j\leqslant n\\ 1\leqslant k\leqslant \ell}}$ ist die $m \times \ell$-Matrix

$$A \cdot B \;:=\; \Big(\sum_{j=1}^{n} a_{ij} b_{jk}\Big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant k\leqslant \ell}}.$$

3.2 Grundeigenschaften

Proposition: Für alle Matrizen A, B, C über K und alle λ, $\mu \in K$ gilt, sofern die Matrizen die für die jeweilige Formel richtige Grösse besitzen:

$$
\begin{array}{rcll}
A + (B + C) & = & (A + B) + C & \text{(Assoziativität der Addition)} \\
A + B & = & B + A & \text{(Kommutativität der Addition)} \\
O_{m,n} + A & = & A & \text{(Neutrales Element der Addition)} \\
A + (-1) \cdot A & = & O_{m,n} & \text{(Inverses Element der Addition)} \\
\\
\lambda \cdot (A + B) & = & \lambda \cdot A + \lambda \cdot B & \text{(Linksdistributivität der skalaren Multiplikation)} \\
(\lambda + \mu) \cdot A & = & \lambda \cdot A + \mu \cdot A & \text{(Rechtsdistributivität der skalaren Multiplikation)} \\
\lambda \cdot (\mu \cdot A) & = & (\lambda \cdot \mu) \cdot A & \text{(Assoziativität der skalaren Multiplikation)} \\
1 \cdot A & = & A & \text{(Einselement und skalare Multiplikation)} \\
0 \cdot A & = & O_{m,n} & \text{(Nullelement und skalare Multiplikation)} \\
\\
A \cdot (B \cdot C) & = & (A \cdot B) \cdot C & \text{(Assoziativität der Multiplikation)} \\
A \cdot (B + C) & = & A \cdot B + A \cdot C & \text{(Linksdistributivität der Multiplikation)} \\
(A + B) \cdot C & = & A \cdot C + B \cdot C & \text{(Rechtsdistributivität der Multiplikation)} \\
I_m \cdot A & = & A & \text{(Linksneutrales Element der Multiplikation)} \\
A \cdot I_n & = & A & \text{(Rechtsneutrales Element der Multiplikation)} \\
\\
\lambda \cdot (A \cdot B) & = & (\lambda \cdot A) \cdot B & \text{(Assoziativität für gemischte Multiplikation)} \\
\lambda \cdot (A \cdot B) & = & A \cdot (\lambda \cdot B) & \text{(Assoziativität für gemischte Multiplikation)}
\end{array}
$$

Insbesondere ist die Menge aller $m \times n$-Matrizen über K zusammen mit der Addition und dem neutralen Element $O_{m,n}$ eine abelsche Gruppe.

Die Menge aller $m \times m$-Matrizen über K zusammen mit der Addition und Multiplikation und dem Nullelement $O_{m,m}$ und dem Einselement I_m ist ein unitärer Ring.

Vorsicht: Für $m \geqslant 2$ ist dieser Ring weder kommutativ noch ein Schiefkörper.

3.3 Invertierbare Matrizen

Definition: Eine $m \times m$-Matrix A, zu der eine $m \times m$-Matrix A' mit $A \cdot A' = A' \cdot A = I_m$ existiert, heisst *invertierbar*. Die betreffende Matrix A' heisst dann *Inverse von* A und wird mit A^{-1} bezeichnet. Die Menge aller invertierbaren $m \times m$-Matrizen über K wird mit $\mathrm{GL}_m(K)$ bezeichnet.

Proposition: Die Menge $\mathrm{GL}_m(K)$ zusammen mit der Matrixmultiplikation und dem Einselement I_m bildet eine Gruppe. Insbesondere gilt für alle $m \times m$-Matrizen A und B:

(a) Ist A invertierbar, so ist die Inverse A^{-1} eindeutig bestimmt.
(b) Ist A invertierbar, so ist auch A^{-1} invertierbar mit $(A^{-1})^{-1} = A$.
(c) Sind A und B invertierbar, so ist $A \cdot B$ invertierbar und $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.
(d) Ist A invertierbar, so ist B invertierbar genau dann wenn $A \cdot B$ invertierbar ist genau dann wenn $B \cdot A$ invertierbar ist.

Proposition: Für jede invertierbare $m \times m$-Matrix A gilt:

(a) Für alle $m \times n$-Matrizen B, C gilt $A \cdot B = C$ genau dann wenn $B = A^{-1} \cdot C$ ist.
(b) Für alle $n \times m$-Matrizen B, C gilt $B \cdot A = C$ genau dann wenn $B = C \cdot A^{-1}$ ist.

Proposition: Für jede $m \times m$-Matrix A über K sind äquivalent:

(a) Die Matrix A ist invertierbar.
(b) Es existiert eine $m \times m$-Matrix A' über K mit $A' \cdot A = I_m$ (Linksinverse).
(c) Es existiert eine $m \times m$-Matrix A'' über K mit $A \cdot A'' = I_m$ (Rechtsinverse).

In jedem der Fälle (b) und (c) ist die Matrix A' die eindeutig bestimmte Inverse A^{-1}.

Vorsicht: Dies beweisen wir erst nach der Dreieckszerlegung.

3.4 Transposition

Definition: Die *Transponierte* einer $m \times n$-Matrix $A = \left(a_{ij}\right)_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$ ist die durch Vertauschen der Zeilenindizes mit den Spaltenindizes entstehende $n \times m$-Matrix

$$A^T := \left(a_{ij}\right)_{\substack{1 \leqslant j \leqslant n \\ 1 \leqslant i \leqslant m}} = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

Proposition: Für alle Matrizen A, B über K und alle $\lambda \in K$ gilt, sofern die Matrizen die für die jeweilige Formel richtige Grösse besitzen:

$$\begin{aligned} (A^T)^T &= A \\ (A+B)^T &= A^T + B^T \\ (\lambda \cdot A)^T &= \lambda \cdot A^T \\ (A \cdot B)^T &= B^T \cdot A^T \\ O_{m,n}^T &= O_{n,m} \\ I_m^T &= I_m \end{aligned}$$

Ausserdem ist A invertierbar genau dann, wenn A^T invertierbar ist, und dann gilt

$$(A^T)^{-1} = (A^{-1})^T.$$

3.5 Lineare Gleichungssysteme

Sei K ein Körper.

Definition: Ein System von Gleichungen $\sum_{j=1}^{n} a_{ij}x_j = b_i$ für alle $1 \leqslant i \leqslant m$, das heisst

$$\begin{array}{ccccc} a_{11}x_1 & + \ldots + & a_{1n}x_1 & = & b_1 \\ \vdots & & \vdots & & \vdots \\ a_{11}x_1 & + \ldots + & a_{mn}x_n & = & b_m \end{array}$$

für gegebene $a_{ij}, b_i \in K$ und zu bestimmenden Variablen x_j, heisst *lineares Gleichungssystem* (kurz LGS) *über* K. Sind alle $b_j = 0$, so heisst das Gleichungssystem *homogen*.

Fakt: Jedes lineare Gleichungssystem kann man schreiben in Matrixform

$$A \cdot x = b$$

für die Matrix $A = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m \\ 1\leqslant j\leqslant n}}$ und die Spaltenvektoren $b = (b_i)_{1\leqslant i\leqslant m}$ und $x = (x_i)_{1\leqslant j\leqslant n}$. Insgesamt stellt man das LGS oft durch die zusammengesetzte $m \times (n+1)$-Matrix dar:

$$(A, b) := \left(\begin{array}{ccc|c} a_{11} & \ldots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} & b_m \end{array}\right)$$

Definition: *Elementare Zeilenumformungen* eines linearen Gleichungssystems oder einer Matrix über K sind:

(a) das Addieren von $\lambda \in K$ mal einer Zeile zu einer anderen,
(b) das Multiplizieren einer Zeile mit $\lambda \in K^\times$,
(c) das Vertauschen zweier Zeilen.

Die Wirkung jeder elementaren Zeilenumformung auf ein Gleichungssystem $Ax = b$ entspricht der Wirkung derselben elementaren Zeilenumformung auf die Matrix (A, b).

Fakt: Jede elementare Zeilenumformung ist umkehrbar, nämlich in den betreffenden Fällen durch

(a) das Addieren von $-\lambda$ mal derselben Zeile zu derselben anderen,
(b) das Multiplizieren derselben Zeile mit λ^{-1},
(c) das nochmalige Vertauschen derselben Zeilen.

Für das Gleichungssystem insgesamt erhalten wir daher eine Äquivalenzumformung.

Definition: Ein Gleichungssystem oder eine Matrix heisst *in Zeilenstufenform*, wenn die von Null verschiedenen Terme in jeder Zeile echt später beginnen als in der Zeile davor.

Satz: *(Gauss-Elimination)* Jedes Gleichungssystem und jede Matrix lässt sich durch eine Folge elementarer Zeilenumformungen in Zeilenstufenform bringen.

Beweis: Später zusammen mit der Dreieckszerlegung.

Definition: Eine Matrix mit genau einem Eintrag 1 und allen übrigen Einträgen 0 heisst *Elementarmatrix*. Genauer betrachten wir für fest gegebene $m, n \geqslant 1$ und alle $1 \leqslant i \leqslant m$ und $1 \leqslant j \leqslant n$ die Elementarmatrix $E_{ij} := \big(\delta_{ii'}\delta_{jj'}\big)_{\substack{1\leqslant i'\leqslant m\\ 1\leqslant j'\leqslant n}}$.

Proposition: Seien A eine $m \times n$-Matrix und $\lambda \in K$, und $1 \leqslant i, j \leqslant m$ mit $i \neq j$.

(a) Addieren von λ mal der j-ten Zeile zur i-ten ergibt die Matrix $(I_m + \lambda E_{ij}) \cdot A$.
(b) Die Matrix $I_m + \lambda E_{ij}$ ist invertierbar mit der Inversen $I_m - \lambda E_{ij}$.

Definition: Eine $m \times m$-Matrix $A = \big(a_{ij}\big)_{1\leqslant i,j\leqslant m}$ mit $a_{ij} = 0$ für alle $i \neq j$ heisst *Diagonalmatrix*.

Proposition: Seien A eine $m \times n$-Matrix und $\lambda \in K$ und $1 \leqslant i \leqslant m$. Sei D die $m \times m$-Diagonalmatrix mit Eintrag λ an der Stelle (i, i) und allen übrigen Diagonaleinträgen 1.

(a) Multiplizieren der i-ten Zeile von A mit λ ergibt die Matrix $D \cdot A$.
(b) Die Matrix D ist invertierbar; ihre Inverse ist die Diagonalmatrix mit Eintrag λ^{-1} an der Stelle (i, i) und allen übrigen Diagonaleinträgen 1.

Definition: Eine $m \times m$-Matrix, in der jede Zeile und jede Spalte genau eine 1 und sonst nur Einträge 0 besitzt, heisst *Permutationsmatrix*. In Formeln: $P = \big(p_{ij}\big)_{1\leqslant i,j\leqslant m}$ ist eine Permutationsmatrix wenn gilt:

$$\begin{array}{l}\forall i\ \exists j\colon p_{ij} = 1\ \wedge\ \forall j' \neq j\colon p_{ij'} = 0, \quad \text{und}\\ \forall j\ \exists i\colon p_{ij} = 1\ \wedge\ \forall i' \neq i\colon p_{i'j} = 0.\end{array}$$

Proposition: Jede Permutationsmatrix P ist invertierbar mit $P^{-1} = P^T$.

Proposition: Seien A eine $m \times n$-Matrix und $1 \leqslant i, j \leqslant m$ mit $i \neq j$. Sei P die $m \times m$-Permutationsmatrix, die der Einheitsmatrix gleicht ausser dass die Einträge in Position (i, i) und (j, j) gleich 0 und die Einträge in Position (i, j) und (j, i) gleich 1 sind.

(a) Vertauschen der i-ten und j-ten Zeile von A ergibt die Matrix $P \cdot A$.
(b) Die Matrix P ist invertierbar mit der Inversen $P^{-1} = P$.

Satz: Jede Folge elementarer Zeilenoperationen wird durch Multiplikation von links mit einer geeigneten invertierbaren Matrix repräsentiert.

Satz: Jede Matrix lässt sich durch eine Folge elementarer Zeilenoperationen und Vertauschen von Spalten in die folgende Blockmatrixform bringen für ein geeignetes k:

$$\left(\begin{array}{c|c} I_k & C \\ \hline O & O \end{array}\right)$$

Bemerkung: Vertauschen von Spalten in der zu der linken Seite eines linearen Gleichungssystems assoziierten Matrix bedeutet Vertauschen von Variablen. Sobald man die linke Seite in die obige Form gebracht hat, ergeben sich die Lösungen wie folgt:

Satz: Ein lineares Gleichungssystem der Form

$$\left(\begin{array}{c|c} I_k & C \\ \hline O & O_{*,\ell} \end{array}\right) \cdot \underline{y} \;=\; \underline{d}$$

besitzt eine Lösung genau dann, wenn der i-te Eintrag von $\underline{d}$ gleich Null ist für alle $i > k$. In diesem Fall ist $\underline{y} = \underline{d}$ eine *partikuläre Lösung*, und die *allgemeine Lösung* ist

$$\underline{y} \;=\; \underline{d} \;+\; \left(\begin{array}{c} -\,C \\ \hline I_\ell \end{array}\right) \cdot \underline{z}$$

für einen beliebigen Spaltenvektor $\underline{z}$ der Länge ℓ.

3.6 Dreiecksmatrizen

Definition: Eine $m \times m$-Matrix $A = \left(a_{ij}\right)_{1 \leqslant i,j \leqslant m}$ mit

(a) $a_{ij} = 0$ für alle $i > j$ heisst *obere Dreiecksmatrix*,
(b) $a_{ij} = 0$ für alle $i < j$ heisst *untere Dreiecksmatrix*.

Proposition: Für jede obere (bzw. untere) Dreiecksmatrix A ist A^T eine untere (bzw. obere) Dreiecksmatrix.

Proposition: Sei $W := \left(\delta_{i,m+1-j}\right)_{1 \leqslant i,j \leqslant m}$ die Permutationsmatrix mit allen Einträgen auf der *Antidiagonalen* gleich 1 und allen übrigen Einträgen 0. Dann gilt $W^{-1} = W$, und für jede obere (bzw. untere) Dreiecksmatrix A ist die Matrix $W^{-1} \cdot A \cdot W$ eine untere (bzw. obere) Dreiecksmatrix.

Bemerkung: Durch beide Transformationen kann man die meisten Aussagen für obere Dreiecksmatrizen in analoge Aussagen für untere Dreiecksmatrizen übersetzen, und umgekehrt.

Proposition: Für jede obere Dreiecksmatrix mit allen Diagonaleinträgen gleich 1 gilt:

(a) Sie kann durch eine Folge elementarer Zeilenumformungen der Art „Addiere ein Vielfaches einer späteren Zeile zu einer früheren“ in die Einheitsmatrix überführt werden.
(b) Sie ist ein Produkt von Matrizen $I_m + \lambda E_{ij}$ für gewisse $\lambda \in K$ und $i < j$.

Proposition: Für je zwei obere Dreiecksmatrizen A und B derselben Grösse sind auch $A + B$ und $A \cdot B$ obere Dreiecksmatrizen, und deren Diagonaleinträge sind die Summe bzw. das Produkt der entsprechenden Diagonaleinträge von A und B.

Proposition: Für jede obere Dreiecksmatrix A gilt:

(a) Jede Links- oder Rechtsinverse von A ist wieder eine obere Dreiecksmatrix.
(b) Die Matrix A ist invertierbar genau dann, wenn kein Diagonaleintrag Null ist.

3.7 Dreieckszerlegung von Matrizen

Satz: Für jede Matrix A existieren eine Permutationsmatrix P und eine invertierbare untere Dreiecksmatrix U, so dass UPA Zeilenstufenform hat.

Satz: *(LR-Zerlegung)* Für jede invertierbare Matrix A existieren eine Permutationsmatrix P, eine invertierbare untere Dreiecksmatrix L, und eine invertierbare obere Dreiecksmatrix R, so dass gilt

$$A = PLR.$$

Satz: Für jede $n \times n$-Matrix sind äquivalent:

(a) Die Matrix ist invertierbar.
(b) Die Matrix lässt sich durch elementare Zeilenoperationen in die Einheitsmatrix überführen.
(c) Die Matrix ist ein Produkt von Matrizen der Form $I_n + \lambda E_{ij}$ für $i \neq j$ und Permutationsmatrizen und invertierbaren Diagonalmatrizen.
(d) Während der Gauss-Elimination bleiben alle Zeilen ungleich Null.

Proposition: Sei A eine $m \times n$-Matrix, und sei (A, I_m) die durch Zusammensetzen entstehende $m \times (n+m)$-Matrix. Seien B eine $m \times n$-Matrix und U eine $m \times m$-Matrix, so dass (B, U) durch eine Folge elementarer Zeilenoperationen aus (A, I_m) entsteht. Dann gilt

$$B = UA.$$

Folge: Für jede invertierbare $m \times m$ Matrix A führt die vollständige Gauss-Elimination von der Matrix (A, I_m) auf die Matrix (I_m, A^{-1}).

4 Vektorräume

4.1 Definition

Sei K ein Körper.

Definition: Ein *Vektorraum über K*, oder kurz ein *K-Vektorraum*, ist ein Tupel $(V, +, \cdot, 0_V)$ bestehend aus einer Menge V mit zwei Abbildungen

$$\begin{aligned} + :&\quad V \times V \to V, \quad (v, v') \mapsto v + v' \\ \cdot :&\quad K \times V \to V, \quad (\lambda, v) \mapsto \lambda \cdot v \end{aligned}$$

und einem ausgezeichneten Element $0_V \in V$, so dass die folgenden *Vektorraumaxiome* gelten:

(I)	$(V, +, 0_V)$ ist eine abelsche Gruppe.	
(II)	$\forall \lambda \in K\ \forall v, v' \in V:\ \lambda \cdot (v + v') = \lambda \cdot v + \lambda \cdot v'$	(Linksdistributivität)
(III)	$\forall \lambda, \lambda' \in K\ \forall v \in V:\ (\lambda + \lambda') \cdot v = \lambda \cdot v + \lambda' \cdot v$	(Rechtsdistributivität)
(IV)	$\forall \lambda, \mu \in K\ \forall v \in V:\ \lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$	(Assoziativität)
(V)	$\forall v \in V:\ 1_K \cdot v = v$	(Einselement)

Die Elemente von V nennt man *Vektoren*.

Proposition: Für jeden K-Vektorraum V und alle $v \in V$ und $\lambda \in K$ gilt:

(a) $0_K \cdot v = \lambda \cdot 0_V = 0_V$.
(b) $\lambda \cdot v = 0_V$ genau dann, wenn $\lambda = 0_K$ oder $v = 0_V$ ist.
(c) $(-1) \cdot v = -v$.

Beispiel: Die Menge $\mathrm{Mat}_{m\times n}(K)$ aller $m\times n$-Matrizen über K mit den oben definierten Operationen ist ein K-Vektorraum.

Spezialfall: Der Raum $\mathrm{Mat}_{n\times 1}(K)$ der Spaltenvektoren der Länge n, bzw. der Raum $\mathrm{Mat}_{1\times n}(K)$ der Zeilenvektoren der Länge n über K. Beide kürzt man meist mit K^n ab. Sobald man das Matrixprodukt verwenden möchte, muss man dann aber dazu sagen, ob man Spaltenvektoren oder Zeilenvektoren meint.

Verallgemeinerung: Für jede Menge X ist die Menge K^X aller Abbildungen $f\colon X \to K$ mit komponentenweiser Addition und skalarer Multiplikation sowie der Nullabbildung $x \mapsto 0$ als Nullelement ein K-Vektorraum.

Beispiel: Jede Menge mit einem Element besitzt eine eindeutige Struktur als K-Vektorraum und heisst dann *Nullraum*. Zum Beispiel $\mathrm{Mat}_{m\times n}(K)$ im Fall $m = 0$ oder $n = 0$, oder K^n im Fall $n = 0$, oder K^X im Fall $X = \varnothing$. Einen Nullraum bezeichnet man oft kurz mit $O = \{0\}$.

4.2 Unterräume

Definition: Ein *Unterraum* oder *Teilraum* eines K-Vektorraums V ist eine Teilmenge $U \subset V$ mit den Eigenschaften:

(a) $U \neq \varnothing$.

(b) $\forall u, u' \in U\colon\ u + u' \in U$.

(c) $\forall \lambda \in K\ \forall u \in U\colon\ \lambda \cdot u \in U$.

Proposition: Eine Teilmenge $U \subset V$ ist ein Unterraum genau dann, wenn sie zusammen mit den Restriktionen der Addition und der skalaren Multiplikation von V selbst einen K-Vektorraum bildet.

Beispiel: Der Nullraum $O = \{0_V\}$ und V selbst sind Unterräume von V.

Beispiel: Für jede $m \times n$-Matrix A über K ist die Lösungsmenge des homogenen linearen Gleichungssystems $\{x \in K^n \mid Ax = 0\}$ ein Unterraum von K^n (Spaltenvektoren).

Beispiel: Sei $F := K^{\mathbb{Z}^{\geqslant 0}}$ der Raum aller unendlichen Folgen $\underline{x} = (x_0, x_1, \ldots)$ in K. Die Teilmenge

$$F_0 \ := \ \big\{\underline{x} \in F \ \big|\ \exists i_0\, \forall i > i_0\colon x_i = 0\big\}$$

aller Folgen, die schliesslich Null werden, ist ein Unterraum von F.

4.3 Durchschnitte und Summen

Proposition: Der Durchschnitt jeder nichtleeren Kollektion von Unterräumen von V ist ein Unterraum von V.

Proposition: Die Vereinigung zweier Unterräume von V ist ein Unterraum von V genau dann, wenn einer der beiden Unterräume in dem anderen enthalten ist.

Die Vereinigung von Unterräumen ist daher im allgemeinen kein vernünftiger Begriff. Stattdessen hat man den folgenden:

Definition: Die Summe jeder Kollektion von Unterräumen $\{V_i \mid i \in I\}$ von V ist die Teilmenge

$$\sum_{i\in I} V_i \;:=\; \left\{ \sum_{i\in I}{}' v_i \;\middle|\; \begin{array}{l} v_i \in V_i \text{ für alle } i \in I, \\ v_i = 0 \text{ für fast alle } i \in I \end{array} \right\} \subset V.$$

Proposition: Die Summe $\sum_{i\in I} V_i$ ist der eindeutige kleinste Unterraum von V, welcher alle V_i enthält. Genauer gilt

$$\sum_{i\in I} V_i \;=\; \bigcap_{U} U,$$

wobei der Durchschnitt über alle Unterräume $U \subset V$ genommen wird, welche alle V_i enthalten.

Abkü: Die Summe endlich vieler Unterräume $V_1, \ldots, V_n$ schreibt man auch so:

$$V_1 + \ldots + V_n \;:=\; \left\{ v_1 + \ldots + v_n \;\middle|\; \forall i = 1, \ldots, n : v_i \in V_i \right\}.$$

4.4 Erzeugnis, Erzeugendensystem

Definition: Gegeben seien eine Menge I sowie Vektoren $v_i \in V$ für alle $i \in I$. Jeder Ausdruck der Form

$$\sum_{i \in I}{}' a_i \cdot v_i$$

mit $a_i \in K$ für alle $i \in I$, und $a_i = 0$ für fast alle $i \in I$, heisst eine *Linearkombination* der Elemente v_i für $i \in I$.

Linearkombinationen einer beliebigen Teilmenge $S \subset V$ kann man elegant und sparsam hinschreiben mit der Indexmenge $I := S$ und $v_s := s$ für alle $s \in S$.

Definition: Für jede Teilmenge S eines K-Vektorraums V heisst die Menge

$$\langle S \rangle \;:=\; \left\{ \sum_{s \in S}{}' a_s \cdot s \;\middle|\; \begin{array}{l} a_s \in K \text{ für alle } s \in S, \\ a_s = 0 \text{ für fast alle } s \in S \end{array} \right\} \subset V$$

das *Erzeugnis von* S.

Proposition: Das Erzeugnis $\langle S \rangle$ ist der eindeutige kleinste Unterraum von V, welcher die Teilmenge S umfasst. Genauer gilt

$$\langle S \rangle \;=\; \bigcap_{U} U,$$

wobei der Durchschnitt über alle Unterräume $U \subset V$ genommen wird, welche S umfassen.

Beispiel: Das Erzeugnis der leeren Menge ist der Nullraum $\langle \varnothing \rangle = \{0\}$.

Definition: Eine Teilmenge $S \subset V$ mit $\langle S \rangle = V$ heisst ein *Erzeugendensystem von* V.

Eine Teilmenge S ist also ein Erzeugendensystem von V genau dann, wenn jeder Vektor in V eine Darstellung als Linearkombination der Elemente von S besitzt.

Beispiel: Die Menge V ist ein Erzeugendensystem von V.

4.5 Lineare Unabhängigkeit

Definition: Eine Teilmenge $S \subset V$ heisst *linear abhängig*, wenn Koeffizienten $a_s \in K$ für alle $s \in S$ existieren, so dass $a_s = 0$ ist für fast alle, aber nicht für alle $s \in S$, und

$$\sum_{s \in S}{}' a_s \cdot s = 0.$$

Existieren keine solchen a_s, so heisst S *linear unabhängig*.

Eine Teilmenge S ist also linear unabhängig genau dann, wenn der Nullvektor keine nicht-triviale Darstellung als Linearkombination der Elemente von S besitzt.

Proposition: Für jede Teilmenge $S \subset V$ sind äquivalent:

(a) S ist linear unabhängig.

(b) Kein Element von S ist eine Linearkombination der übrigen Elemente von S.

(c) Jeder Vektor in V besitzt höchstens eine Darstellung als Linearkombination der Elemente von S.

4.6 Basis

Definition: Ein linear unabhängiges Erzeugendensystem von V heisst eine *Basis von V*.

Eine Teilmenge S ist also eine Basis von V genau dann, wenn jeder Vektor in V genau eine Darstellung als Linearkombination der Elemente von S besitzt.

Satz: Für jede Teilmenge $S \subset V$ sind äquivalent:

(a) S ist eine Basis von V.

(b) S ist ein minimales Erzeugendensystem von V.

(c) S ist eine maximale linear unabhängige Teilmenge von V.

Satz: Für jedes Erzeugendensystem E von V und jede linear unabhängige Teilmenge $L \subset E$ existiert eine Basis B von V mit $L \subset B \subset E$.

Folge: (a) Jedes Erzeugendensystem von V enthält eine Basis von V.

(b) Jede linear unabhängige Teilmenge von V lässt sich zu einer Basis von V erweitern.

(c) Jeder Vektorraum besitzt eine Basis.

Vorsicht: Ein Vektorraum besitzt im allgemeinen viele verschiedene Basen. Bevor man eine spezielle Basis gefunden oder gewählt hat, darf man daher nur mit dem unbestimmten Artikel von „einer Basis“ sprechen.

4.7 Dimension

Kleiner Austauschsatz: Für jede Basis B von V und jeden Vektor $v_0 \in V \smallsetminus (B \cup \{0\})$ existiert ein $b_0 \in B$, so dass $(B \smallsetminus \{b_0\}) \cup \{v_0\}$ eine Basis von V ist.

Grosser Austauschsatz: Für jede Basis B von V und jede linear unabhängige Teilmenge $L \subset V$ existiert eine Injektion $i\colon L \hookrightarrow B$, so dass $L \cup (B \smallsetminus i(L))$ eine Basis von V ist und die Vereinigung disjunkt ist.

Folge: Für jedes Erzeugendensystem E von V und jede linear unabhängige Teilmenge L von V gilt

$$\operatorname{card}(L) \leqslant \operatorname{card}(E).$$

Folge: Je zwei Basen von V haben dieselbe Kardinalität.

Definition: Diese Kardinalität heisst die *Dimension von* V, geschrieben $\dim_K(V)$ oder $\dim(V)$.

Beispiel: Für jede natürliche Zahl n bilden die Vektoren $e_i := \big(\delta_{i,j}\big)_{j=1,\ldots,n}$ für alle $i = 1, \ldots, n$ eine Basis von K^n, genannt die *Standardbasis von* K^n. Insbesondere ist also $\dim_K(K^n) = n$.

Beispiel: Der Nullraum hat die Basis $\varnothing$ und folglich die Dimension 0.

Satz: Für jeden Vektorraum V der Dimension $n < \infty$ gilt:

(a) Jedes Erzeugendensystem E von V mit $|E| = n$ ist eine Basis von V.

(b) Jede linear unabhängige Teilmenge L von V mit $|L| = n$ ist eine Basis von V.

Satz: Gegebene Spaltenvektoren $v_1, \ldots, v_n \in K^n$ bilden eine Basis von K^n genau dann, wenn die $n \times n$-Matrix $(v_1, \ldots, v_n)$ invertierbar ist.

Beispiel: Für jedes $1 \leqslant i \leqslant n$ sei $v_i = (a_{ij})_j$ mit $a_{ii} \neq 0$ und $a_{ij} = 0$ für alle $j > i$. Dann ist $(v_1, \ldots, v_n)$ eine obere Dreiecksmatrix mit allen Diagonaleinträgen ungleich Null und folglich invertierbar; also ist $\{v_1, \ldots, v_n\}$ eine Basis von K^n.

Beispiel: Die Addition und Multiplikation von reellen und komplexen Zahlen macht $\mathbb{C}$ zu einem Vektorraum über $\mathbb{R}$. Dieser hat die Basis $\{1, i\}$ und folglich die Dimension $\dim_{\mathbb{R}}(\mathbb{C}) = 2$.

Satz: Die Addition und Multiplikation von reellen und rationalen Zahlen macht $\mathbb{R}$ zu einem Vektorraum über $\mathbb{Q}$. Dieser ist unendlich-dimensional; genauer gilt

$$\dim_{\mathbb{Q}}(\mathbb{R}) = \operatorname{card}(\mathbb{R}).$$

4.8 Direkte Summen, Komplemente

Definition: Seien V_1 und V_2 Unterräume von V mit $V = V_1 + V_2$ und $V_1 \cap V_2 = \{0\}$. Dann heisst V die *(innere) direkte Summe von V_1 und V_2*, und wir schreiben

$$V \;=\; V_1 \oplus V_2.$$

Weiter heisst dann V_2 ein *Komplement von V_1 in V*, und umgekehrt.

Proposition: Die genannten Bedingungen an V_1 und V_2 sind äquivalent zu der Bijektivität der Abbildung

$$V_1 \times V_2 \longrightarrow V,\ (v_1, v_2) \mapsto v_1 + v_2.$$

Satz: Jeder Unterraum eines Vektorraums besitzt ein Komplement.

Beispiel: Es gilt $V = V \oplus \{0\}$; somit sind V und $\{0\}$ Komplemente voneinander in V.

Beispiel: Sei $V_1 := \langle \binom{a}{b} \rangle \subset K^2$ für gegebene $a, b \in K$. Dann ist $\langle \binom{1}{0} \rangle$ ein Komplement von V_1 genau dann, wenn $b \neq 0$ ist, und $\langle \binom{0}{1} \rangle$ ist ein Komplement von V_1 genau dann, wenn $a \neq 0$ ist.

Variante: Analog zu Summen beliebig vieler Teilräume von V kann man auch beliebige direkte Summen definieren. Siehe dazu §5.4 sowie [Kowalsky-Michler, §2.3].

4.9 Dimension von Unterräumen

Satz: Sei V' ein Unterraum von V.

(a) Dann gilt $\dim(V') \leqslant \dim(V)$.

(b) Gilt weiter $\dim(V') = \dim(V) < \infty$, so folgt $V' = V$.

Beispiel: Wie früher sei $F := K^{\mathbb{Z}^{\geqslant 0}}$ der Raum aller unendlichen Folgen in K und F_0 der Unterraum aller Folgen, die schliesslich Null werden. Für jedes $i \geqslant 0$ sei $e_i := (\delta_{ij})_{j \geqslant 0}$ die Folge mit Eintrag 1 an der Stelle i und allen übrigen Einträgen 0. Dann ist $\{e_i \mid i \geqslant 0\}$ eine Basis von F_0. Insbesondere ist $\dim(F_0)$ abzählbar unendlich. Dagegen ist $\dim(F)$ überabzählbar.

Satz: Ist $V = V_1 \oplus V_2$, so gilt

$$\dim(V) = \dim(V_1) + \dim(V_2).$$

Satz: Für beliebige Unterräume V_1 und V_2 von V gilt

$$\dim(V_1 + V_2) + \dim(V_1 \cap V_2) \;=\; \dim(V_1) + \dim(V_2).$$

4.10 Quotientenvektorräume

Sei U ein Unterraum eines K-Vektorraums V. Für jedes $v \in V$ betrachte die Teilmenge

$$v + U \;:=\; \{v + u \mid u \in U\} \;\subset\; V.$$

Betrachte die Menge der Teilmengen

$$V/U \;:=\; \{v + U \mid v \in V\}.$$

Geometrisch ist dies die Menge aller *zu U parallelen* affin-linearen Teilräume von V.

Proposition: Je zwei Teilmengen der Form $v + U$ sind entweder gleich oder disjunkt, und die Vereinigung aller ist V.

Genauer ist $v + U = v' + U \Leftrightarrow v - v' \in U \Leftrightarrow v' \in v + U \Leftrightarrow v \in v' + U$.

Proposition: Die Menge V/U besitzt eine eindeutige Struktur eines K-Vektorraums, so dass gilt:

(a) $\forall v, v' \in V : (v + U) + (v' + U) = (v + v') + U$.
(b) $\forall v \in V \; \forall \lambda \in K : \lambda \cdot (v + U) = (\lambda v) + U$.

Für diese gilt weiter:

(c) Das Nullelement von V/U ist $0_{V/U} = 0_V + U = U$.
(d) Das additive Inverse jedes Elements $v + U$ ist $-(v + U) = (-v) + U$.

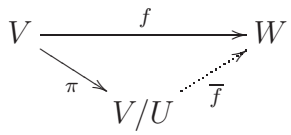
Definition: V/U heisst der *Quotientenvektorraum* oder *Faktorraum von V nach U*.

Proposition: Die Abbildung $\pi\colon V \to V/U$, $v \mapsto v + U$ ist linear und surjektiv und hat Kern U.

Beispiel: (a) Es ist $U = V$ genau dann, wenn $V/U = 0$ ist.

(b) Es ist $U = 0$ genau dann, wenn π ein Isomorphismus ist.

Proposition: *(Universelle Eigenschaft)* Für jeden K-Vektorraum W und jede lineare Abbildung $f\colon V \to W$ mit $U \subset \mathrm{Kern}(f)$ existiert genau eine lineare Abbildung $\overline{f}\colon V/U \to W$ mit $\overline{f} \circ \pi = f$, das heisst, so dass das folgende Diagramm kommutiert:



Proposition: Ein Unterraum $U' \subset V$ ist ein Komplement von U genau dann, wenn die Abbildung $\pi|_{U'}\colon U' \to V/U$ ein Isomorphismus ist.

Beispiel: Ist V ein endlich-dimensionaler euklidischer oder unitärer Vektorraum über $K = \mathbb{R}$ bzw. $\mathbb{C}$, so gilt stets $V = U \oplus U^\perp$ und somit $U^\perp \xrightarrow{\sim} V/U$.

Beispiel: Für $m \leqslant n$ betrachte die Injektion $i\colon K^m \hookrightarrow K^n$, $x \mapsto \binom{x}{0}$ sowie die dazu komplementäre Injektion $j\colon K^{n-m} \hookrightarrow K^n$, $y \mapsto \binom{0}{y}$. Dann induziert j einen Isomorphismus $K^{n-m} \xrightarrow{\sim} K^n/i(K^m)$.

Proposition: Sei B eine Basis von V, so dass $B \cap U$ eine Basis von U ist. Dann ist $\{v + U \mid v \in B \smallsetminus U\}$ eine Basis von V/U. Insbesondere gilt

$$\dim_K(V) \;=\; \dim_K(U) + \dim_K(V/U).$$

Variante: Sei $B = (b_1, \ldots, b_n)$ eine geordnete Basis von V, deren Anfangssegment $B' := (b_1, \ldots, b_m)$ eine geordnete Basis von U ist. Dann ist $B'' := (b_{m+1}+U, \ldots, b_n+U)$ eine geordnete Basis von V/U.

Proposition: Für jedes $i = 1, 2$ sei V_i ein K-Vektorraum mit geordneter Basis B_i, sei $U_i \subset V_i$ der von einem Anfangssegment B_i' von B_i erzeugte Unterraum, und sei B_i'' die wie oben induzierte geordnete Basis von V_i/U_i. Jede lineare Abbildung $f\colon V_1 \to V_2$ mit $f(U_1) \subset U_2$ induziert natürliche lineare Abbildungen

$$f'\colon U_1 \to U_2,\ u \mapsto f(u),$$

$$f''\colon V_1/U_1 \to V_2/U_2,\ v + U_1 \mapsto f(v) + U_2,$$

und die Darstellungsmatrix von f hat die Blockdreiecksgestalt

$$M_{B_2,B_1}(f) \;=\; \begin{pmatrix} M_{B_2',B_1'}(f') & * \\ 0 & M_{B_2'',B_1''}(f'') \end{pmatrix}.$$

Beispiel: Von der Raumzeit der Newtonschen Mechanik aus gesehen ist die Zeit auf natürliche Weise ein Faktorraum und kein Unterraum.

5 Lineare Abbildungen

5.1 Definition

Gegeben seien Vektorräume U, V, W über einem Körper K.

Definition: Eine Abbildung $f\colon V \to W$ heisst *K-linear*, wenn gilt:

(a) $\forall v, v' \in V : f(v + v') = f(v) + f(v')$ und

(b) $\forall v \in V\ \forall \lambda \in K : f(\lambda v) = \lambda \cdot f(v)$.

Wenn der Körper K aus dem Zusammenhang ersichtlich ist, sagt man nur kurz *linear*.

Beispiel: Die *Nullabbildung* $V \to W$, $v \mapsto 0_W$ ist linear.

Beispiel: Die *identische Abbildung* $\mathrm{id}_V\colon V \to V$, $v \mapsto v$ ist linear.

Beispiel: Sei A eine $m \times n$-Matrix über K. Dann ist *Linksmultiplikation* mit A eine lineare Abbildung der Räume von Spaltenvektoren

$$L_A\colon\ K^n \to K^m,\ v \mapsto Av,$$

und *Rechtsmultiplikation* mit A eine lineare Abbildung der Räume von Zeilenvektoren

$$R_A\colon\ K^m \to K^n,\ v \mapsto vA.$$

Proposition: Für jede lineare Abbildung der Räume von Spaltenvektoren $f\colon K^n \to K^m$ existiert genau eine $m \times n$-Matrix A über K mit $L_A = f$. Die i-te Spalte der Matrix A ist genau das Bild des i-ten Standardbasisvektors e_i unter f.

Tipp: Die Matrix A klar von der linearen Abbildung L_A bzw. R_A unterscheiden!

Proposition: Für jede Basis B von V und jede Abbildung $f_0\colon B \to W$ existiert genau eine lineare Abbildung $f\colon V \to W$ mit $\forall\, b \in B\colon f(b) = f_0(b)$, nämlich

$$\sum_{b\in B}{}' x_b \cdot b \ \mapsto\ \sum_{b\in B}{}' x_b \cdot f_0(b).$$

5.2 Kern und Bild

Sei $f\colon V \to W$ eine lineare Abbildung.

Definition: Der *Kern von f* und das *Bild von f* sind die Teilmengen

$$\begin{aligned} \operatorname{Kern}(f) &:= \{v \in V \mid f(v) = 0\} \subset V, \\ \operatorname{Bild}(f) &:= \{f(v) \mid v \in V\} \subset W. \end{aligned}$$

Proposition: (a) $\operatorname{Kern}(f)$ ist ein Unterraum von V.

(b) f ist injektiv genau dann, wenn $\operatorname{Kern}(f) = 0$ ist.

Proposition: (a) $\operatorname{Bild}(f)$ ist ein Unterraum von W.

(b) f ist surjektiv genau dann, wenn $\operatorname{Bild}(f) = W$ ist.

Beispiel: Ein inhomogenes lineares Gleichungssystem $Ax = b$ ist lösbar genau dann, wenn der Vektor b im Bild $\operatorname{Bild}(L_A)$ der linearen Abbildung L_A liegt. Ist es lösbar, so ist die Lösung eindeutig genau dann, wenn $\operatorname{Kern}(L_A) = 0$ ist.

Satz: Es gilt

$$\dim(V) = \dim \operatorname{Kern}(f) + \dim \operatorname{Bild}(f).$$

Insbesondere ist $\dim \operatorname{Kern}(f) \leqslant \dim(V)$ und $\dim \operatorname{Bild}(f) \leqslant \dim(V)$.

Satz: Ist $\dim(V) = \dim(W) < \infty$, so gilt

$$f \text{ injektiv} \iff f \text{ surjektiv} \iff f \text{ bijektiv}.$$

5.3 Komposition, Isomorphismen

Proposition: Für je zwei lineare Abbildungen $f\colon U \to V$ und $g\colon V \to W$ ist die Komposition $g \circ f\colon U \to W$ linear.

Proposition: Für jede $m \times n$-Matrix A und jede $n \times \ell$-Matrix B gilt $L_A \circ L_B = L_{AB}$.

Definition: Eine lineare Abbildung $f\colon V \to W$, zu welcher eine lineare Abbildung $g\colon W \to V$ existiert mit $g \circ f = \mathrm{id}_V$ und $f \circ g = \mathrm{id}_W$, heisst ein *Isomorphismus*.

Satz: Eine lineare Abbildung f ist ein Isomorphismus genau dann, wenn sie bijektiv ist. Die beidseitige *Inverse* g in der obigen Definition ist dann eindeutig bestimmt und gleich der Umkehrabbildung $f^{-1}\colon W \to V$.

Beispiel: Die Abbildung $L_A\colon K^n \to K^m$ ist ein Isomorphismus genau dann, wenn $m = n$ ist und A invertierbar ist. Ihre Umkehrabbildung ist dann $L_A^{-1} = L_{A^{-1}}$.

Definition: Zwei Vektorräume V und W über einem Körper K heissen *isomorph*, in Symbolen $V \cong W$, wenn ein Isomorphismus zwischen ihnen existiert.

Vorsicht: Gibt es einen Isomorphismus, so gibt es im allgemeinen viele, und möglicherweise keinen besonders ausgezeichneten. Isomorphe Vektorräume darf man also nicht ohne weiteres miteinander identifizieren.

Satz: Die Relation $\cong$ ist eine Äquivalenzrelation. Genauer ist jede Komposition zweier Isomorphismen ein Isomorphismus, die identische Abbildung auf jedem Vektorraum ein Isomorphismus, und die Inverse jedes Isomorphismus ein Isomorphismus.

Satz: Es gilt $V \cong W$ genau dann, wenn $\dim(V) = \dim(W)$ ist.

Definition:

(a) Ein *Monomorphismus* ist eine injektive lineare Abbildung, geschrieben $V \hookrightarrow W$.
(b) Ein *Epimorphismus* ist eine surjektive lineare Abbildung, geschrieben $V \twoheadrightarrow W$.
(c) Ein *Isomorphismus* ist ... (siehe oben) ... , geschrieben $V \xrightarrow{\sim} W$.
(d) Ein *Endomorphismus von* V ist eine lineare Abbildung $V \to V$.
(e) Ein *Automorphismus von* V ist ein Isomorphismus $V \xrightarrow{\sim} V$.

Proposition: Die Menge $\mathrm{Aut}(V)$ aller Automorphismen von V zusammen mit der Komposition $\circ$ und dem neutralen Element id_V ist eine Gruppe, genannt die *Automorphismengruppe von* V.

Beispiel: Die Abbildung $A \mapsto L_A$ induziert eine Bijektion $\mathrm{GL}_n(K) \to \mathrm{Aut}(K^n)$, welche mit der Gruppenoperation auf beiden Seiten verträglich ist, also einen *Gruppen-Isomorphismus*.

5.4 Direkte Produkte und Summen

Definition: Das kartesische Produkt von K-Vektorräumen

$$\bigtimes_{i\in I} V_i \;:=\; \big\{(v_i)_{i\in I} \;\big|\; \forall i\colon v_i \in V_i\big\}$$

versehen mit den Operationen

$$\begin{aligned}(v_i)_i + (v_i')_i &:= (v_i + v_i')_i\\ \lambda\cdot(v_i)_i &:= (\lambda v_i)_i\end{aligned}$$

und dem Nullelement $(0_{V_i})_i$ heisst das *(direkte) Produkt von* $(V_i)_{i\in I}$. Ihre Teilmenge

$$\bigboxplus_{i\in I} V_i \;:=\; \big\{(v_i)_{i\in I} \;\big|\; \forall i\colon v_i \in V_i,\ \text{fast alle } v_i = 0\big\}$$

heisst die *äussere direkte Summe von* $(V_i)_{i\in I}$.

Konvention: Sind alle Faktoren gleich, so schreibt man oft $V^I := \bigtimes_{i\in I} V$ und $V^{(I)} := \bigboxplus_{i\in I} V$. Die Elemente von $\bigboxplus_{i\in I} V_i$, insbesondere von $\bigboxplus_{i\in I} K$, schreibt man oft als formale Linearkombinationen $(v_i)_i = \sum_{i\in I}' v_i\cdot X_i$ mit neugewählten Symbolen X_i.

Proposition: Das Produkt $\bigtimes_{i\in I} V_i$ ist ein K-Vektorraum und $\bigboxplus_{i\in I} V_i$ ist ein Unterraum, und für jedes $j \in I$ sind die folgenden Abbildungen linear:

$$\begin{aligned}\mathrm{proj}_j\colon &\ \bigtimes_{i\in I} V_i \to V_j,\quad (v_i)_i \mapsto v_j\\ \mathrm{incl}_j\colon &\ V_j \to \bigboxplus_{i\in I} V_i,\quad v_j \mapsto \left(\begin{cases} v_j & \text{falls } i = j\\ 0 & \text{falls } i \neq j\end{cases}\right)_i\end{aligned}$$

Proposition: Für beliebige Unterräume V_i eines Vektorraums V ist die folgende Abbildung linear:

$$\bigboxplus_{i\in I} V_i \longrightarrow V,\ (v_i)_i \mapsto \sum_{i\in I}' v_i.$$

Definition: Der Vektorraum V heisst die *innere direkte Summe von* V_i *für* $i \in I$, wenn die obige Abbildung ein Isomorphismus ist, und dann schreiben wir

$$\bigoplus_{i\in I} V_i \;=\; V.$$

Konvention: Im Fall $I = \{1,\ldots,r\}$ schreiben wir auch

$$\begin{aligned}V_1 \times \ldots \times V_r &= \bigtimes_{i=1}^r V_i,\\ V_1 \boxplus \ldots \boxplus V_r &= \bigboxplus_{i=1}^r V_i,\\ V_1 \oplus \ldots \oplus V_r &= \bigoplus_{i=1}^r V_i.\end{aligned}$$

Für $r = 2$ stimmt diese Definition von $V_1 \oplus V_2$ mit derjenigen in §4.8 überein.

Konvention: Oft werden innere und äussere direkte Summe mit demselben Symbol $\bigoplus$ bezeichnet. Welche dann jeweils gemeint ist, muss man aus dem Zusammenhang erschliessen.

5.5 Geordnete Basen

Definition: Ein Tupel $(v_1, \ldots, v_n)$ von Vektoren in V heisst

(a) *linear unabhängig*, wenn $\forall x_1, \ldots, x_n \in K\colon \big(0 = \sum_{i=1}^n x_i v_i\big) \to \big(x_1 = \ldots = x_n = 0\big)$.

(b) *Erzeugendensystem von V*, wenn $\forall v \in V\ \exists\, x_1, \ldots, x_n \in K\colon\ v = \sum_{i=1}^n x_i v_i$.

(c) *geordnete Basis von V*, wenn $\forall v \in V\ \exists!\, x_1, \ldots, x_n \in K\colon\ v = \sum_{i=1}^n x_i v_i$.

Der Begriff „geordnete Basis“ setzt also voraus, dass der Vektorraum endlich-dimensional ist.

Proposition: Ein Tupel $(v_1, \ldots, v_n)$ von Vektoren in V ist

(a) linear unabhängig genau dann, wenn $v_1, \ldots, v_n$ paarweise verschieden sind und die Menge $\{v_1, \ldots, v_n\}$ linear unabhängig ist.

(b) ein Erzeugendensystem von V genau dann, wenn die Menge $\{v_1, \ldots, v_n\}$ ein Erzeugendensystem von V ist.

(c) eine geordnete Basis von V genau dann, wenn $v_1, \ldots, v_n$ paarweise verschieden sind und die Menge $\{v_1, \ldots, v_n\}$ eine Basis von V ist.

Proposition: Für jedes Tupel $T := (v_1, \ldots, v_n)$ von Vektoren in V ist die Abbildung

$$\varphi_T\colon K^n \to V,\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{i=1}^n x_i v_i$$

linear. Sie ist

(a) injektiv genau dann, wenn T linear unabhängig ist.

(b) surjektiv genau dann, wenn T ein Erzeugendensystem von V ist.

(c) ein Isomorphismus genau dann, wenn T eine geordnete Basis von V ist.

5.6 Darstellungsmatrix

Definition: Ein Diagramm bestehend aus Mengen und durch Pfeile dargestellte Abbildungen zwischen diesen heisst *kommutativ*, wenn für je zwei Wege in Pfeilrichtung mit demselben Startpunkt und demselben Endpunkt die zusammengesetzten Abbildungen übereinstimmen.

Definition: Seien $B = (v_1, \ldots, v_n)$ eine geordnete Basis von V und $B' = (w_1, \ldots, w_m)$ eine geordnete Basis von W. Die *Darstellungsmatrix* einer linearen Abbildung $f\colon V \to W$ *bezüglich der Basen* B *und* B' ist die eindeutig bestimmte $m \times n$-Matrix A, für die $f \circ \varphi_B = \varphi_{B'} \circ L_A$ gilt, das heisst, für die das folgende Diagramm kommutiert:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \wr\uparrow{\scriptstyle \varphi_B} & & \wr\uparrow{\scriptstyle \varphi_{B'}} \\ K^n & \xrightarrow{L_A} & K^m. \end{array}$$

Wir bezeichnen diese Matrix A mit $M_{B',B}(f)$.

Eine explizite Rechnung mit dem Ansatz $A = \big(a_{ij}\big)_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ liefert für alle $1 \leqslant j \leqslant n$:

$$f(v_j) \;=\; f(\varphi_B(e_j)) \;=\; \varphi_{B'}(L_A(e_j)) \;=\; \varphi_{B'}(Ae_j) \;=\; \varphi_{B'}\Big(\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}\Big) \;=\; \sum_{i=1}^{m} a_{ij} w_i.$$

Beispiel: Die Darstellungsmatrix von $L_A\colon K^n \to K^m$ bezüglich der jeweiligen Standardbasis ist A.

Satz: Für jede geordnete Basis $B = (v_1, \ldots, v_n)$ von V gilt

$$M_{B,B}(\mathrm{id}_V) = I_n.$$

Proposition: Sei B, B', B'' je eine geordnete Basis von U, V, bzw. von W. Für alle linearen Abbildungen $f\colon U \to V$ und $g\colon V \to W$ gilt dann

$$M_{B'',B}(g \circ f) \;=\; M_{B'',B'}(g) \cdot M_{B',B}(f).$$

Proposition: Seien $B = (v_1, \ldots, v_n)$ eine geordnete Basis von V und $B' = (w_1, \ldots, w_m)$ eine geordnete Basis von W. Eine lineare Abbildung $f\colon V \to W$ ist ein Isomorphismus genau dann, wenn die Darstellungsmatrix $M_{B',B}(f)$ quadratisch und invertierbar ist, und dann gilt $M_{B,B'}(f^{-1}) = M_{B',B}(f)^{-1}$.

5.7 Basiswechsel

Definition: Die Matrix $M_{\tilde{B},B}(\mathrm{id}_V)$ für geordnete Basen B und $\tilde{B}$ desselben Vektorraums V heisst die zu B und $\tilde{B}$ assoziierte *Basiswechselmatrix*.

Proposition: Die Basiswechselmatrix $M_{\tilde{B},B}(\mathrm{id}_V)$ ist invertierbar, und ihre Inverse ist gleich $M_{\tilde{B},B}(\mathrm{id}_V)^{-1} = M_{B,\tilde{B}}(\mathrm{id}_V)$.

Proposition: Seien B, $\tilde{B}$ geordnete Basen von V, und B', $\tilde{B}'$ geordnete Basen von W. Dann gilt für jede lineare Abbildung $f\colon V \to W$

$$M_{\tilde{B}',\tilde{B}}(f) \;=\; M_{\tilde{B}',B'}(\mathrm{id}_W) \cdot M_{B',B}(f) \cdot M_{B,\tilde{B}}(\mathrm{id}_V).$$

Spezialfall: Seien B und $\tilde{B}$ geordnete Basen von V. Dann gilt für jede lineare Abbildung $f\colon V \to V$

$$\begin{aligned} M_{\tilde{B},\tilde{B}}(f) &= M_{\tilde{B},B}(\mathrm{id}_V) \cdot M_{B,B}(f) \cdot M_{B,\tilde{B}}(\mathrm{id}_V) \\ &= M_{\tilde{B},B}(\mathrm{id}_V) \cdot M_{B,B}(f) \cdot M_{\tilde{B},B}(\mathrm{id}_V)^{-1} \\ &= M_{B,\tilde{B}}(\mathrm{id}_V)^{-1} \cdot M_{B,B}(f) \cdot M_{B,\tilde{B}}(\mathrm{id}_V). \end{aligned}$$

5.8 Rang

Definition: Der *Rang* einer linearen Abbildung f ist $\mathrm{Rang}(f) := \dim \mathrm{Bild}(f)$.

Proposition: Für jede lineare Abbildung $f\colon V \to W$ und beliebige Isomorphismen $\varphi\colon V' \xrightarrow{\sim} V$ und $\psi\colon W \xrightarrow{\sim} W'$ gilt

$$\mathrm{Rang}(\psi \circ f \circ \varphi) = \mathrm{Rang}(f).$$

Satz: Für jede $m \times n$-Matrix A existieren invertierbare Matrizen U und V, so dass UAV eine Blockmatrix der Form

$$\left(\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right)$$

ist für ein geeignetes $0 \leqslant r \leqslant \min\{m, n\}$, wobei jeweils O die Nullmatrix bezeichnet.

Satz: Die Zahl r hängt nur von A ab.

Definition: Die Zahl r heisst der *Rang* von A und wird bezeichnet mit $\mathrm{Rang}(A)$.

Satz: Für jede Matrix A sind die folgenden Zahlen gleich:

(a) Der Rang von A.
(b) Der Rang von A^T.
(c) Die maximale Anzahl linear unabhängiger Spalten von A *(Spaltenrang)*.
(d) Die maximale Anzahl linear unabhängiger Zeilen von A *(Zeilenrang)*.

Proposition: Für jede lineare Abbildung f und je zwei geordnete Basen B und B' gilt

$$\mathrm{Rang}(f) = \mathrm{Rang}\, M_{B',B}(f).$$

Beispiel: Eine $m \times n$-Matrix hat Rang m genau dann, wenn man aus ihren Spalten Vektoren $v_1, \ldots, v_m$ auswählen kann, so dass die Matrix $(v_1, \ldots, v_m)$ invertierbar ist.

Beispiel: Eine $m \times m$-Matrix hat Rang m genau dann, wenn sie invertierbar ist.

5.9 Abbildungsräume

Definition: Für je zwei K-Vektorräume V und W setzen wir

$$\mathrm{Hom}_K(V,W) \;:=\; \{h\colon V \to W \mid h \text{ } K\text{-linear}\}.$$

Proposition: Dies ist ein Untervektorraum des Raums W^V aller Abbildungen $V \to W$.

Proposition: Für je zwei lineare Abbildungen $f\colon V' \to V$ und $g\colon W \to W'$ ist die Abbildung

$$C_{g,f}\colon \mathrm{Hom}_K(V,W) \to \mathrm{Hom}_K(V',W'),\; h \mapsto g \circ h \circ f$$

wohldefiniert und linear. Sind f und g Isomorphismen, so auch $C_{g,f}$.

Proposition: Die folgende Abbildung ist ein Isomorphismus von K-Vektorräumen:

$$\mathrm{Mat}_{m\times n}(K) \xrightarrow{\sim} \mathrm{Hom}_K(K^n,K^m),\; A \mapsto L_A.$$

Proposition: Für jede geordnete Basis $B = (v_1,\ldots,v_n)$ von V und jede geordnete Basis $B' = (w_1,\ldots,w_m)$ von W ist die folgende Abbildung ein Isomorphismus:

$$\mathrm{Hom}_K(V,W) \longrightarrow \mathrm{Mat}_{m\times n}(K),\; f \mapsto M_{B',B}(f).$$

Satz: Für je zwei endlichdimensionale K-Vektorräume V und W gilt

$$\dim_K \mathrm{Hom}_K(V,W) \;=\; \dim_K(V)\cdot\dim_K(W).$$

Definition: Für jeden K-Vektorraum V setzen wir

$$\mathrm{End}_K(V) \;:=\; \mathrm{Hom}_K(V,V).$$

Proposition: Mit der Addition und Komposition von Endomorphismen sowie der Nullabbildung und der identischen Abbildung ist $(\mathrm{End}_K(V),+,\circ,0_V,\mathrm{id}_V)$ ein Ring, genannt der *Endomorphismenring von V*.

5.10 Dualraum

Definition: Der Vektorraum

$$V^* := \mathrm{Hom}_K(V, K)$$

heisst der *Dualraum von* V, und seine Elemente heissen *Linearformen auf* V.

Proposition-Definition: Sei $B = (v_1, \ldots, v_n)$ eine geordnete Basis eines K-Vektorraums V. Für jedes $1 \leqslant i \leqslant n$ sei $\ell_i \in V^*$ die lineare Abbildung

$$\ell_i \colon V \longrightarrow K, \ \textstyle\sum_{j=1}^n x_j v_j \mapsto x_i.$$

Dann ist $B^* := (\ell_1, \ldots, \ell_n)$ eine geordnete Basis von V^*, genannt die *duale Basis zu* B.

Satz: Für jeden K-Vektorraum V gilt

$$\dim_K(V^*) = \begin{cases} \dim_K(V) & \text{falls } \dim_K(V) < \infty, \\ \infty & \text{falls } \dim_K(V) = \infty. \end{cases}$$

Beispiel: Wie früher sei $F := K^{\mathbb{Z}^{\geqslant 0}}$ der Raum aller unendlichen Folgen und F_0 der Teilraum aller Folgen, die schliesslich Null werden. Der Dualraum F_0^* von F_0 ist dann isomorph zu F, und folglich gilt $\dim F_0 < \dim F_0^*$ im Sinne unendlicher Kardinalzahlen.

Proposition: (a) Für jede lineare Abbildung $f \colon V \to W$ ist die Abbildung

$$f^* \colon W^* \to V^*, \ \ell \mapsto \ell \circ f$$

wohldefiniert und linear. Sie heisst die *duale Abbildung zu* f.

(b) Für je zwei lineare Abbildungen $f \colon U \to V$ und $g \colon V \to W$ gilt $(g \circ f)^* = f^* \circ g^*$.

(c) Die duale Abbildung der identischen Abbildung ist $\mathrm{id}_V^* = \mathrm{id}_{V^*}$.

(d) Für jeden Isomorphismus $f \colon V \to W$ ist $f^* \colon W^* \to V^*$ ein Isomorphismus.

Proposition: Seien B eine geordnete Basis von V und B' eine geordnete Basis von W, und sei $f \colon V \to W$ eine lineare Abbildung. Dann gilt

$$M_{B^*, B'^*}(f^*) = M_{B', B}(f)^T.$$

Proposition: Für jeden K-Vektorraum V und jedes Element $v \in V$ ist die Auswertungs-Abbildung (evaluation)

$$\mathrm{ev}_v \colon V^* \longrightarrow K, \ \ell \mapsto \ell(v)$$

linear, also ein Element des *Bidualraums* $(V^*)^*$. Die induzierte Abbildung

$$V \longrightarrow (V^*)^*, \ v \mapsto \mathrm{ev}_v$$

ist linear und injektiv. Sie ist ein Isomorphismus genau dann, wenn V endlich-dimensional ist.

6 Determinanten

6.1 Symmetrische Gruppe

Definition: Eine bijektive Abbildung von einer Menge X auf sich selbst heisst eine *Permutation von X*.

Proposition-Definition: Die Menge aller Permutationen der Menge $\{1, \ldots, n\}$ zusammen mit der Komposition von Abbildungen und der identischen Abbildung id als neutrales Element ist eine Gruppe, genannt die *symmetrische Gruppe vom Grad n*.

Elemente von S_n bezeichnet man üblicherweise mit kleinen griechischen Buchstaben und schreibt ihre Operation klammernlos in der Form $\sigma\colon i \mapsto \sigma i$.

Proposition: Es gilt $|S_n| = n!$.

Definition: Ein Paar (i, j) mit $1 \leqslant i < j \leqslant n$ und $\sigma i > \sigma j$ heisst ein *Fehlstand von σ*. Die Zahl

$$\operatorname{sgn}(\sigma) \;:=\; \big(-1\big)^{\text{Anzahl Fehlstände von } \sigma}$$

heisst das *Signum* oder die *Signatur* oder das *Vorzeichen von σ*. Eine Permutation mit $\operatorname{sgn}(\sigma) = 1$ heisst *gerade*, eine mit $\operatorname{sgn}(\sigma) = -1$ heisst *ungerade*.

Beispiel: Eine Permutation, die zwei verschiedene Ziffern vertauscht und alle übrigen Ziffern festlässt, heisst *Transposition*. Jede Transposition hat Signum -1.

Beispiel: Eine Permutation, die k verschiedene Ziffern zyklisch vertauscht und alle übrigen Ziffern festlässt, hat Signum $(-1)^{k-1}$.

Lemma: Für jedes $\sigma \in S_n$ gilt

$$\prod_{1\leqslant i<j\leqslant n} (\sigma j - \sigma i) \;=\; \operatorname{sgn}(\sigma) \cdot \prod_{1\leqslant i<j\leqslant n} (j - i).$$

Proposition: Für alle $\sigma, \tau \in S_n$ gilt:

$$\begin{aligned} \operatorname{sgn}(\mathrm{id}) &= 1 \\ \operatorname{sgn}(\sigma \circ \tau) &= \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau) \\ \operatorname{sgn}(\sigma^{-1}) &= \operatorname{sgn}(\sigma) \end{aligned}$$

Das bedeutet, dass die Abbildung $\operatorname{sgn}\colon S_n \to \{\pm 1\}$ ein Gruppenhomomorphismus ist.

Definition: Für jedes $\sigma \in S_n$ betrachte die $n \times n$-Matrix

$$P_\sigma \;:=\; \big(\delta_{i,\sigma j}\big)_{1\leqslant i,j\leqslant n}.$$

Proposition: Die Matrix P_σ ist eine Permutationsmatrix. Umgekehrt ist jede $n \times n$-Permutationsmatrix gleich P_σ für genau ein $\sigma \in S_n$. Ausserdem gilt für alle $\sigma, \tau \in S_n$

$$P_{\sigma\tau} \;=\; P_\sigma \cdot P_\tau.$$

Das bedeutet, dass $\sigma \mapsto P_\sigma$ einen Gruppenisomorphismus von S_n auf die Gruppe aller $n \times n$-Permutationsmatrizen induziert.

6.2 Konstruktion und Grundeigenschaften

In diesem Kapitel rechnen wir in einem beliebigen kommutativen unitären Ring R.

Definition: Die *Determinante* einer $n \times n$-Matrix $A = \big(a_{ij}\big)_{1 \leqslant i,j \leqslant n}$ ist

$$\det(A) \;:=\; \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot \prod_{i=1}^{n} a_{i,\sigma i}.$$

Beispiel: Explizite Formeln für $n = 0, 1, 2, 3$.

Proposition: Die Determinante ist eine lineare Abbildung jeder einzelnen Zeile, das heisst:

(a) Stimmen A, A', A'' ausserhalb der i-ten Zeile überein, und ist die i-te Zeile von A'' die Summe der i-ten Zeilen von A und A', so gilt $\det(A'') = \det(A) + \det(A')$.

(b) Entsteht A' aus A durch Multiplikation einer Zeile mit einem Skalar $\lambda \in R$, so gilt $\det(A') = \lambda \cdot \det(A)$.

Dieselben Aussagen gelten für Spalten anstelle von Zeilen.

Proposition: Es gilt $\det(A^T) = \det(A)$.

Proposition: Für jede Permutation $\sigma \in S_n$ gilt $\det(P_\sigma) = \operatorname{sgn}(\sigma)$.

Proposition: Für jede Blockdreiecksmatrix der Form

$$A \;=\; \left(\begin{array}{c|c} A' & B \\ \hline O & A'' \end{array}\right) \quad \text{oder} \quad \left(\begin{array}{c|c} A' & O \\ \hline B' & A'' \end{array}\right)$$

mit quadratischen Matrizen A' und A'' und der jeweiligen Nullmatrix O gilt

$$\det(A) = \det(A') \cdot \det(A'').$$

Proposition: Für jede obere oder untere Dreiecksmatrix $A = \big(a_{ij}\big)_{1 \leqslant i,j \leqslant n}$ gilt

$$\det(A) = \prod_{i=1}^{n} a_{ii}.$$

Folge: Insbesondere gilt $\det(I_n) = 1$.

Satz: Für je zwei $n \times n$-Matrizen A und B gilt $\det(AB) = \det(A) \cdot \det(B)$.

6.3 Berechnung der Determinante

Proposition: Die Determinante verhält sich unter elementaren Zeilenoperationen wie folgt: Entsteht A' aus A durch ...

(a) Addition eines Vielfachen einer Zeile zu einer anderen, so gilt $\det(A') = \det(A)$.

(b) Multiplikation einer Zeile mit einem Skalar $\lambda \in R$, so gilt $\det(A') = \lambda \cdot \det(A)$.

(c) Vertauschen zweier Zeilen, so gilt $\det(A') = -\det(A)$.

Die entsprechenden Aussagen gelten für elementare Spaltenoperationen.

Über einem Körper lässt sich die Determinante daher mit Gauss-Elimination berechnen.

Beispiel: Ist eine Zeile oder Spalte von A eine Linearkombination der übrigen, so gilt $\det(A) = 0$.

Beispiel: Für beliebige $a_1, \ldots, a_n \in R$ hat die $n \times n$-Matrix

$$A := \left(a_i^{j-1}\right)_{1 \leqslant i,j \leqslant n} = \begin{pmatrix} 1 & a_1 & a_1^2 & \ldots & a_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \ldots & a_n^{n-1} \end{pmatrix}$$

die *Vandermonde-Determinante*

$$\det(A) = \prod_{1 \leqslant i < j \leqslant n} (a_j - a_i).$$

6.4 Zeilen- und Spaltenentwicklung

Konstruktion: Sei A eine $n \times n$-Matrix mit $n > 0$. Für jedes Paar von Indizes $1 \leqslant i, j \leqslant n$ sei A_{ij} diejenige $(n-1) \times (n-1)$-Matrix, die durch Streichen der i-ten Zeile und der j-ten Spalte aus A entsteht.

Satz: Für jedes $1 \leqslant i \leqslant n$ gilt

$$\det(A) \;=\; \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}).$$

Für jedes $1 \leqslant j \leqslant n$ gilt

$$\det(A) \;=\; \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}).$$

Definition: Die *Adjunkte* von A ist die Matrix

$$\tilde{A} \;:=\; \big((-1)^{i+j} \cdot \det(A_{ji})\big)_{1 \leqslant i,j \leqslant n}.$$

Satz: Es gilt

$$A \cdot \tilde{A} \;=\; \tilde{A} \cdot A \;=\; \det(A) \cdot I_n.$$

Satz: Eine quadratische Matrix A über einem Körper ist invertierbar genau dann, wenn $\det(A) \neq 0$ ist. In diesem Fall gilt weiter

$$\begin{aligned} \det(A^{-1}) &= \det(A)^{-1} \qquad \text{und} \\ A^{-1} &= \det(A)^{-1} \cdot \tilde{A}. \end{aligned}$$

Beispiel: Für jede invertierbare 2×2-Matrix gilt

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \;=\; \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

6.5 Ähnlichkeit und Determinante eines Endomorphismus

Sei K ein Körper.

Definition: Zwei $n \times n$-Matrizen A und B über K heissen *ähnlich* oder *zueinander konjugiert*, wenn eine invertierbare $n \times n$-Matrix U über K existiert mit $B = UAU^{-1}$.

Proposition: Dies ist eine Äquivalenzrelation.

Proposition: Ähnliche Matrizen haben dieselbe Determinante.

Proposition: Sei f ein Endomorphismus eines endlich-dimensionalen K-Vektorraums V. Dann sind die Darstellungsmatrizen $M_{B,B}(f)$ bezüglich beliebiger geordneter Basen B von V zueinander ähnlich. Insbesondere ist $\det(M_{B,B}(f))$ unabhängig von B.

Definition: Das Element $\det(f) := \det M_{B,B}(f) \in K$ heisst die *Determinante von* f.

Proposition: Für alle Endomorphismen f und g eines endlich-dimensionalen K-Vektorraums V gilt:

(a) $\det(\mathrm{id}_V) = 1$.
(b) $\det(g \circ f) = \det(g) \cdot \det(f)$.
(c) Die Abbildung f ist ein Automorphismus genau dann, wenn $\det(f) \neq 0$ ist.
(d) Für jeden Automorphismus f gilt $\det(f^{-1}) = \det(f)^{-1}$.
(e) Für jeden Isomorphismus $h\colon V \xrightarrow{\sim} W$ gilt $\det(h \circ f \circ h^{-1}) = \det(f)$.

7 Polynome

7.1 Polynome einer Variablen

Sei R ein kommutativer unitärer Ring, und sei X ein noch nicht verwendetes Symbol.

Definition: Eine Abbildung der Form $f\colon R \to R$, $x \mapsto f(x) = \sum'_{i\geqslant 0} a_i x^i$, mit Koeffizienten $a_i \in R$ und fast allen $a_i = 0$, heisst eine *Polynomfunktion*.

Definition: Ein „formaler Ausdruck“ der Form $F(X) = \sum'_{i\geqslant 0} a_i X^i$, mit Koeffizienten $a_i \in R$ und fast allen $a_i = 0$, heisst ein *Polynom in X über R*. Für jedes solche F, jeden kommutativen unitären Ring R', der R enthält, und jedes Element $x \in R'$, ist der *Wert von F an der Stelle x* das Element

$$F(x) \;:=\; \sum_{i\geqslant 0}{}' a_i x^i \;\in R'.$$

Insbesondere induziert F eine Polynomfunktion $R \to R$, $x \mapsto F(x)$.

Beispiel: Ein Polynom mit reellen Koeffizienten kann man auf ganz $\mathbb{C}$ auswerten.

Vorsicht: Ist R endlich, so können verschiedene Polynome über R dieselbe Polynomfunktion $R \to R$ induzieren, z.B. die beiden Polynome 0 und $X^2 - X$ über $\mathbb{F}_2$.

Definition: Für je zwei Polynome $F(X) = \sum'_{i\geqslant 0} a_i X^i$ und $G(X) = \sum'_{i\geqslant 0} b_i X^i$ und jedes $\lambda \in R$ setzen wir

$$\begin{array}{rcl} (F+G)(X) & := & F(X)+G(X) \;:=\; \sum'_{i\geqslant 0} (a_i+b_i)\cdot X^i, \\ (F\cdot G)(X) & := & F(X)\cdot G(X) \;:=\; \sum'_{i\geqslant 0} \big(\sum_{j=0}^{i} a_{i-j}b_j\big)\cdot X^i \\ (\lambda\cdot F)(X) & := & \lambda\cdot F(X) \;:=\; \sum'_{i\geqslant 0} \lambda a_i X^i. \end{array}$$

Proposition: Diese Operationen sind verträglich mit der Auswertung, das heisst, für alle $x \in R'$ wie oben gilt $(F+G)(x) = F(x)+G(x)$ und so weiter.

Proposition-Definition: Die Menge $R[X]$ aller Polynome in X über R mit den obigen Operationen, dem *Nullpolynom* $0 = \sum'_{i\geqslant 0} 0\cdot X^i$ und dem *Einspolynom* $1 = \sum'_{i\geqslant 0} \delta_{i,0}\cdot X^i$ ist ein kommutativer unitärer Ring, genannt der *Polynomring in X über R*.

Konstruktion: Ein Polynom $\sum'_{i\geqslant 0} a_i X^i$ anzugeben ist nach Definition äquivalent dazu, seine Koeffizienten anzugeben. Man kann den Polynomring daher konkret realisieren als die Menge aller Folgen $(a_i)_{i\geqslant 0}$ in R, die schliesslich Null werden, mit komponentenweiser Addition und skalarer Multiplikation sowie dem Produkt $(a_i)_i \cdot (b_i)_i := \big(\sum_{j=0}^{i} a_{i-j}b_j\big)_i$. Dabei repräsentieren die Folge $(0,0,\ldots)$ das Nullelement, die Folge $(1,0,0,\ldots)$ das Einselement, und die Folge $(0,1,0,0,\ldots)$ das Variablensymbol X.

7.2 Grad eines Polynoms

Definition: Der *Grad* eines Polynom $F(X) = \sum'_{i\geqslant 0} a_i X^i$ über einem Körper K ist

$$\deg(F) := \begin{cases} \max\{i \geqslant 0 \mid a_i \neq 0\} & \text{falls } F \neq 0, \\ -\infty & \text{falls } F = 0. \end{cases}$$

Ein Polynom vom Grad $\leqslant 0$ heisst *konstant*, eines vom Grad > 0 heisst *nichtkonstant*. Ein Polynom vom Grad $\leqslant 1$ heisst *linear*, eines vom Grad $\leqslant 2$ *quadratisch*, eines vom Grad $\leqslant 3$ *kubisch*. Ein Polynom vom Grad $n \geqslant 0$ mit höchstem Koeffizienten $a_n = 1$ heisst *normiert*. Der Koeffizient a_0 von $X^0 = 1$ in F heisst der *konstante Koeffizient von F*.

Proposition: Für je zwei Polynome F und G über K und jedes $\lambda \in K^\times$ gilt:

$$\begin{aligned} \deg(F+G) &\leqslant \max\{\deg(F), \deg(G)\}, \\ \deg(F \cdot G) &= \deg(F) + \deg(G), \\ \deg(\lambda \cdot F) &= \deg(F). \end{aligned}$$

Proposition: Für je drei Polynome F, G, H über K mit $H \neq 0$ gilt:

$$\begin{aligned} F \cdot G = 0 &\Leftrightarrow (F = 0) \vee (G = 0), \\ FH = GH &\Leftrightarrow F = G. \end{aligned}$$

Proposition: Der Ring $K[X]$ ist ein Vektorraum über K, isomorph zu dem Folgenraum F_0 aus §4.2. Eine Basis von $K[X]$ ist $\{X^i \mid i \geqslant 0\}$. Die Polynome vom Grad $\leqslant n$ bilden einen Unterraum der Dimension $n+1$ mit Basis $\{X^i \mid 0 \leqslant i \leqslant n\}$.

7.3 Nullstellen

Satz: Für je zwei Polynome $F(X)$ und $G(X) \in K[X]$ mit $G \neq 0$ existieren eindeutige Polynome $Q(X)$ und $R(X) \in K[X]$ mit $\deg(R) < \deg(G)$ und $F = Q \cdot G + R$.

Diese *Polynomdivision mit Rest* erfolgt von Hand wie die Division ganzer Zahlen.

Sei nun $F(X) \in K[X]$ beliebig.

Definition: Ein Element $\lambda \in K$ mit $F(\lambda) = 0$ heisst eine *Nullstelle von F*.

Proposition: Ein Element $\lambda \in K$ ist eine Nullstelle von F genau dann, wenn ein Polynom G über K existiert mit $F(X) = (X - \lambda) \cdot G(X)$.

Definition: Für jedes $\lambda \in K$ heisst

$$\mu_\lambda \;:=\; \sup\big\{\, m \geqslant 0 \;\big|\; \exists\, G \in K[X] \colon F(X) = (X-\lambda)^m \cdot G(X) \,\big\} \;\in\; \mathbb{Z}^{\geqslant 0} \cup \{\infty\}$$

die *Nullstellenordnung von F in λ*. Ist $\mu_\lambda > 0$, so heisst λ eine *Nullstelle der Vielfachheit* oder *Multiplizität μ_λ von F*.

Bemerkung: Es gilt $\mu_\lambda \leqslant \deg(F)$ falls $F \neq 0$ ist, und $\mu_\lambda = \infty$ falls $F = 0$ ist.

Satz: Jedes von Null verschiedene Polynom F über K lässt sich schreiben in der Form

$$F(X) \;=\; (X - \lambda_1)^{\mu_1} \cdots (X - \lambda_r)^{\mu_r} \cdot G(X)$$

mit geeigneten $r \geqslant 0$ und $\mu_i \geqslant 1$ und paarweise verschiedenen $\lambda_i \in K$, sowie einem Polynom G über K ohne Nullstellen in K. Dabei sind $\lambda_1, \ldots, \lambda_r$ genau die Nullstellen von F in K und $\mu_1, \ldots, \mu_r$ deren jeweilige Nullstellenordnung. Ausserdem sind r und G, sowie die Paare (λ_i, μ_i) bis auf Permutation der i, eindeutig bestimmt.

Definition: Ist in der obigen Zerlegung G eine Konstante in $K^\times$, so sagen wir, dass *F über K in Linearfaktoren zerfällt.*

Folge: Für jedes von Null verschiedene Polynom F über K ist die Anzahl der Nullstellen von F in K, mit Vielfachheiten gezählt, kleiner oder gleich $\deg(F)$.

Proposition: Ist K ein unendlicher Körper, so ist jedes Polynom F über K durch die Polynomfunktion $K \to K$, $x \mapsto F(x)$ eindeutig bestimmt.

Satz: Für jedes normierte Polynom $F(X) \in \mathbb{Z}[X]$ gilt: Jede Nullstelle von F in $\mathbb{Q}$ liegt schon in $\mathbb{Z}$ und teilt den konstanten Koeffizienten von F.

Folge: Es gibt einen effektiven Algorithmus, der für jedes Polynom in $\mathbb{Q}[X]$ alle Nullstellen in $\mathbb{Q}$ bestimmt.

7.4 Algebraisch abgeschlossene Körper

Satz: Für jeden Körper K sind äquivalent:

(a) Jedes nichtkonstante Polynom über K besitzt eine Nullstelle in K.

(b) Jedes Polynom über K zerfällt in Linearfaktoren über K.

(c) Jedes Polynom vom Grad $n \geqslant 0$ über K besitzt, mit Vielfachheiten gezählt, genau n Nullstellen in K.

Definition: Ein Körper mit den obigen Eigenschaften heisst *algebraisch abgeschlossen*.

Beispiel: Ist K ein endlicher Körper, so ist $F(X) := 1+\prod_{\lambda\in K}(X-\lambda)$ ein Polynom vom Grad $|K| > 0$ über K ohne Nullstellen in K. Also ist K nicht algebraisch abgeschlossen.

Satz: Jeder Körper ist in einem algebraisch abgeschlossenen Körper enthalten.

(Beweis in der Vorlesung Algebra II.)

Fundamentalsatz für die Algebra: Der Körper $\mathbb{C}$ ist algebraisch abgeschlossen.

Beispiel: Der Körper $\mathbb{R}$ ist nicht algebraisch abgeschlossen.

Proposition: Für jedes Polynom F über $\mathbb{R}$ und jede Nullstelle $\lambda \in \mathbb{C}$ von F ist die komplex Konjugierte $\overline{\lambda}$ eine Nullstelle von F derselben Multiplizität wie λ.

Satz: Jedes Polynom über $\mathbb{R}$ ist ein Produkt von Polynomen vom Grad $\leqslant 2$ über $\mathbb{R}$.

7.5 Irreduzible Polynome

Sei K ein Körper.

Definition: Ein Polynom in $K[X]$, welches sich als Produkt zweier Polynome in $K[X]$ vom Grad $\geqslant 1$ schreiben lässt, heisst *reduzibel*. Ein Polynom in $K[X]$ vom Grad $\geqslant 1$, welches nicht reduzibel ist, heisst *irreduzibel*.

Beispiel: Jedes Polynom vom Grad 1 ist irreduzibel.

Beispiel: Die irreduziblen Polynome in $\mathbb{R}[X]$ sind genau

(a) die Polynome vom Grad 1, also $aX + b$ für alle $a, b \in \mathbb{R}$ mit $a \neq 0$, sowie

(b) die Polynome vom Grad 2 ohne reelle Nullstellen, also $aX^2 + bX + c$ für alle $a, b, c \in \mathbb{R}$ mit $b^2 - 4ac < 0$.

Beispiel: (a) Das Polynom $X^2 + 1$ ist irreduzibel in $\mathbb{R}[X]$ und reduzibel in $\mathbb{C}[X]$.

(b) Das Polynom $X^2 - 2$ ist irreduzibel in $\mathbb{Q}[X]$ und reduzibel in $\mathbb{R}[X]$.

(c) Das Polynom $X^7 - 3X^2 + 12$ ist irreduzibel in $\mathbb{Q}[X]$ und reduzibel in $\mathbb{R}[X]$.

(d) Das Polynom $X^{65536} + X^{65535} + \ldots + X^2 + X + 1$ ist irreduzibel in $\mathbb{Q}[X]$.

(e) Das Polynom $X^3 + X + 1$ ist irreduzibel in $\mathbb{F}_2[X]$.

Proposition: Für jedes irreduzible Polynom $p(X) \in K[X]$ und je zwei Polynome φ', $\varphi'' \in K[X]$ gilt:

$$p \text{ teilt } \varphi'\varphi'' \implies p \text{ teilt } \varphi' \text{ oder } \varphi''.$$

Beweis siehe Algebra I.

Satz: Jedes normierte Polynom in $K[X]$ ist ein Produkt von normierten irreduziblen Polynomen in $K[X]$, und diese sind bis auf Vertauschung eindeutig bestimmt.

Bemerkung: Ob und wie man diese Faktorisierung konkret bestimmen kann, hängt vom Körper K ab. Ist K endlich, so gibt es überhaupt nur endlich viele Polynome von kleinerem Grad, und es genügt, jedes davon auf Teilbarkeit zu überprüfen. Für ein effektives Verfahren z.B. im Fall $K = \mathbb{Q}$ siehe Algebra II.

8 Endomorphismen

8.1 Charakteristisches Polynom

Definition: Das *charakteristische Polynom* einer $n \times n$-Matrix A über K ist

$$\mathrm{char}_A(X) \;:=\; \det(X \cdot I_n - A) \in K[X].$$

Proposition: Das Polynom $\mathrm{char}_A(X)$ ist normiert vom Grad n, und sein konstanter Koeffizient ist $(-1)^n \cdot \det(A)$.

Proposition: Je zwei ähnliche Matrizen über K haben dasselbe charakteristische Polynom.

Sei nun f ein Endomorphismus eines endlich-dimensionalen K-Vektorraums V, und sei B eine geordnete Basis von V.

Proposition: Das charakteristische Polynom von $M_{B,B}(f)$ ist unabhängig von B.

Definition: Dieses heisst das *charakteristische Polynom von* f, bezeichnet mit

$$\mathrm{char}_f(X) \;:=\; \mathrm{char}_{M_{B,B}(f)}(X).$$

Dies ist ein normiertes Polynom vom Grad $\dim_K(V)$ über K, und sein konstanter Koeffizient ist $(-1)^n \cdot \det(f)$.

Proposition: Für jeden Isomorphismus von Vektorräumen $\varphi\colon V \xrightarrow{\sim} W$ gilt

$$\mathrm{char}_{\varphi \circ f \circ \varphi^{-1}}(X) \;=\; \mathrm{char}_f(X).$$

Proposition: Für jedes $\lambda \in K$ gilt

$$\mathrm{char}_f(\lambda) \;=\; \det(\lambda \cdot \mathrm{id}_V - f).$$

8.2 Eigenwerte und Eigenvektoren

Sei f ein Endomorphismus eines K-Vektorraums V.

Definition: Ein Vektor $v \in V$ heisst *Eigenvektor von f zum Eigenwert $\lambda \in K$*, falls $v \neq 0$ ist und $f(v) = \lambda \cdot v$. Ein Element $\lambda \in K$, welches als Eigenwert zu einem geeigneten Eigenvektor von f auftritt, heisst schlechthin ein *Eigenwert von f*.

Proposition-Definition: Ein Element $\lambda \in K$ ist Eigenwert von f genau dann, wenn der Endomorphismus $\lambda \cdot \mathrm{id}_V - f \colon V \to V$ nicht injektiv ist. Der nichttriviale Unterraum

$$\mathrm{Eig}_{\lambda,f} \;:=\; \mathrm{Kern}(\lambda \cdot \mathrm{id}_V - f) \subset V$$

heisst dann der zu λ gehörende *Eigenraum von f*. Seine von Null verschiedenen Elemente sind genau die Eigenvektoren von f zum Eigenwert λ.

Definition: Die *geometrische Vielfachheit* eines Eigenwerts $\lambda \in K$ von f ist die Dimension des zugehörigen Eigenraums $\mathrm{Eig}_{\lambda,f}$.

Proposition: Seien $\lambda_1, \ldots, \lambda_r \in K$ paarweise verschiedene Eigenwerte von f. Dann ist die folgende lineare Abbildung injektiv:

$$\mathrm{Eig}_{\lambda_1,f} \times \ldots \times \mathrm{Eig}_{\lambda_r,f} \longrightarrow V, \quad (v_1, \ldots, v_r) \mapsto v_1 + \ldots + v_r.$$

Ab jetzt sei V endlich-dimensional.

Satz: Die Eigenwerte von f sind genau die Nullstellen von $\mathrm{char}_f(X)$ in K.

Folge: Ist K algebraisch abgeschlossen und $V \neq 0$, so besitzt jeder Endomorphismus von V mindestens einen Eigenwert.

Definition: Die *arithmetische Vielfachheit* eines Eigenwerts $\lambda \in K$ von f ist die Vielfachheit von λ als Nullstelle von $\mathrm{char}_f(X)$.

Folge: Die Anzahl der Eigenwerte von f, mit arithmetischen Vielfachheiten gezählt, ist $\leqslant \dim_K(V)$, und sogar $= \dim_K(V)$ wenn K algebraisch abgeschlossen ist.

Nach Konstruktion sind die geometrische Vielfachheit und die arithmetische Vielfachheit jedes Eigenwerts > 0.

Satz: Die geometrische Vielfachheit ist stets $\leqslant$ der arithmetischen Vielfachheit.

Definition: Die Eigenwerte, Eigenvektoren, und Eigenräume einer $n \times n$-Matrix A über K sind die Eigenwerte, Eigenvektoren, bzw. Eigenräume des Endomorphismus $L_A \colon K^n \to K^n$.

8.3 Diagonalisierbarkeit

Sei weiterhin f ein Endomorphismus eines endlich-dimensionalen K-Vektorraums V.

Proposition: Für jede geordnete Basis B von V sind äquivalent:

(a) Die Darstellungsmatrix $M_{B,B}(f)$ ist eine Diagonalmatrix.

(b) Die Basis B besteht aus Eigenvektoren von f.

Definition: Besitzt V eine solche Basis B, so heisst f *diagonalisierbar*.

Satz: Ein Endomorphismus f ist diagonalisierbar genau dann, wenn $\mathrm{char}_f(X)$ über K in Linearfaktoren zerfällt und für alle Eigenwerte von f die geometrische Vielfachheit gleich der arithmetischen Vielfachheit ist.

Folge: Zerfällt $\mathrm{char}_f(X)$ über K in Linearfaktoren und haben alle seine Nullstellen die Vielfachheit 1, so ist f diagonalisierbar.

Proposition: Für jede $n \times n$-Matrix A über K sind äquivalent:

(a) Der Endomorphismus $L_A\colon K^n \to K^n$ ist diagonalisierbar.

(b) Die Matrix A ist ähnlich über K zu einer Diagonalmatrix.

Für eine invertierbare Matrix U ist dann $A = UDU^{-1}$ eine Diagonalmatrix genau dann, wenn die Spalten von U eine Basis aus Eigenvektoren von A bilden.

Definition: Eine solche quadratische Matrix A heisst *diagonalisierbar über K*.

Beispiel: Die reelle Matrix $A := \left(\begin{smallmatrix} 5 & 3 \\ 3 & 5 \end{smallmatrix}\right)$ hat charakteristisches Polynom $X^2 - 10X + 16 = (X-2)(X-8)$ und somit zwei verschiedene Eigenwerte der arithmetischen Vielfachheit 1 in $\mathbb{R}$. Also ist A diagonalisierbar über $\mathbb{R}$.

Beispiel: Die reelle Matrix $B := \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ repräsentiert eine Drehung um 90° in der Ebene $\mathbb{R}^2$. Ihr charakteristisches Polynom ist $X^2 + 1$ und hat somit keine Nullstellen in $\mathbb{R}$. Also ist B nicht diagonalisierbar über $\mathbb{R}$. Über $\mathbb{C}$ hat das charakterische Polynom jedoch zwei verschiedene einfache Nullstellen $\pm i$. Also ist B diagonalisierbar über $\mathbb{C}$.

Beispiel: Die Matrix $C := \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ über einem beliebigen Körper hat charakteristisches Polynom $(X-1)^2$ und somit nur den Eigenwert 1 der arithmetischen Vielfachheit 2. Der zugehörige Eigenraum $\langle \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) \rangle$ hat aber Dimension 1, also ist die geometrische Vielfachheit gleich 1. Somit ist C nicht diagonalisierbar.

Anwendung: Sei A diagonalisierbar, also $A = UDU^{-1}$ für eine invertierbare Matrix U und eine Diagonalmatrix D mit Diagonaleinträgen $\lambda_1, \ldots, \lambda_n$. Dann sind die Potenzen von A gegeben durch:

$$A^m \;=\; UD^mU^{-1} \;=\; \sum_{i=1}^{n} \lambda_i^m \cdot UE_{ii}U^{-1}.$$

8.4 Minimalpolynom

Sei weiterhin f ein Endomorphismus eines K-Vektorraums V.

Definition: Für alle natürliche Zahlen i sind die Endomorphismen $f^i \in \mathrm{End}_K(V)$ definiert durch $f^0 := \mathrm{id}_V$ und $f^{i+1} := f \circ f^i$. Für jedes Polynom $\varphi(X) = \sum'_{i\geqslant 0} a_i X^i \in K[X]$ ist der Endomorphismus $\varphi(f) \in \mathrm{End}_K(V)$ definiert durch

$$\varphi(f) := \sum_{i\geqslant 0}{}' a_i f^i \colon\ V \to V,\ v \mapsto \sum_{i\geqslant 0}{}' a_i f^i(v).$$

Grundeigenschaften: Für alle $i, j \geqslant 0$ und $\varphi, \psi \in K[X]$ und $a \in K$ gilt:

(a) $f^i \circ f^j = f^{i+j}$.
(b) $(a\varphi)(f) = a\varphi(f)$.
(c) $(\varphi + \psi)(f) = \varphi(f) + \psi(f)$.
(d) $(\varphi\psi)(f) = \varphi(f) \circ \psi(f)$.

Proposition: Für jedes normierte Polynom $\varphi(X) \in K[X]$ sind äquivalent:

(a) Es ist $\varphi(f) = 0$, und für alle $\psi(X) \in K[X]$ mit $\psi(f) = 0$ gilt $\varphi|\psi$.
(b) Es ist $\varphi(f) = 0$, und für alle $\psi(X) \in K[X]\smallsetminus\{0\}$ mit $\psi(f) = 0$ gilt $\deg\varphi \leqslant \deg\psi$.

Ausserdem ist ein solches φ eindeutig bestimmt, wenn es existiert.

Definition: Dieses φ heisst das *Minimalpolynom von f*.

Definition: Für jede quadratische Matrix A über K definieren wir genauso

$$\varphi(A) := \sum_{i\geqslant 0}{}' a_i A^i \quad \text{falls} \quad \varphi(X) = \sum_{i\geqslant 0}{}' a_i X^i$$

sowie das Minimalpolynom von A, mit den entsprechenden Eigenschaften. Dann gilt insbesondere $L_{\varphi(A)} = \varphi(L_A)$, und das Minimalpolynom von A ist gleich dem von L_A.

Proposition: Ist $\dim_K(V) < \infty$, so existiert das Minimalpolynom von f.

Beispiel: (a) Die Matrix $A := \left(\begin{smallmatrix} 2 & 1 \\ 3 & 2 \end{smallmatrix}\right)$ hat Minimalpolynom $X^2 - 4X + 1$.

(b) Die Matrix $A := \left(\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right)$ hat Minimalpolynom $X - 2$.

(c) Die Matrix $A := \left(\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right)$ hat Minimalpolynom $(X - 2)^2$.

Proposition: Für jeden Eigenwert $\lambda \in K$ von f gilt $\varphi(f) = 0 \Rightarrow \varphi(\lambda) = 0$.

Folge: Ein Endomorphismus eines unendlich-dimensionalen Vektorraums mit unendlich vielen Eigenwerten besitzt kein Minimalpolynom.

8.5 Satz von Cayley-Hamilton

Ab jetzt sei zusätzlich $\dim_K(V) < \infty$.

Proposition: Ist f diagonalisierbar mit den paarweise verschiedenen Eigenwerten λ_i der Multiplizität m_i für $i = 1, \ldots, r$, so hat f das

$$\begin{array}{rl} \text{charakteristische Polynom} & \prod_{i=1}^{r}(X-\lambda_i)^{m_i} \\ \text{Minimal-Polynom} & \prod_{i=1}^{r}(X-\lambda_i). \end{array}$$

Beispiel: Besitzt V eine Basis B mit Darstellungsmatrix der Form

$$M_{BB}(f) \;=\; \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

so hat f charakteristisches und Minimal-Polynom gleich $(X-\lambda)^n$.

Satz: *(Cayley-Hamilton)* Für das charakteristische Polynom $\mathrm{char}_f(X)$ von f gilt

$$\mathrm{char}_f(f) = 0.$$

Folge: Das Minimalpolynom von f teilt das charakteristische Polynom von f.

Folge: Das Minimalpolynom von f hat dieselben Nullstellen wie das charakteristische Polynom von f.

Proposition: Für jede invertierbare Matrix A mit charakteristischem oder Minimal-Polynom $\sum'_{i\geqslant 0} a_i X^i$ ist $a_0 \neq 0$ und

$$A^{-1} \;=\; -\sum_{i\geqslant 1}{}' \frac{a_i}{a_0} A^{i-1}.$$

Beispiel: Für die Matrix $A := \left(\begin{smallmatrix} 2 & 1 \\ 3 & 2 \end{smallmatrix}\right)$ gilt $A^{-1} = 4I_2 - A$.

8.6 Blocktrigonalisierung

Proposition: Für das charakteristische Polynom einer Blockdreiecksmatrix der Form

$$A = \begin{pmatrix} A_1 & * & \dots & * \\ & \ddots & \ddots & \vdots \\ & & \ddots & * \\ & & & A_r \end{pmatrix}$$

gilt

$$\operatorname{char}_A(X) = \prod_{i=1}^{r} \operatorname{char}_{A_i}(X).$$

Definition: Ein Unterraum U mit $f(U) \subset U$ heisst *f-invariant*.

Proposition: Ist $\dim_K(V) \geqslant 2$, so sind die folgenden Aussagen äquivalent:

(a) Der Endomorphismus f ist *blocktrigonalisierbar*, das heisst, es existiert eine geordnete Basis $B = (b_1, \dots, b_n)$ von V, so dass die Darstellungsmatrix $M_{B,B}(f)$ eine Blockdreiecksmatrix der Form $\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)$ ist für eine Zerlegung $n = n_1 + n_2$ mit $n_1, n_2 \geqslant 1$.

(b) Es existiert ein f-invarianter Unterraum $\{0\} \neq U \subsetneqq V$.

(c) Das charakteristische Polynom $\operatorname{char}_f(X)$ ist reduzibel in $K[X]$.

Definition: Die *Begleitmatrix* eines normierten Polynoms

$$\varphi(X) = X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0$$

ist die folgende $n \times n$-Matrix:

$$\begin{pmatrix} -a_{n-1} & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ -a_0 & 0 & \dots & \dots & 0 \end{pmatrix}$$

(Oft nimmt man auch die Transponierte, oder die durch Umkehren der Reihenfolge der Zeilen sowie der Spalten entstehende Matrix, oder die Transponierte davon.)

Proposition: Die Begleitmatrix von φ hat das charakteristische Polynom φ.

Proposition: Ist das charakteristische Polynom $\operatorname{char}_f(X)$ irreduzibel in $K[X]$, so existiert eine geordnete Basis B von V, so dass die Darstellungsmatrix $M_{B,B}(f)$ die Begleitmatrix von $\operatorname{char}_f(X)$ ist.

Satz: Es existiert eine geordnete Basis B von V, so dass die Darstellungsmatrix $M_{B,B}(f)$ die Blockdreiecksgestalt

$$\begin{pmatrix} A_1 & * & \dots & * \\ & \ddots & \ddots & \vdots \\ & & \ddots & * \\ & & & A_r \end{pmatrix}$$

hat, wobei jedes A_i die Begleitmatrix eines irreduziblen Polynoms in $K[X]$ ist. Dabei sind die A_i bis auf Vertauschung durch f eindeutig bestimmt.

Bemerkung: Im Fall $K = \mathbb{R}$ hat dann jeder Block die Grösse 1 oder 2; siehe §7.5.

8.7 Trigonalisierbarkeit

Definition: Eine Menge $\mathcal{F}$ von Teilräumen von V, die durch Inklusion total geordnet ist, heisst eine *Fahne* oder *Flagge*.

(b) Eine Fahne, die für jedes $0 \leqslant m \leqslant \dim_K(V) < \infty$ einen Unterraum der Dimension m enthält, heisst *vollständig* oder *maximal*.

(c) Eine Fahne, welche aus f-invarianten Unterräumen besteht, heisst *f-invariant*.

Satz: Die folgenden Aussagen sind äquivalent:

(a) Der Endomorphismus f ist *trigonalisierbar*, das heisst, es existiert eine geordnete Basis B von V, so dass die Darstellungsmatrix $M_{B,B}(f)$ eine obere Dreiecksmatrix ist.

(b) Es existiert eine f-invariante vollständige Fahne von V.

(c) Das charakteristische Polynom $\mathrm{char}_f(X)$ zerfällt in Linearfaktoren in $K[X]$.

Folge: Ist K algebraisch abgeschlossen, so ist jeder Endomorphismus eines endlichdimensionalen K-Vektorraums trigonalisierbar.

Beispiel: Für beliebige $a, b, c \in \mathbb{Q}$ mit $b \neq -c$ gilt:

$$\begin{pmatrix} 1 & 1 & a \\ -1 & 0 & b \\ 1 & 0 & c \end{pmatrix}^{-1} \cdot \begin{pmatrix} -1 & -3 & -4 \\ -1 & 0 & 3 \\ 1 & -2 & -5 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & a \\ -1 & 0 & b \\ 1 & 0 & c \end{pmatrix} = \begin{pmatrix} -2 & * & * \\ 0 & -2 & * \\ 0 & 0 & -2 \end{pmatrix}$$

Beispiel: Über dem Körper $\mathbb{F}_2$ mit 2 Elementen gilt:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

8.8 Hauptraumzerlegung

Proposition: Das Minimalpolynom und das charakteristische Polynom von f haben dieselben irreduziblen Faktoren.

Schreibe nun

$$\mathrm{char}_f(X) \;=\; \prod_{i=1}^{r} p_i(X)^{m_i}$$

mit verschiedenen normierten irreduziblen $p_i(X) \in K[X]$ und Exponenten $m_i \geqslant 1$.

Definition: Für jedes i heisst der Unterraum

$$\mathrm{Hau}_{p_i}(f) \;:=\; \mathrm{Kern}(p_i(f)^{m_i})$$

der *Hauptraum* oder *verallgemeinerte Eigenraum* von f zum irreduziblen Faktor $p_i(X)$.

Bemerkung: Für jeden Eigenwert λ von f gilt $\mathrm{Eig}_\lambda(f) \subset \mathrm{Hau}_{X-\lambda}(f)$.

Satz: Jeder Hauptraum ist f-invariant, und die Einschränkung $f|_{\mathrm{Hau}_{p_i}(f)}$ hat das charakteristische Polynom $p_i(X)^{m_i}$. Ausserdem gilt

$$V \;=\; \bigoplus_{i=1}^{r} \mathrm{Hau}_{p_i}(f).$$

Lemma: Seien $\varphi(X)$, $\psi(X) \in K[X]$ teilerfremd mit $(\varphi\psi)(f) = 0$. Dann gilt

$$V \;=\; \mathrm{Kern}(\varphi(f)) \oplus \mathrm{Kern}(\psi(f)),$$

und beide Summanden sind f-invariant.

Beispiel: Die reelle Matrix

$$A := \begin{pmatrix} 1 & 1 & 4 & 5 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

hat charakteristisches Polynom $(X-0)^2(X-1)^2$ und die Haupträume

$$\mathrm{Hau}_{X-0}(L_A) = \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -7 \\ 0 \\ 3 \\ -1 \end{pmatrix} \right\rangle, \qquad \mathrm{Hau}_{X-1}(L_A) = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \\ 0 \end{pmatrix} \right\rangle.$$

8.9 Nilpotente Endomorphismen

Definition: Existiert ein $m \geqslant 1$ mit $f^m = 0$, so heisst f *nilpotent*.

Für jedes $j \geqslant 1$ heisst die folgende $j \times j$-Matrix ein *Jordanblock zum Eigenwert* 0:

$$N_j := \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}$$

Satz: *(Jordansche Normalform)* Ist f nilpotent, so existiert eine geordnete Basis B von V, so dass die Darstellungsmatrix $M_{B,B}(f)$ eine Blockdiagonalmatrix der Form

$$\begin{pmatrix} N_{j_1} & & \\ & \ddots & \\ & & N_{j_r} \end{pmatrix}$$

ist. Dabei ist für jedes $j \geqslant 1$ die Anzahl der $1 \leqslant i \leqslant r$ mit $j_i = j$ gleich

$$2 \dim \mathrm{Kern}(f^j) - \dim \mathrm{Kern}(f^{j-1}) - \dim \mathrm{Kern}(f^{j+1}).$$

Insbesondere sind die Diagonalblöcke bis auf Vertauschung unabhängig von B. Die Anzahl r der Jordanblöcke ist die Dimension des Eigenraums $\mathrm{Eig}_0(f)$.

Konstruktion: Für jedes $i \geqslant 0$ betrachte den Unterraum $V_i := \mathrm{Kern}(f^i)$. Durch absteigende Induktion über $i \geqslant 1$ konstruiere Unterräume W_i mit $V_i = V_{i-1} \oplus W_i$ und $f(W_{i+1}) \subset W_i$. Sodann wähle für jedes $j \geqslant 1$ einen Unterraum U_j mit $W_j = f(W_{j+1}) \oplus U_j$, sowie eine geordnete Basis $(u_{j1}, \dots, u_{jd_j})$ von U_j. Dann ist

$$B := \big(\dots; f^{j-1}(u_{jk}), f^{j-2}(u_{jk}), \dots, u_{jk}; \dots\big)$$

mit allen $j \geqslant 1$ und $1 \leqslant k \leqslant d_j$ die gesuchte Basis von V.

Beispiel: Jede $n \times n$-Matrix der Form

$$\begin{pmatrix} 0 & a_2 & * & \dots & * \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & * \\ \vdots & & & \ddots & a_n \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}$$

mit allen $a_i \neq 0$ ist ähnlich zu N_n.

Beispiel: Die folgenden reellen Matrizen haben die Jordansche Normalform

$$\begin{pmatrix} 4 & 3 & 0 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 4 & 3 & 0 & 0 \\ 0 & 2 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \left(\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right),$$

$$\begin{pmatrix} 2 & 0 & 3 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 3 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

8.10 Jordansche Normalform

Definition: Für jedes $k \geqslant 1$ heisst eine $k \times k$-Matrix der Form

$$\begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix}$$

ein *Jordanblock der Grösse k zum Eigenwert λ*.

Satz: *(Jordansche Normalform)* Ist f trigonalisierbar, so existiert eine geordnete Basis B von V, so dass die Darstellungsmatrix $M_{B,B}(f)$ eine Blockdiagonalmatrix mit Jordanblöcken auf der Blockdiagonalen ist. Dabei ist für jedes $k \geqslant 1$ die Anzahl der Jordanblöcke der Grösse k zum Eigenwert λ gleich

$$2\dim \operatorname{Kern}\big((f-\lambda\cdot \mathrm{id}_V)^k\big) - \dim \operatorname{Kern}\big((f-\lambda\cdot \mathrm{id}_V)^{k-1}\big) - \dim \operatorname{Kern}\big((f-\lambda\cdot \mathrm{id}_V)^{k+1}\big).$$

Insbesondere sind die Jordanblöcke bis auf Vertauschung unabhängig von B.

Folge: Der Endomorphismus f ist diagonalisierbar genau dann, wenn sein Minimalpolynom gleich $\prod_i (X-\lambda_i)$ ist für paarweise verschiedene $\lambda_i \in K$.

Beispiel: Die folgende reelle Matrix hat die Jordansche Normalform

$$\begin{pmatrix} 1 & -7 & 6 & 0 \\ -1 & 0 & 0 & 2 \\ 0 & 3 & 0 & 1 \\ 0 & -1 & 0 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 & 1 & 4 & 5 \\ 0 & 0 & 2 & 6 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & -7 & 6 & 0 \\ -1 & 0 & 0 & 2 \\ 0 & 3 & 0 & 1 \\ 0 & -1 & 0 & 0 \end{pmatrix} = \left(\begin{array}{c|c|cc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right).$$

Die obige Version der Jordanschen Normalform beinhaltet insbesondere den Fall, dass K algebraisch abgeschlossen ist. Für den allgemeinen Fall müssen wir die Jordanblöcke aus Begleitmatrizen anstatt aus Eigenwerten zusammensetzen:

Definition: Sei $P \in \mathrm{Mat}_{d\times d}(K)$ die Begleitmatrix eines irreduziblen Polynoms $p(X) \in K[X]$, und betrachte die $d \times d$-Elementarmatrix

$$E_{d1} = \begin{pmatrix} 0 & \dots & \dots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & & \ddots & \vdots \\ 1 & 0 & \dots & 0 \end{pmatrix}$$

Für $k \geqslant 1$ heisst eine $kd \times kd$-Matrix der Blockdiagonalgestalt

$$\begin{pmatrix} P & E_{d1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & E_{d1} \\ 0 & \dots & \dots & 0 & P \end{pmatrix}$$

ein *Jordanblock der Grösse kd zum irreduziblen Polynom* $p(X)$.

Satz: *(Jordansche Normalform)* Es existiert eine geordnete Basis B von V, so dass die Darstellungsmatrix $M_{B,B}(f)$ eine Blockdiagonalmatrix mit Jordanblöcken auf der Blockdiagonalen ist. Dabei ist für jedes $k \geqslant 1$ die Anzahl der Jordanblöcke der Grösse kd zu einem irreduziblen Polynom $p(X)$ vom Grad d gleich

$$\frac{1}{d} \cdot \Big(2\dim \mathrm{Kern}(p(f)^k) - \dim \mathrm{Kern}(p(f)^{k-1}) - \dim \mathrm{Kern}(p(f)^{k+1})\Big).$$

Insbesondere sind die Jordanblöcke bis auf Vertauschung unabhängig von B.

Folge: Zwei Matrizen sind ähnlich genau dann, wenn sie dieselbe Jordansche Normalform haben bis auf Vertauschung der Jordanblöcke.

Folge: Jede quadratische Matrix ist ähnlich zu ihrer Transponierten.

Beispiel: Die folgende reelle Matrix hat die Jordansche Normalform

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 0 & 0 & 1 & -1 \\ 1 & 4 & -3 & -3 \\ 0 & 3 & -2 & -2 \\ 1 & 2 & -1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{array}\right).$$

9 Euklidische Vektorräume

9.1 Normierte Körper

Sei K ein Körper.

Definition: Eine *Norm auf* K ist eine Abbildung

$$|\ |\colon\ K \to \mathbb{R}^{\geqslant 0},\quad x \mapsto |x|$$

mit den folgenden Eigenschaften für alle $x, y \in K$:

$$\begin{array}{lll} |x| = 0 \ \Leftrightarrow\ x = 0 & & \text{(Separiertheit)} \\ |x \cdot y| \ =\ |x| \cdot |y| & & \text{(Multiplikativität)} \\ |x + y| \ \leqslant\ |x| + |y| & & \text{(Dreiecksungleichung)} \end{array}$$

Ein Körper zusammen mit einer Norm heisst ein *normierter Körper* $(K, |\ |)$.

Beispiel: Die *Standard-Norm* auf $\mathbb{R}$ ist der Absolutbetrag

$$\mathbb{R} \to \mathbb{R}^{\geqslant 0},\ x \mapsto \begin{cases} x & \text{falls } x \geqslant 0, \\ -x & \text{falls } x < 0. \end{cases}$$

Beispiel: Die *Standard-Norm* auf $\mathbb{C}$ ist der Absolutbetrag

$$\mathbb{C} \to \mathbb{R}^{\geqslant 0},\ z \mapsto |z| := \sqrt{z\bar{z}}.$$

Wenn nichts anderes gesagt wird, meinen wir mit $|\ |$ die Standard-Norm auf $\mathbb{R}$ oder $\mathbb{C}$.

Beispiel: Die *triviale Norm* auf einem beliebigen Körper K ist die Abbildung

$$K \to \mathbb{R}^{\geqslant 0},\quad x \mapsto |x| := \begin{cases} 0 & \text{falls } x = 0, \\ 1 & \text{falls } x \neq 0. \end{cases}$$

Lemma: Für alle reelle Zahlen $\alpha, \beta \geqslant 0$ und $p \geqslant 1$ gilt $\alpha^p + \beta^p \leqslant (\alpha + \beta)^p$.

Proposition: Für jede Norm $|\ |$ auf K und jede reelle Zahl $0 < c \leqslant 1$ ist auch die Abbildung $x \mapsto |x|^c$ eine Norm auf K.

Bemerkung: Weitere interessante Normen, insbesondere auf dem Körper $\mathbb{Q}$, werden in Algebra I behandelt.

9.2 Normierte Vektorräume

Sei $(K, |\ |)$ ein normierter Körper, und sei V ein K-Vektorraum.

Definition: Eine *Norm auf V* ist eine Abbildung

$$\|\ \|\colon\ V \to \mathbb{R}^{\geqslant 0},\quad v \mapsto \|v\|$$

mit den folgenden Eigenschaften für alle $v, v' \in V$ und $x \in K$:

$$\begin{array}{ll} \|v\| = 0 \ \Leftrightarrow\ v = 0 & \text{(Separiertheit)} \\ \|x \cdot v\| \ =\ |x| \cdot \|v\| & \text{(Multiplikativität)} \\ \|v + v'\| \ \leqslant\ \|v\| + \|v'\| & \text{(Dreiecksungleichung)} \end{array}$$

Ein Vektorraum zusammen mit einer Norm heisst ein *normierter Vektorraum* $(V, \|\ \|)$.

Variante: Eine Abbildung, die alle obigen Eigenschaften ausser möglicherweise die Separiertheit erfüllt, heisst eine *Seminorm auf V*.

Proposition: Für jede Norm $\|\ \|$ und alle $v, w \in V$ gilt $\big|\, \|v\| - \|w\| \,\big| \leqslant \|v - w\|$.

Definition: Zwei Normen $\|\ \|$ und $\|\ \|'$ auf V heissen *äquivalent*, wenn reelle Zahlen $c, c' > 0$ existieren mit

$$\forall\, v \in V\colon \quad c \cdot \|v\| \ \leqslant\ \|v\|' \ \leqslant\ c' \cdot \|v\|.$$

Proposition: Dies ist eine Äquivalenzrelation auf der Menge aller Normen auf V.

Beispiel: Für jedes $p \in \mathbb{R}^{\geqslant 1} \cup \{\infty\}$ ist die ℓ_p-Norm auf K^n definiert durch

$$|(x_1, \ldots, x_n)|_p \ :=\ \begin{cases} \big(|x_1|^p + \ldots |x_n|^p\big)^{1/p} & \text{falls } p < \infty, \\ \max\big\{|x_1|, \ldots, |x_n|\big\} & \text{falls } p = \infty. \end{cases}$$

Beispiel: Seien V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und E, E' endliche Erzeugendensysteme des Dualraums V^*. Dann sind

$$\begin{array}{rl} \|v\|_{1,E} \ :=\ \sum_{\ell \in E} |\ell(v)| & \quad \text{und} \\ \|v\|_{\infty,E'} \ :=\ \max\,\{|\ell(v)| : \ell \in E'\} & \end{array}$$

Normen auf V. Ein klassischer Satz zur Geometrie der konvexen Polyeder besagt, dass sich jede Norm der ersten Form in der zweiten Form schreiben lässt und umgekehrt.

Satz: Je zwei Normen auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum sind äquivalent.

9.3 Bilinearformen

Definition: Eine *Bilinearform* auf einem K-Vektorraum V ist eine Abbildung

$$\beta\colon\ V \times V \to K,\ (v,w) \mapsto \beta(v,w)$$

so dass für alle $v, v', w, w' \in V$ und $\lambda \in K$ gilt:

$$\begin{array}{rcll} \beta(v, w+w') & = & \beta(v,w) + \beta(v,w') & \text{(rechts additiv)} \\ \beta(v+v', w) & = & \beta(v,w) + \beta(v',w) & \text{(links additiv)} \\ \beta(v, \lambda w) & = & \lambda\beta(v,w) & \text{(rechts homogen)} \\ \beta(\lambda v, w) & = & \lambda\beta(v,w) & \text{(links homogen)} \end{array}$$

Eine Bilinearform β heisst *symmetrisch*, wenn zusätzlich für alle $v, w \in V$ gilt

$$\beta(v,w) \ = \ \beta(w,v).$$

Definition: Eine quadratische Matrix A mit $A = A^T$ heisst *symmetrisch*.

Beispiel: Für jede $n \times n$-Matrix A über K ist die folgende Abbildung eine Bilinearform:

$$\beta_A\colon\ K^n \times K^n \to K,\ (x,y) \mapsto x^T A\, y.$$

Diese ist symmetrisch genau dann, wenn A symmetrisch ist.

9.4 Darstellungsmatrix

Sei $B = (v_1, \ldots, v_n)$ eine geordnete Basis von V.

Definition: Die *Darstellungsmatrix* einer Bilinearform β auf V bezüglich B ist die $n \times n$-Matrix

$$M_B(\beta) \ := \ \big(\beta(v_i, v_j)\big)_{i,j=1,\ldots,n}.$$

Proposition: Für jede $n \times n$-Matrix über K existiert genau eine Bilinearform β auf V mit $M_B(\beta) = A$.

Proposition: Eine Bilinearform auf V ist symmetrisch genau dann, wenn ihre Darstellungsmatrix bezüglich B symmetrisch ist.

Proposition: Die Darstellungsmatrix von β bezüglich jeder weiteren geordneten Basis B' von V ist

$$M_{B'}(\beta) \ = \ M_{B,B'}(\mathrm{id}_V)^T \cdot M_B(\beta) \cdot M_{B,B'}(\mathrm{id}_V)$$

9.5 Reelle Skalarprodukte

Sei V ein $\mathbb{R}$-Vektorraum.

Definition: Eine symmetrische Bilinearform

$$\langle\ ,\ \rangle\colon\ V \times V \to \mathbb{R},\ (v,w) \mapsto \langle v,w\rangle$$

heisst *positiv definit*, wenn zusätzlich gilt:

$$\forall\, v \in V \smallsetminus \{0\}\colon\ \langle v,v\rangle > 0.$$

Der *Betrag* eines Vektors $v \in V$ bezüglich $\langle\ ,\ \rangle$ ist dann die Zahl

$$\|v\|\ :=\ \sqrt{\langle v,v\rangle}\ \in\ \mathbb{R}^{\geqslant 0}.$$

Definition: Eine positiv definite symmetrische Bilinearform heisst ein *Skalarprodukt*. Ein $\mathbb{R}$-Vektorraum zusammen mit einem Skalarprodukt heisst *euklidischer Vektorraum* $(V, \langle\ ,\ \rangle)$.

Definition: Das *Standard-Skalarprodukt auf* $\mathbb{R}^n$ ist für $x = (x_i)_i$ und $y = (y_i)_i$ gegeben durch

$$\langle x,y\rangle\ :=\ x^T y\ =\ x_1y_1 + \ldots + x_ny_n.$$

Der zugehörige Betrag ist die ℓ_2-Norm

$$\|x\|\ :=\ \sqrt{x_1^2 + \ldots + x_n^2}.$$

Beispiel: Sei V ein Unterraum des Raums der stetigen Funktionen $[a,b] \to \mathbb{R}$ für reelle Zahlen $a < b$. Sei φ eine stetige Funktion $[a,b] \to \mathbb{R}^{>0}$. Dann ist

$$\langle f,g\rangle\ :=\ \int_a^b f(t)g(t)\varphi(t)\,dt$$

ein Skalarprodukt auf V.

Definition: Eine reelle symmetrische $n \times n$-Matrix A mit der Eigenschaft $x^TA\,x > 0$ für alle $0 \neq x \in \mathbb{R}^n$ heisst *positiv definit*.

Proposition: Sei B eine geordnete Basis von V. Eine symmetrische Bilinearform $\langle\ ,\ \rangle$ auf V ist positiv definit genau dann, wenn die Darstellungsmatrix $M_B(\langle\ ,\ \rangle)$ positiv definit ist.

Beispiel: Für jede positiv definite symmetrische reelle $n \times n$-Matrix A ist die folgende Abbildung ein Skalarprodukt auf $\mathbb{R}^n$:

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},\ (x,y) \mapsto x^TA\,y.$$

9.6 Grundeigenschaften

Im diesem und den folgenden Abschnitten sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum mit der zugehörigen Betragsfunktion $\|\ \|$.

Satz: *(Cauchy-Schwarz Ungleichung)* Für beliebige $v, w \in V$ gilt

$$|\langle v, w\rangle| \ \leqslant\ \|v\| \cdot \|w\|.$$

Weiter gilt Gleichheit genau dann, wenn v und w linear abhängig sind.

Proposition: Der zugehörige Betrag $\|\ \|$ ist eine Norm auf V.

Proposition: Für beliebige $v, w \in V$ gilt $\|v + w\| = \|v\| + \|w\|$ genau dann, wenn einer der Vektoren ein nicht-negatives skalares Vielfaches des anderen ist.

Definition: Ein Vektor $v \in V$ mit Betrag $\|v\| = 1$ heisst *normiert*.

Proposition: Für jeden von Null verschiedenen Vektor $v \in V$ ist $\dfrac{v}{\|v\|}$ normiert.

Definition: Der *Winkel* zwischen zwei von Null verschiedenen Vektoren $v, w \in V$ ist die eindeutige reelle Zahl $0 \leqslant \alpha \leqslant \pi$ mit

$$\cos\alpha \ =\ \frac{\langle v, w\rangle}{\|v\| \cdot \|w\|} \ =\ \Big\langle \frac{v}{\|v\|}, \frac{w}{\|w\|} \Big\rangle.$$

Definition: Zwei Vektoren $v, w \in V$ mit $\langle v, w\rangle = 0$ heissen zueinander *orthogonal*, und man schreibt $v \perp w$.

Ist ein Vektor $v \in V$ orthogonal zu jedem Element einer Teilmenge $S \subset V$, so heissen v und S *orthogonal* und man schreibt $v \perp S$ oder $S \perp v$.

Ist jedes Element einer Teilmenge $S \subset V$ orthogonal zu jedem Element einer Teilmenge $S' \subset V$, so heissen S und S' *orthogonal* und man schreibt $S \perp S'$.

Bemerkung: Für jeden Vektor $v \in V$ gilt $v \perp v \Leftrightarrow v = 0$.

9.7 Orthonormalbasen

Definition: Eine Teilmenge $S \subset V \smallsetminus \{0\}$, so dass je zwei verschiedene Elemente von S zueinander orthogonal sind, heisst *Orthogonalsystem*. Ist zusätzlich jeder Vektor in S normiert, so heisst S ein *Orthonormalsystem*. Ist zusätzlich S eine Basis von V, so heisst S je nach Fall eine *Orthogonalbasis* bzw. eine *Orthonormalbasis von* V.

Ein Tupel $(v_1, \ldots, v_n)$ mit $v_i \in V \smallsetminus \{0\}$ und $v_i \perp v_j$ für alle $i \neq j$ heisst *Orthogonalsystem*. Ist zusätzlich jedes v_i normiert, so heisst das Tupel ein *Orthonormalsystem*. Ist das Tupel zusätzlich eine geordnete Basis von V, so heisst es je nach Fall eine *geordnete Orthogonalbasis* bzw. *geordnete Orthonormalbasis von* V.

Proposition: Jedes Orthonormalsystem ist linear unabhängig.

Proposition: Sei B eine Orthonormalbasis von V, und sei $v \in V$ beliebig. Dann gilt $\langle b, v\rangle = 0$ für fast alle $b \in B$, und v hat die Entwicklung

$$v \;=\; \sum_{b\in B}{}' \langle b, v\rangle \cdot b.$$

Proposition: Eine geordnete Basis B von V ist eine Orthogonalbasis genau dann, wenn die Darstellungsmatrix des Skalarprodukts $M_B(\langle\ ,\ \rangle)$ eine Diagonalmatrix ist. Sie ist eine Orthonormalbasis genau dann, wenn die Darstellungsmatrix die Einheitsmatrix ist.

Satz: *(Gram-Schmidt-Orthogonalisierung)* Für jedes linear unabhängige Tupel $T = (v_1, \ldots, v_n)$ in V existiert genau ein Orthonormalsystem $B = (b_1, \ldots, b_n)$, so dass für alle $1 \leqslant j \leqslant n$ gilt

$$v_j = \sum_{i=1}^{j} a_{ij} b_i \quad \text{für geeignete } a_{ij} \in \mathbb{R} \text{ und } a_{jj} \in \mathbb{R}^{>0}.$$

Ist ausserdem T eine Basis von V, so ist B eine Orthonormalbasis von V.

Folge: Jeder endlich-dimensionale euklidische Vektorraum hat eine Orthonormalbasis.

Bemerkung: Die Bedingung im obigen Satz ist äquivalent dazu, dass die Basiswechselmatrix $M_{B,T}(\mathrm{id}_V)$ eine obere Dreiecksmatrix mit allen Diagonaleinträgen > 0 ist.

Satz: *(Cholesky-Zerlegung)* Für jede positiv definite reelle symmetrische Matrix A existiert eine reelle obere Dreiecksmatrix R mit allen Diagonaleinträgen > 0, so dass $A = R^T R$ ist.

9.8 Orthogonale Gruppe

Definition: Ein Isomorphismus $f\colon V \xrightarrow{\sim} W$ zwischen zwei euklidischen Vektorräumen $(V, \langle\ ,\ \rangle_V)$ und $(W, \langle\ ,\ \rangle_W)$ mit der Eigenschaft

$$\forall v, v' \in V\colon \langle f(v), f(v')\rangle_W = \langle v, v'\rangle_V$$

heisst *orthogonal* oder eine *Isometrie*.

Bedeutung: Eine Isometrie erhält alle Abstände und Winkel.

Proposition: Zwischen beliebigen euklidischen Vektorräumen derselben endlichen (!) Dimension existiert eine Isometrie. Jede Komposition von Isometrien ist eine Isometrie. Der identische Endomorphismus ist eine Isometrie.

Definition: Eine reelle $n \times n$-Matrix A, für welche die Abbildung $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$ für das jeweilige Standard-Skalarprodukt eine Isometrie ist, heisst *orthogonal*. Die Menge $\mathrm{O}(n) = \mathrm{O}_n(\mathbb{R})$ aller orthogonalen $n \times n$-Matrizen heisst die *orthogonale Gruppe vom Grad n*.

Proposition: Für jede reelle $n \times n$-Matrix Q sind äquivalent:

(a) Q ist orthogonal.

(b) Die Spalten von Q bilden eine Orthonormalbasis von $\mathbb{R}^n$ mit dem Standard-Skalarprodukt.

(c) $Q^TQ = I_n$.

(d) $QQ^T = I_n$.

Proposition: Die Menge $\mathrm{O}(n)$ ist eine Gruppe bezüglich Matrixmultiplikation.

Bemerkung: Die $\mathrm{O}(n)$ ist eine kompakte Teilmenge von $\mathrm{Mat}_{n\times n}(\mathbb{R}) \cong \mathbb{R}^{n^2}$.

Satz: *(Variante der Gram-Schmidt-Orthogonalisierung)* Sei V ein euklidischer Vektorraum der Dimension $\geqslant n$. Für alle $v_1, \ldots, v_n \in V$ existiert ein Orthonormalsystem $(b_1, \ldots, b_n)$ in V, so dass für alle $1 \leqslant j \leqslant n$ gilt

$$v_j = \sum_{i=1}^{j} a_{ij} b_i \quad \text{für geeignete } a_{ij} \in \mathbb{R}.$$

Satz: *(QR-Zerlegung)* Für jede reelle $n \times n$-Matrix A existiert eine orthogonale Matrix Q und eine reelle obere Dreiecksmatrix R, so dass $A = QR$ ist.

Beispiel: Für jedes $x \in \mathbb{R}$ gilt:

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{1+x^2}} & \frac{-x}{\sqrt{1+x^2}} \\ \frac{x}{\sqrt{1+x^2}} & \frac{1}{\sqrt{1+x^2}} \end{pmatrix} \cdot \begin{pmatrix} \sqrt{1+x^2} & \frac{x}{\sqrt{1+x^2}} \\ 0 & \frac{1}{\sqrt{1+x^2}} \end{pmatrix}$$

Für $x \to \infty$ bleibt Q in dem Kompaktum $\mathrm{O}(n)$ und das Wachstum findet nur in dem Faktor R der Zerlegung QR statt.

9.9 Volumen

„Satz“: Für jede reelle $n \times n$-Matrix A und jede Teilmenge $X \subset \mathbb{R}^n$ gilt

$$\mathrm{vol}_n(L_A(X)) \;=\; |\det(A)| \cdot \mathrm{vol}_n(X),$$

sofern beide Seiten wohldefiniert sind.

Bemerkung: Die Anführungszeichen in diesem Satz und den folgenden beziehen sich darauf, dass das Volumen erst in der mehrdimensionalen Analysis richtig definiert wird.

„Folge“: Das Volumen des von beliebigen $v_1, \ldots, v_n \in \mathbb{R}^n$ aufgespannten *Parallelotops* (oder *Parallelepipeds* oder *Raumspats*) ist

$$\mathrm{vol}_n\big(\big\{\textstyle\sum_{i=1}^n t_i v_i \;\big|\; \forall\, i\colon 0 \leqslant t_i \leqslant 1\big\}\big) \;=\; \big|\det(v_1, \ldots, v_n)\big|.$$

„Proposition-Definition“: Sei $(V, \langle\ ,\ \rangle)$ ein euklidischer Vektorraum der Dimension n, und sei $f\colon \mathbb{R}^n \xrightarrow{\sim} V$ eine Isometrie (für das Standard-Skalarprodukt auf $\mathbb{R}^n$). Dann ist für jede Teilmenge $X \subset V$ das *Volumen*

$$\mathrm{vol}_n(X) \;:=\; \mathrm{vol}(f^{-1}(X))$$

unabhängig von f, sofern es existiert.

Betrachte nun Vektoren $v_1, \ldots, v_m$ in einem beliebigen euklidischen Vektorraum $(V, \langle\ ,\ \rangle)$.

Definition: Die *Gram-sche Determinante* von $v_1, \ldots, v_m$ ist

$$G \;:=\; \det(\langle v_i, v_j\rangle)_{i,j=1,\ldots,m} \;\in\; \mathbb{R}.$$

„Satz“: (a) Es gilt stets $G \geqslant 0$.

(b) Es ist $G > 0$ genau dann, wenn $v_1, \ldots, v_m$ linear unabhängig sind.

(c) Es gilt

$$\mathrm{vol}_m\big(\big\{\textstyle\sum_{i=1}^m t_i v_i \;\big|\; \forall\, i\colon 0 \leqslant t_i \leqslant 1\big\}\big) \;=\; \sqrt{G}.$$

Beispiel: Für $m = 1$ ist das m-dimensionale Volumen die *Länge*

$$\mathrm{vol}_1\big(\big\{t v_1 \;\big|\; 0 \leqslant t \leqslant 1\big\}\big) \;=\; \sqrt{\langle v_1, v_1\rangle} \;=\; |v_1|.$$

Für $m = 2$ ist es der *Flächeninhalt*, zum Beispiel gilt für alle $a, b \in \mathbb{R}$ mit $v_1 := (1, 0, a)^T$ und $v_2 := (0, 1, b)^T$

$$\mathrm{vol}_2\Big(\Big\{\begin{pmatrix} s \\ t \\ as + bt \end{pmatrix} \;\Big|\; s, t \in [0, 1]\Big\}\Big) \;=\; \sqrt{1 + a^2 + b^2}.$$

9.10 Unterräume, orthogonales Komplement

Ab diesem Abschnitt sei $(V, \langle\ ,\ \rangle)$ wieder ein beliebiger euklidischer Vektorraum.

Satz: Ist $\dim V < \infty$, so lässt sich jede Orthonormalbasis jedes Unterraums U zu einer Orthonormalbasis von V erweitern.

Definition: Das *orthogonale Komplement* einer Teilmenge $S \subset V$ ist die Teilmenge

$$S^\perp \ := \ \{v \in V \mid v \perp S\,\}.$$

Proposition: Das orthogonale Komplement ist ein Unterraum.

Beispiel: Es gilt $\{0\}^\perp = \varnothing^\perp = V$ und $V^\perp = \{0\}$.

Proposition: (a) Für jede Teilmenge $S \subset V$ gilt $S \subset (S^\perp)^\perp$.

(b) Für je zwei Teilmengen $S \subset T \subset V$ gilt $S^\perp \supset T^\perp$.

(c) Für jeden Unterraum $U \subset V$ gilt $U \cap U^\perp = \{0\}$.

Proposition: Für jeden endlich-dimensionalen Unterraum $U \subset V$ gilt $V = U \oplus U^\perp$.

Sei nun $U \subset V$ ein Unterraum mit $V = U \oplus U^\perp$.

Proposition: (a) Jede Orthonormalbasis von U zusammen mit jeder Orthonormalbasis von $U^\perp$ bildet eine Orthonormalbasis von V.

(b) $\dim V = \dim U + \dim U^\perp$.

(c) $U = (U^\perp)^\perp$.

(d) $U = V$ genau dann, wenn $U^\perp = 0$ ist.

Definition: Die *orthogonale Projektion von V auf U* ist die lineare Abbildung

$$\begin{aligned} \pi_U \colon V = U \oplus U^\perp &\to U, \\ u + u' &\mapsto u \qquad \text{für alle } u \in U \text{ und } u' \in U^\perp. \end{aligned}$$

Ist $(b_1, \ldots, b_n)$ eine Orthonormalbasis von U, so ist π_U gegeben durch die Formel

$$\pi_U(v) \ = \ \sum_{i=1}^{n} \langle b_i, v\rangle \cdot b_i.$$

Proposition: Für jeden Vektor $v \in V$ ist $\pi_U(v)$ der eindeutige Vektor $u \in U$, für den $\|v - u\|$ minimal ist.

Beispiel: Sei V der Vektorraum aller stetigen Funktionen $[a,b] \to \mathbb{R}$ mit dem Skalarprodukt $\langle f,g\rangle := \int_a^b f(x)g(x)\,dx$, und sei $U \subset V$ der Unterraum aller Polynomfunktionen vom Grad $\leqslant n$. Für jede Funktion $f \in V$ ist dann $\pi_U(f)$ die eindeutige Funktion in U, welche f bestmöglich *im quadratischen Mittel approximiert*.

Vorsicht: Wo oben endliche Dimension vorausgesetzt ist, ist dies notwendig.

Beispiel: Sei V der Vektorraum aller beschränkten reellen Folgen $\underline{x} = (x_n)_{n\geqslant 1}$ mit dem Skalarprodukt

$$\langle \underline{x}, \underline{y}\rangle \;:=\; \sum_{n\geqslant 1} \frac{x_n y_n}{n^2}.$$

Sei U der Unterraum aller Folgen mit der Eigenschaft $\forall x \gg 0\colon x_n = 0$. Dann gilt $U^\perp = 0$. Die Folgen $(m\cdot\delta_{nm})_{n\geqslant 1}$ für alle $m \geqslant 1$ bilden eine Orthonormalbasis von U, welche keine Fortsetzung zu einer Orthonormalbasis von V hat.

9.11 Skalarprodukte und Dualraum

Proposition: (a) Die folgende Abbildung ist ein injektiver Homomorphismus:

$$\delta\colon V \longrightarrow V^* := \mathrm{Hom}_{\mathbb{R}}(V,\mathbb{R}), \quad v \mapsto \delta(v) := \langle v,\ \rangle$$

(b) Diese ist ein Isomorphismus genau dann, wenn $\dim V < \infty$ ist.

Vorsicht: Einen endlich-dimensionalen euklidischen Vektorraum kann man deshalb leicht mit seinem Dualraum verwechseln. Man soll diese aber nicht ohne Not identifizieren. Wenn man sie auseinander hält, kann man Fehler vermeiden, zum Beispiel beim Verhalten unter linearen Abbildungen. Eine lineare Abbildung $f\colon V \to V$ entspricht unter dem obigen Isomorphismus im allgemeinen nicht ihrer dualen Abbildung $f^*\colon V^* \to V^*$, denn

$$\begin{aligned}
& & f &= \delta^{-1}\circ f^*\circ\delta \\
\Leftrightarrow & & \delta\circ f &= f^*\circ\delta \\
\Leftrightarrow & \forall v\in V: & \delta(f(v)) &= f^*(\delta(v)) \overset{\text{def}}{=} \delta(v)\circ f \\
\Leftrightarrow & \forall v,w\in V: & \langle f(v),w\rangle &\overset{\text{def}}{=} \delta(f(v))(w) = \delta(v)(f(w)) \overset{\text{def}}{=} \langle v,f(w)\rangle,
\end{aligned}$$

und die letzte Eigenschaft ist nicht immer erfüllt.

Bemerkung: Sei V ein K-Vektorraum für einen beliebigen Körper K, und sei U ein Unterraum. Dann ist

$$\big\{\ell \in V^* \;\big|\; \forall u\in U\colon \ell(u) = 0\big\}$$

ein Unterraum von V^*, welcher auch oft das *orthogonale Komplement von U* genannt wird. Die Rechtfertigung dafür besteht darin, dass dieser für jeden endlich-dimensionalen euklidischen Vektorraum gleich $\delta(U^\perp)$ ist mit dem obigen Isomorphismus δ.

9.12 Adjungierte Abbildungen

Seien V und W euklidische Vektorräume und $f\colon V \to W$ eine lineare Abbildung.

Proposition: Es gibt höchstens eine lineare Abbildung $f^*\colon W \to V$ mit der Eigenschaft

$$\forall v \in V\ \forall w \in W\colon\ \langle f(v), w\rangle = \langle v, f^*(w)\rangle.$$

Definition: Diese heisst die *Adjungierte (Abbildung) von* f, wenn sie existiert.

Proposition: Ist f^* die Adjungierte von f, so ist auch f die Adjungierte von f^*, das heisst, es gilt $(f^*)^* = f$. Man nennt f und f^* daher auch *zueinander adjungiert*.

Beispiel: Die Adjungierte der Nullabbildung $V \to W$ ist die Nullabbildung $W \to V$.

Beispiel: Sei $U \subset V$ ein Unterraum mit $V = U \oplus U^\perp$ und mit dem von V induzierten Skalarprodukt. Dann ist die zur Inklusion $i_U\colon U \hookrightarrow V$, $u \mapsto u$ adjungierte Abbildung die orthogonale Projektion $\pi_U\colon V \twoheadrightarrow U$, und umgekehrt.

Vorsicht: Die Adjungierte existiert nicht immer, zum Beispiel nicht für die Inklusion $U \hookrightarrow V$ in dem letzten Beispiel von §9.10.

Proposition: Ist $\dim V < \infty$, so existiert die adjungierte Abbildung f^*. Genauer gilt für jede geordnete Orthonormalbasis $(b_1, \ldots, b_n)$ von V

$$\forall w \in W\colon\ f^*(w) \ = \ \sum_{i=1}^{n} \langle f(b_i), w\rangle \cdot b_i.$$

Proposition: Seien B eine geordnete Orthonormalbasis von V und B' eine geordnete Orthonormalbasis von W. Dann gilt

$$M_{B,B'}(f^*) \ = \ M_{B',B}(f)^T.$$

Proposition: Für jede reelle $m \times n$-Matrix A ist die Adjungierte von $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ gleich $L_{A^T}\colon \mathbb{R}^m \to \mathbb{R}^n$.

Beispiel: Jede orthogonale lineare Abbildung $f\colon V \xrightarrow{\sim} W$ hat Adjungierte $f^* = f^{-1}$.

Definition: Eine lineare Abbildung $f\colon V \to V$, die ihre eigene Adjungierte ist, heisst *selbstadjungiert*.

Beispiel: Die Abbildung $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$ ist selbstadjungiert genau dann, wenn A symmetrisch ist.

9.13 Spektralsatz für selbstadjungierte Endomorphismen

Sei f ein selbstadjungierter Endomorphismus eines endlich-dimensionalen euklidischen Vektorraums V.

Lemma 1: Für jeden Eigenvektor $v \in V$ von f ist die Zerlegung $V = \mathbb{R}v \oplus (\mathbb{R}v)^\perp$ invariant unter f, das heisst, es gilt $f(\mathbb{R}v) \subset \mathbb{R}v$ und $f((\mathbb{R}v)^\perp) \subset (\mathbb{R}v)^\perp$.

Lemma 2: Ist $\dim V > 0$, so besitzt f einen Eigenwert in $\mathbb{R}$.

Spektralsatz: Für jeden selbstadjungierten Endomorphismus eines endlich-dimensionalen euklidischen Vektorraums V existiert eine Orthonormalbasis von V bestehend aus Eigenvektoren von f.

Folge: Jeder selbstadjungierte Endomorphismus eines endlich-dimensionalen euklidischen Vektorraums V ist diagonalisierbar.

Geometrische Bedeutung: Die von der Orthonormalbasis erzeugten Geraden heissen *Hauptachsen von f*. Diese bilden ein kartesisches Koordinatensystem, in welchem f durch je eine Streckung um voneinander unabhängige reelle Faktoren in alle Koordinatenrichtungen beschrieben wird.

Hauptachsentransformation 1: Für jede reelle symmetrische Matrix A existiert eine orthogonale Matrix Q, so dass $Q^{-1}AQ = Q^TAQ$ Diagonalgestalt hat.

Folge: Alle komplexen Eigenwerte einer reellen symmetrischen Matrix sind reell.

Methode: Eine Orthonormalbasis aus Eigenvektoren findet man, indem zuerst anhand des charakteristischen Polynoms alle Eigenwerte bestimmt, dann zu jedem Eigenwert eine Basis des zugehörigen Eigenraums berechnet, auf die Basis jedes Eigenraums das Gram-Schmidt-Orthogonalisierungsverfahren anwendet, und zuletzt die so erhaltenen Basisvektoren zu einer Basis des Gesamtraums zusammenfügt.

Beispiel: Die reelle Matrix

$$A := \frac{1}{15}\begin{pmatrix} 10 & 5 & 10 \\ 5 & -14 & 2 \\ 10 & 2 & -11 \end{pmatrix}$$

ist gleichzeitig symmetrisch und orthogonal. Ihr charakteristisches Polynom lautet $(X-1)(X+1)^2$; nach dem obigen Satz existiert also eine orthogonale Matrix Q mit

$$Q^{-1}AQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Die lineare Abbildung L_A setzt sich somit aus einer Punktspiegelung in einer Ebene und der Identität in der dazu orthogonalen Geraden zusammen. Konkret sind die Spalten von Q eine Orthonormalbasis aus Eigenvektoren, zum Beispiel findet man

$$Q = \begin{pmatrix} 5/\sqrt{30} & 0 & 1/\sqrt{6} \\ 1/\sqrt{30} & -2/\sqrt{5} & -1/\sqrt{6} \\ 2/\sqrt{30} & 1/\sqrt{5} & -2/\sqrt{6} \end{pmatrix}.$$

9.14 Normalform symmetrischer Bilinearformen

Eine reelle $n \times n$-Matrix A kann genauso gut einen Endomorphismus von $\mathbb{R}^n$ darstellen wie eine Bilinearform auf $\mathbb{R}^n$, nämlich

$$\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},\ (x, y) \mapsto x^T A y.$$

Eine äquivalente Formulierung des Hauptachsentransformationssatzes ist daher:

Hauptachsentransformation 2: Für jede reelle symmetrische Matrix A existiert eine Orthonormalbasis von $\mathbb{R}^n$, bezüglich welcher die obige symmetrische Bilinearform Diagonalgestalt erhält.

Definition: Eine symmetrische Bilinearform β auf einem reellen Vektorraum V heisst

(a) *nicht-ausgeartet* wenn $\forall v \in V \smallsetminus \{0\}\colon\ \exists\, w \in V\colon\ \beta(v, w) \neq 0$;

(b) *ausgeartet*, wenn $\exists\, v \in V \smallsetminus \{0\}\colon\ \forall w \in V\colon\ \beta(v, w) = 0$;

(c) *positiv definit*, wenn $\forall v \in V \smallsetminus \{0\}\colon\ \beta(v, v) > 0$.

(d) *positiv semi-definit*, wenn $\forall v \in V\colon\ \beta(v, v) \geqslant 0$.

(e) *negativ definit*, wenn $\forall v \in V \smallsetminus \{0\}\colon\ \beta(v, v) < 0$.

(f) *negativ semi-definit*, wenn $\forall v \in V\colon\ \beta(v, v) \leqslant 0$.

(g) *indefinit*, wenn $\exists\, v, w \in V\colon\ \beta(v, v) > 0 > \beta(w, w)$.

Eine reelle symmetrische $n \times n$-Matrix A heisst entsprechend, wenn die zugehörige Bilinearform die jeweilige Eigenschaft hat.

Bemerkung: Es ist β negativ definit genau dann, wenn $-\beta$ positiv definit ist. Analog für semidefinit, usw., sowie für A.

Sei β eine symmetrische Bilinearform auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum V.

Satz: (a) Es existiert eine geordnete Basis von V, bezüglich welcher die Darstellungsmatrix von β eine Diagonalmatrix mit Diagonaleinträgen aus $\{+1, -1, 0\}$ ist.

(b) Die Anzahl d_0 der Diagonaleinträge 0 ist die Dimension des Kerns der linearen Abbildung $V \to V^*$, $v \mapsto \beta(v,\ )$.

(c) Die Anzahl $d_\pm$ der Diagonaleinträge ± 1 ist die maximale Dimension eines Teilraums $U \subset V$, für den die Einschränkung $\pm\beta|_{U \times U}$ positiv definit ist.

(d) Insbesondere sind die Diagonaleinträge bis auf Vertauschung unabhängig von der in (a) gewählten Basis.

Definition: Das Tupel (d_+, d_-) oder (d_0, d_+, d_-) heisst die *Signatur von* β.
Die Zahl $d_+ + d_-$ heisst der *Rang von* β.
Die Zahl $d_+ - d_-$ heisst der *Index von* β.

Zusatz: Für die einer reellen symmetrischen $n \times n$-Matrix A zugeordnete Bilinearform $(x, y) \mapsto x^T Ay$ ist

d_+ die Anzahl der positiven Eigenwerte von A,
d_- die Anzahl der negativen Eigenwerte von A,
$d_+ + d_-$ der Rang von A.

Beispiel: Die Matrix $A := \left(\begin{smallmatrix} 1 & 3 \\ 3 & 1 \end{smallmatrix}\right)$ hat die Eigenwerte 4 und -2; der zugehörige Endomorphismus von $\mathbb{R}^2$ ist daher diagonalisierbar zu $\left(\begin{smallmatrix} 4 & 0 \\ 0 & -2 \end{smallmatrix}\right)$. Die zugehörige Bilinearform hat dagegen die Normalform $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Diese erreicht man durch Bestimmung der Eigenräume oder durch geeigneten Basiswechsel mit einer Dreiecksmatrix, z.B. mit

$$\begin{pmatrix} 1 & 3/\sqrt{8} \\ 0 & -1/\sqrt{8} \end{pmatrix}^T \cdot \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3/\sqrt{8} \\ 0 & -1/\sqrt{8} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Beispiel: Das charakteristische Polynom der reellen symmetrischen Matrix

$$\begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix}$$

ist $(X - (a-1))^2(X - (a+2))$, woraus man schnell die Werte d_+, d_-, d_0 bestimmt.

9.15 Kriterien für Positiv-Definitheit

Satz: Für jede reelle symmetrische $n \times n$-Matrix $A = \left(a_{ij}\right)_{i,j=1,\ldots,n}$ sind äquivalent:

(a) Die Matrix A ist positiv definit.
(b) Alle Eigenwerte von A sind positiv.
(c) Es existiert eine invertierbare Matrix B mit $A = B^T B$.
(d) Es existiert eine invertierbare obere Dreiecksmatrix R mit $A = R^T R$.
(e) Es existiert eine invertierbare symmetrische Matrix C mit $A = C^T C = C^2$.
(f) Die Determinante der Matrix $A_k := \left(a_{ij}\right)_{i,j=1,\ldots,k}$ ist positiv für jedes $1 \leqslant k \leqslant n$. (**Hauptminorenkriterium**)

Beispiel: Die Hauptminoren der reellen symmetrischen Matrix

$$\begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix}$$

sind a und $a^2 - 1 = (a-1)(a+1)$ und $a^3 - 3a + 2 = (a-1)^2(a+2)$; daher ist die Matrix positiv definit genau dann, wenn $a > 1$ ist. Im Fall $a = 2$ gilt zum Beispiel:

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ \sqrt{\frac{1}{2}} & \sqrt{\frac{7}{4}} & 0 \\ \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{7}} & \sqrt{\frac{19}{14}} \end{pmatrix} \cdot \begin{pmatrix} \sqrt{2} & \sqrt{\frac{1}{2}} & \sqrt{\frac{1}{2}} \\ 0 & \sqrt{\frac{7}{4}} & \sqrt{\frac{1}{7}} \\ 0 & 0 & \sqrt{\frac{19}{14}} \end{pmatrix} = \left[\frac{1}{3}\begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}\right]^2$$

9.16 Singulärwertzerlegung

In den Abschnitten 3.7, 8.3, 9.7, 9.8, 9.13, 9.15 haben wir schon verschiedene Matrixzerlegungen kennengelernt. Eine weitere ist:

Satz: Für jede reelle $m \times n$-Matrix A vom Rang r existieren eine orthogonale $m \times m$-Matrix Q, eine orthogonale $n \times n$-Matrix R, und eine $m \times n$-Matrix der Form

$$D \;=\; \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ \hline & & & \end{array}\right)$$

für reelle Zahlen $\sigma_1 \geqslant \ldots \geqslant \sigma_r > 0$ und allen übrigen Einträgen 0, so dass gilt

$$A \;=\; QDR.$$

Dabei sind die Zahlen $\sigma_1, \ldots, \sigma_r$ durch A eindeutig bestimmt. Genauer sind $\sigma_1^2, \ldots, \sigma_r^2$ genau die von Null verschiedenen Eigenwerte von $A^T A$, mit Vielfachheiten.

Definition: Die Zahlen $\sigma_1, \ldots, \sigma_r$ heissen die *Singulärwerte* von A.

Tipp: Vergleiche mit der LR-Zerlegung aus §3.7 und der QR-Zerlegung aus §9.8.

Beispiel: Die Matrix $\left(\begin{smallmatrix} 2 & 3 \\ 0 & 2 \end{smallmatrix}\right)$ hat die Singulärwertzerlegung

$$\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix} \;=\; \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} \cdot \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}.$$

Beispiel: Für jedes $x \in \mathbb{R}$ sind die Singulärwerte der Matrix $\left(\begin{smallmatrix} 1 & x \\ 0 & 1 \end{smallmatrix}\right)$ gleich

$$\sqrt{\tfrac{1}{2} \cdot \big(2 + x^2 \pm \sqrt{x^2(4 + x^2)}\big)}.$$

Für $|x| \to \infty$ verhalten sie sich asymptotisch wie $|x|$ und $|x|^{-1}$, während die Faktoren Q und R der Singulärwertzerlegung in dem Kompaktum $\mathrm{O}(n)$ bleiben.

9.17 Quadratische Formen

Zuerst seien K ein beliebiger Körper und n eine positive natürliche Zahl.

Definition: Ein Polynom der Form

$$q(x_1, \ldots, x_n) \;=\; a + \sum_{1 \leqslant i \leqslant n} a_i x_i + \sum_{1 \leqslant i \leqslant j \leqslant n} a_{ij} x_i x_j$$

mit a, a_i, $a_{ij} \in K$ heisst eine *quadratische Form in n Variablen über K*. Sind dabei $a = a_1 = \ldots = a_n = 0$, so heisst q *homogen*, sonst *inhomogen*. Übersetzt in Spaltenvektoren $x \in K^n$ ist eine quadratische Form also ein Ausdruck der Form

$$q(x) \;=\; \alpha + a^T x + x^T A x$$

für beliebige $\alpha \in K$ und $a \in K^n$ und $A \in \mathrm{Mat}_{n\times n}(K)$.

Proposition: Ist $1 + 1 \neq 0$ in K, so besitzt jede quadratische Form über K eine eindeutige Darstellung $\alpha + a^T x + x^T A x$ mit A symmetrisch.

Betrachte nun eine reelle quadratische Form $q(x) = \alpha + a^T x + x^T A x$ mit A symmetrisch sowie die zugehörige symmetrische Bilinearform

$$\beta \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},\ (x, y) \mapsto x^T A y.$$

Definition: Rang, Index und Signatur von β heissen auch *Rang, Index* und *Signatur von q*. Ist β nicht-ausgeartet (also A invertierbar), so heisst q *nicht-ausgeartet*.

Hauptachsentransformation 3: Für jede reelle quadratische Form q existieren eine orthogonale Matrix Q und $\beta \in \mathbb{R}$ und $b \in \mathbb{R}^n$ sowie eine Diagonalmatrix D, so dass gilt

$$q(Qx) \;=\; \beta + b^T x + x^T D x.$$

Ist zusätzlich q nicht-ausgeartet, so existieren weiter $\gamma \in \mathbb{R}$ und $c \in \mathbb{R}^n$, so dass gilt

$$q(Qx + c) \;=\; \gamma + x^T D x.$$

Definition: Die Nullstellenmenge $\{x \in \mathbb{R}^n \mid q(x) = 0\}$ einer reellen quadratischen Form q wie oben mit $A \neq 0$ heisst eine *reelle Quadrik*.

Bedeutung: Der Satz liefert eine Normalform einer Quadrik bis auf Isometrie und Translation. Ist q nicht-ausgeartet, so erhalten wir ein neues kartesisches Koordinatensystem im $\mathbb{R}^n$ mit Ursprung im Punkt c und den Koordinatenachsen $\{Qte_i + c \mid t \in \mathbb{R}\}$ mit $e_i = (\delta_{ij})_j$. Diese Koordinatenachsen heissen *Hauptachsen von q*. Diese sind gleichzeitig *Symmetrieachsen* der nicht-ausgearteten Quadrik, da die Quadrik invariant unter der Spiegelung an der jeweils dazu orthogonalen Hyperebene ist.

Beispiele für Quadriken:

im $\mathbb{R}^2$: Ellipse, Hyperbel, Parabel, zwei Geraden, doppelte Gerade.

im $\mathbb{R}^3$: Ellipsoid, einschaliges oder zweischaliges Hyperbeloid, Doppelkegel, Paraboloid.

im $\mathbb{R}^4$: Lichtkegel im Minkowski-Raum $x_1^2 + x_2^2 + x_3^2 - x_4^2 = 0$.

9.18 Spektralsatz für normale Endomorphismen

Wie vorher sei V ein euklidischer Vektorraum.

Definition: Eine lineare Abbildung $f\colon V \to V$, deren Adjungierte f^* existiert und die Gleichung $f^* \circ f = f \circ f^*$ erfüllt, heisst *normal*.

Proposition: (a) Jede selbstadjungierte lineare Abbildung ist normal.

(b) Jede orthogonale lineare Abbildung $f\colon V \xrightarrow{\sim} V$ ist normal.

Beispiel: Insbesondere ist nach (a) jede reelle Diagonalmatrix normal.

Beispiel: Für alle $a, b \in \mathbb{R}$ kommutiert die Matrix $A := \left(\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right)$ mit $A^T = \left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$; also ist die Abbildung $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$ normal.

Bemerkung: Die Abbildung

$$\mathbb{C} \longrightarrow \mathrm{Mat}_{2\times 2}(\mathbb{R}),\ a + ib \mapsto \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

ist ein injektiver Ringhomomorphismus, also ein Ringisomorphismus von $\mathbb{C}$ auf sein Bild. Man kann dies zur Konstruktion von $\mathbb{C}$ verwenden anstelle der üblichen Identifikation $\mathbb{C} \cong \mathbb{R}^2$.

Spektralsatz: Für jeden normalen Endomorphismus f eines endlich-dimensionalen euklidischen Vektorraums V existiert eine geordnete Orthonormalbasis B von V, bezüglich welcher die Darstellungsmatrix von f die folgende Blockdiagonalgestalt hat:

$$M_{B,B}(f) = \begin{pmatrix} D_1 & & \\ & \ddots & \\ & & D_r \end{pmatrix} \quad \text{mit} \quad D_k = \begin{cases} a_k & \text{mit } a_k \in \mathbb{R} \text{ oder} \\ \left(\begin{smallmatrix} a_k & b_k \\ -b_k & a_k \end{smallmatrix}\right) & \text{mit } a_k, b_k \in \mathbb{R}. \end{cases}$$

9.19 Klassifikation orthogonaler Endomorphismen

Sei V ein endlich-dimensionaler euklidischer Vektorraum.

Proposition: Für jeden orthogonalen Endomorphismus f von V gilt $\det(f) = \pm 1$.

Satz: Für jeden orthogonalen Endomorphismus f von V existiert eine geordnete Orthonormalbasis B von V, bezüglich welcher die Darstellungsmatrix von f die folgende Blockdiagonalgestalt hat:

$$M_{B,B}(f) = \begin{pmatrix} D_1 & & \\ & \ddots & \\ & & D_r \end{pmatrix} \quad \text{mit} \quad D_k = \begin{cases} \pm 1 & \text{oder} \\ \left(\begin{smallmatrix} a_k & b_k \\ -b_k & a_k \end{smallmatrix}\right) & \text{mit } a_k,\, b_k \in \mathbb{R} \\ & \text{und } a_k^2 + b_k^2 = 1. \end{cases}$$

Gilt weiter $\det(f) = 1$, so kann man alle 1×1-Blockdiagonaleinträge gleich 1 wählen.

Definition: Sind alle Blockdiagonaleinträge ausser einem gleich 1, und ist dieser

(a) gleich -1, so heisst f eine *Spiegelung (an einer Hyperebene)*.
(b) gleich $\left(\begin{smallmatrix} a_k & b_k \\ -b_k & a_k \end{smallmatrix}\right)$, so heisst f eine *Drehung um den Winkel* $\pm \arg(a_k + ib_k)$.

Proposition: Jeder orthogonale Endomorphismus von V ist eine Komposition von Spiegelungen. Insbesondere ist die orthogonale Gruppe $O(n)$ von Spiegelungen erzeugt.

Definition: Ein orthogonaler Endomorphismus f mit $\det(f) = 1$ heisst *speziell orthogonal*. Die Menge $\mathrm{SO}(n) = \mathrm{SO}_n(\mathbb{R})$ aller orthogonalen $n \times n$-Matrizen mit Determinante 1 heisst die *spezielle orthogonale Gruppe vom Grad n*.

Proposition: Jeder spezielle orthogonale Endomorphismus ist eine Komposition von Drehungen. Insbesondere ist die spezielle orthogonale Gruppe $\mathrm{SO}(n)$ von Drehungen erzeugt.

Proposition: *(„Satz vom Fussball“)* Jedes Element von SO(3) ist eine Drehung.

Bemerkung: Dass eine Katze sich in der Luft auf die Füsse drehen kann, hängt unter anderem damit zusammen, dass die Gruppe SO(3) nicht kommutativ ist.

Beispiel: Siehe §9.13.

10 Unitäre Vektorräume

Die Theorie komplexer Vektorräume mit Skalarprodukt folgt denselben Linien wie die Theorie reeller Vektorräume mit Skalarprodukt; die meisten Definitionen und Resultate sind sogar schon wörtlich gleich oder gehen ineinander über durch systematisches Ersetzen der folgenden Begriffe:

$\mathbb{R}$	$\mathbb{C}$
euklidischer Vektorraum	unitärer Vektorraum
Bilinearform	Sesquilinearform
symmetrische Bilinearform	Hermitesche Form
transponierte Matrix A^T	adjungierte Matrix $A^* = \overline{A}^T$
orthogonale Abbildung	unitäre Abbildung

10.1 Hermitesche Formen

Sei V ein $\mathbb{C}$-Vektorraum.

Definition: Eine *Sesquilinearform* auf V ist eine Abbildung

$$\gamma\colon\ V \times V \to \mathbb{C},\ (v,w) \mapsto \gamma(v,w)$$

so dass für alle $v, v', w, w' \in V$ und $\lambda \in \mathbb{C}$ gilt:

$$\begin{aligned}
\gamma(v, w+w') &= \gamma(v,w) + \gamma(v,w') && \text{(rechts additiv)}\\
\gamma(v+v', w) &= \gamma(v,w) + \gamma(v',w) && \text{(links additiv)}\\
\gamma(v, \lambda w) &= \lambda \cdot \gamma(v,w) && \text{(rechts homogen)}\\
\gamma(\lambda v, w) &= \overline{\lambda} \cdot \gamma(v,w) && \text{(links halbhomogen)}
\end{aligned}$$

Eine *hermitesche Form* auf V ist eine Sesquilinearform γ, so dass für alle $v, w \in V$ gilt

$$\gamma(v,w) \ = \ \overline{\gamma(w,v)}.$$

Vorsicht: Welche Variable linear und welche semilinear ist, wird nicht einheitlich gehandhabt. Die hier gewählte Konvention macht gewisse Formeln schöner.

Definition: Die *komplex Konjugierte* einer komplexen Matrix $A = (a_{ij})_{i,j}$ ist die Matrix $\overline{A} := (\overline{a_{ij}})_{i,j}$, und die *Adjungierte* ist $A^* := \overline{A}^T$. Analog für Spalten- und Zeilenvektoren.

Definition: Eine komplexe Matrix A mit $A = A^*$ heisst *selbstadjungiert* oder *hermitesch*.

Beispiel: Für jede komplexe $n \times n$-Matrix A ist die folgende Abbildung eine Sesquilinearform:

$$\gamma_A\colon\ \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C},\ (x,y) \mapsto x^* A\, y.$$

Diese ist hermitesch genau dann, wenn A hermitesch ist.

10.2 Darstellungsmatrix

Sei $B = (v_1, \ldots, v_n)$ eine geordnete Basis von V.

Definition: Die *Darstellungsmatrix* einer Sesquilinearform γ auf V bezüglich B ist die $n \times n$-Matrix

$$M_B(\gamma) \;:=\; \big(\gamma(v_i, v_j)\big)_{i,j=1,\ldots,n}.$$

Proposition: Für jede komplexe $n \times n$-Matrix existiert genau eine Sesquilinearform γ auf V mit $M_B(\gamma) = A$.

Proposition: Eine Sesquilinearform auf V ist hermitesch genau dann, wenn ihre Darstellungsmatrix bezüglich B hermitesch ist.

Proposition: Die Darstellungsmatrix von γ bezüglich jeder weiteren geordneten Basis B' von V ist

$$M_{B'}(\gamma) \;=\; M_{B,B'}(\mathrm{id}_V)^* \cdot M_B(\gamma) \cdot M_{B,B'}(\mathrm{id}_V).$$

10.3 Komplexe Skalarprodukte

Proposition: Für jede hermitesche Form γ auf V und jeden Vektor $v \in V$ gilt

$$\gamma(v, v) \in \mathbb{R}.$$

Definition: Eine hermitesche Form

$$\langle\ ,\ \rangle\colon\ V \times V \to \mathbb{C},\ (v, w) \mapsto \langle v, w\rangle$$

heisst *positiv definit*, wenn zusätzlich gilt:

$$\forall\, v \in V \smallsetminus \{0\}\colon\ \langle v, v\rangle > 0.$$

Der *Betrag* eines Vektors $v \in V$ bezüglich $\langle\ ,\ \rangle$ ist dann die Zahl

$$\|v\| \;:=\; \sqrt{\langle v, v\rangle} \;\in\; \mathbb{R}^{\geqslant 0}.$$

Definition: Eine positiv definite hermitesche Form heisst ein *Skalarprodukt*. Ein $\mathbb{C}$-Vektorraum zusammen mit einem Skalarprodukt heisst *unitärer Vektorraum* $(V, \langle\ ,\ \rangle)$.

Definition: Das *Standard-Skalarprodukt auf* $\mathbb{C}^n$ ist für $x = (x_i)_i$ und $y = (y_i)_i$ gegeben durch

$$\langle x, y\rangle \;:=\; x^* y \;=\; \overline{x}_1 y_1 + \ldots + \overline{x}_n y_n.$$

Der zugehörige Betrag ist die ℓ_2-Norm

$$\|x\| \;:=\; \sqrt{|x_1|^2 + \ldots + |x_n|^2}.$$

Definition: Eine hermitesche $n \times n$-Matrix A mit der Eigenschaft $x^* A\, x > 0$ für alle $0 \neq x \in \mathbb{C}^n$ heisst *positiv definit*.

Proposition: Sei B eine geordnete Basis von V. Eine hermitesche Form $\langle\ ,\ \rangle$ auf V ist positiv definit genau dann wenn die Darstellungsmatrix $M_B(\langle\ ,\ \rangle)$ positiv definit ist.

10.4 Grundeigenschaften

Ab jetzt sei $(V, \langle\ ,\ \rangle)$ ein unitärer Vektorraum mit zugehöriger Betragsfunktion $\| \|$.

Die Begriffe und Grundeigenschaften aus §9.6, insbesondere die Cauchy-Schwarz Ungleichung und die Definition von normiert und orthogonal, gelten wortwörtlich auch für unitäre Vektorräume, mit denselben Beweisen.

10.5 Orthonormalbasen

Die Begriffe geordnetes oder ungeordnetes Orthonormalsystem oder -basis und deren Grundeigenschaften aus §9.7 gelten wortwörtlich auch für unitäre Vektorräume. In den Beweisen muss man lediglich berücksichtigen, dass ein komplexes Skalarprodukt in der ersten Variable nur semilinear ist, und in den Formeln darf man die Reihenfolge im komplexen Skalarprodukt im allgemeinen nicht vertauschen. Genauer haben wir:

Satz: *(Gram-Schmidt-Orthogonalisierung)* Für jedes linear unabhängige Tupel $T = (v_1, \ldots, v_n)$ in V existiert genau ein Orthonormalsystem $B = (b_1, \ldots, b_n)$, so dass für alle $1 \leqslant j \leqslant n$ gilt

$$v_j = \sum_{i=1}^{j} a_{ij} b_i \quad \text{für geeignete } a_{ij} \in \mathbb{C} \text{ und } a_{jj} \in \mathbb{R}^{>0}.$$

Ist ausserdem T eine Basis von V, so ist B eine Orthonormalbasis von V.

Folge: Jeder endlich-dimensionale unitäre Vektorraum hat eine Orthonormalbasis.

Bemerkung: Die Bedingung im obigen Satz ist äquivalent dazu, dass die Basiswechselmatrix $M_{B,T}(\mathrm{id}_V)$ eine obere Dreiecksmatrix mit allen Diagonaleinträgen in $\mathbb{R}^{>0}$ ist.

Satz: *(Cholesky-Zerlegung)* Für jede positiv definite hermitesche Matrix A existiert eine komplexe obere Dreiecksmatrix R mit allen Diagonaleinträgen in $\mathbb{R}^{>0}$, so dass $A = R^* R$ ist.

10.6 Unitäre Gruppe

Die Begriffe und Grundeigenschaften aus §9.8 übertragen sich auf unitäre Vektorräume, indem man jeweils reell, orthogonal, A^T durch komplex, unitär, A^* ersetzt. Genauer haben wir:

Definition: Ein Isomorphismus $f\colon V \xrightarrow{\sim} W$ zwischen zwei unitären Vektorräumen mit der Eigenschaft

$$\forall v, v' \in V\colon\ \langle f(v), f(v')\rangle = \langle v, v'\rangle$$

heisst *unitär* oder eine *Isometrie*.

Proposition: Zwischen beliebigen unitären Vektorräumen derselben endlichen Dimension existiert eine Isometrie. Jede Komposition von Isometrien ist eine Isometrie. Der identische Endomorphismus ist eine Isometrie.

Definition: Eine komplexe $n \times n$-Matrix A, für welche die Abbildung $L_A\colon \mathbb{C}^n \to \mathbb{C}^n$ für das jeweilige Standard-Skalarprodukt eine Isometrie ist, heisst *unitär*. Die Menge $\mathrm{U}(n)$ aller unitären $n \times n$-Matrizen heisst die *unitäre Gruppe vom Grad* n.

Proposition: Für jede komplexe $n \times n$-Matrix Q sind äquivalent:

(a) Q ist unitär.

(b) Die Spalten von Q bilden eine Orthonormalbasis von $\mathbb{C}^n$ mit dem Standard-Skalarprodukt.

(c) $Q^*Q = I_n$.

(d) $QQ^* = I_n$.

Proposition: Die Menge $\mathrm{U}(n)$ ist eine Gruppe bezüglich Matrixmultiplikation.

Bemerkung: Die $\mathrm{U}(n)$ ist eine kompakte Teilmenge von $\mathrm{Mat}_{n\times n}(\mathbb{C}) \cong \mathbb{C}^{n^2}$.

Satz: *(Variante der Gram-Schmidt-Orthogonalisierung)* Sei V ein unitärer Vektorraum der Dimension $\geqslant n$. Für alle $v_1, \ldots, v_n \in V$ existiert ein Orthonormalsystem $(b_1, \ldots, b_n)$ in V, so dass für alle $1 \leqslant j \leqslant n$ gilt

$$v_j = \sum_{i=1}^{j} a_{ij} b_i \quad \text{für geeignete } a_{ij} \in \mathbb{C}.$$

Satz: *(QR-Zerlegung)* Für jede komplexe $n \times n$-Matrix A existiert eine unitäre Matrix Q und eine komplexe obere Dreiecksmatrix R, so dass $A = QR$ ist.

10.7 Unterräume, orthogonales Komplement

Die Begriffe orthogonales Komplement und orthogonale Projektion sowie deren Grundeigenschaften aus §9.10 gelten wortwörtlich auch für unitäre Vektorräume, mit denselben Beweisen.

Für die Beziehung zum Dualraum in §9.11 gilt das dagegen nicht, da für einen unitären Vektorraum die Abbildung $V \to V^* := \mathrm{Hom}_{\mathbb{C}}(V, \mathbb{C})$, $v \mapsto \langle v, \ \rangle$ nicht $\mathbb{C}$-linear, sondern nur semilinear ist.

10.8 Adjungierte Abbildungen

Die Begriffe und Grundeigenschaften aus §9.12 gelten wortwörtlich auch für unitäre Vektorräume, mit denselben Beweisen, wobei jeweils nur A^T durch A^* und symmetrisch durch hermitesch zu ersetzen ist. Insbesondere gilt für jede lineare Abbildung zwischen unitären Vektorräumen $f\colon V \to W$:

Proposition: Es gibt höchstens eine lineare Abbildung $f^*\colon W \to V$ mit der Eigenschaft

$$\forall v \in V \ \forall w \in W\colon \ \langle f(v), w\rangle = \langle v, f^*(w)\rangle.$$

Definition: Diese heisst die *Adjungierte (Abbildung) von* f, wenn sie existiert.

Proposition: Ist f^* die Adjungierte von f, so ist auch f die Adjungierte von f^*, das heisst, es gilt $(f^*)^* = f$. Man nennt f und f^* daher auch *zueinander adjungiert*.

Proposition: Ist $\dim V < \infty$, so existiert die adjungierte Abbildung f^*. Genauer gilt für jede geordnete Orthonormalbasis $(b_1, \ldots, b_n)$ von V

$$\forall w \in W\colon \ f^*(w) \ = \ \sum_{i=1}^{n} \langle f(b_i), w\rangle \cdot b_i.$$

Proposition: Seien B eine geordnete Orthonormalbasis von V und B' eine geordnete Orthonormalbasis von W. Dann gilt

$$M_{B,B'}(f^*) \ = \ M_{B',B}(f)^*.$$

Proposition: Für jede komplexe $m \times n$-Matrix A ist die Adjungierte der linearen Abbildung $L_A\colon \mathbb{C}^n \to \mathbb{C}^m$ die lineare Abbildung $L_{A^*}\colon \mathbb{C}^m \to \mathbb{C}^n$.

Beispiel: Die Abbildung $L_A\colon \mathbb{C}^n \to \mathbb{C}^n$ ist selbstadjungiert genau dann, wenn A selbstadjungiert, das heisst, hermitesch ist.

Bei den Grundeigenschaften ist ausserdem die Semilinearität zu beachten:

Proposition: Für alle adjungierten linearen Abbildungen f, g unitärer Vektorräume gilt, soweit sinnvoll:

(a) $(g \circ f)^* = f^* \circ g^*$,
(b) $(f + g)^* = f^* + g^*$,
(c) $(cf)^* = \overline{c} f^*$ für alle $c \in \mathbb{C}$.

10.9 Spektralsatz für selbstadjungierte Endomorphismen

Auch §9.13 überträgt sich entsprechend:

Spektralsatz: Für jeden selbstadjungierten Endomorphismus f eines unitären Vektorraums V gilt:

(a) Alle komplexen Eigenwerte sind reell.

(b) Ist $\dim V < \infty$, so existiert eine Orthonormalbasis von V bestehend aus Eigenvektoren von f. Insbesondere ist f dann diagonalisierbar.

Hauptachsentransformation: Für jede hermitesche Matrix A existiert eine unitäre Matrix Q, so dass $Q^{-1}AQ = Q^*AQ$ eine reelle Diagonalmatrix ist.

Beispiel: Jede hermitesche 2×2-Matrix hat die Form $\left(\begin{smallmatrix} a & b \\ \overline{b} & d \end{smallmatrix}\right)$ für $a, d \in \mathbb{R}$ und $b \in \mathbb{C}$. Eine direkte Rechnung liefert die reellen Eigenwerte

$$\frac{a + d \pm \sqrt{(a-d)^2 + 4b\overline{b}}}{2}.$$

Weitere Beispiele siehe §10.14.

10.10 Normalform hermitescher Formen

Die Begriffe ausgeartet oder nicht-ausgeartet, positiv oder negativ definit oder semidefinit bzw. indefinit, und den Rang und und die Signatur, sowie deren Eigenschaften aus §9.14 gelten wieder wortwörtlich auch für hermitesche (anstatt symmetrischer) Formen auf unitären Vektorräumen, sowie für hermitesche Matrizen, und mit denselben Beweisen.

10.11 Kriterien für Positiv-Definitheit

Auch §9.15 besitzt ein direktes Analogon, nämlich:

Satz: Für jede hermitesche $n \times n$-Matrix $A = \big(a_{k\ell}\big)_{k,\ell=1,\ldots,n}$ sind äquivalent:

(a) Die Matrix A ist positiv definit.

(b) Alle Eigenwerte von A sind positiv.

(c) Es existiert eine invertierbare Matrix B mit $A = B^*B$.

(d) Es existiert eine invertierbare obere Dreiecksmatrix R mit $A = R^*R$. (Cholesky-Zerlegung)

(e) Es existiert eine invertierbare hermitesche Matrix C mit $A = C^*C = C^2$.

(f) Die Determinante der Matrix $A_m := \big(a_{k\ell}\big)_{k,\ell=1,\ldots,m}$ ist positiv für jedes $1 \leqslant m \leqslant n$. (Hauptminorenkriterium)

10.12 Singulärwertzerlegung

Auch §9.16 besitzt ein direktes Analogon, nämlich:

Satz: Für jede komplexe $m \times n$-Matrix A vom Rang r existieren eine unitäre $m \times m$-Matrix Q, eine unitäre $n \times n$-Matrix R, und eine $m \times n$-Matrix der Form

$$D \;=\; \left(\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_r & \\ \hline & & & \end{array}\right)$$

für reelle Zahlen $\sigma_1 \geqslant \ldots \geqslant \sigma_r > 0$ und allen übrigen Einträgen 0, so dass gilt

$$A \;=\; QDR.$$

Dabei sind die Zahlen $\sigma_1, \ldots, \sigma_r$ durch A eindeutig bestimmt. Genauer sind $\sigma_1^2, \ldots, \sigma_r^2$ genau die von Null verschiedenen Eigenwerte von A^*A, mit Vielfachheiten.

Definition: Die Zahlen $\sigma_1, \ldots, \sigma_r$ heissen die *Singulärwerte* von A.

10.13 Spektralsatz für normale Endomorphismen

Wie vorher sei V ein unitärer Vektorraum.

Definition: Eine lineare Abbildung $f\colon V \to V$, deren Adjungierte f^* existiert und die Gleichung $f^* \circ f = f \circ f^*$ erfüllt, heisst *normal*.

Proposition: (a) Jede selbstadjungierte lineare Abbildung ist normal.

(b) Jede unitäre lineare Abbildung $f\colon V \xrightarrow{\sim} V$ hat Adjungierte f^{-1} und ist normal.

Beispiel: Jede komplexe Diagonalmatrix D kommutiert mit ihrer Adjungierten D^*; also ist der Endomorphismus L_D von $\mathbb{C}^n$ normal.

Proposition: Ist f normal, so ist f^* normal.

Der Begriff ist weniger aus sich heraus interessant als wegen der folgenden Sätze:

Satz: Sei f ein normaler Endomorphismus von V, und sei $\lambda \in \mathbb{C}$.

(a) Der Eigenraum von f zum Eigenwert λ ist gleich dem Eigenraum von f^* zum Eigenwert $\overline{\lambda}$.

(b) Für jeden Eigenvektor v von f ist die Zerlegung $V = \mathbb{C}v \oplus (\mathbb{C}v)^\perp$ invariant unter f, das heisst, es gilt $f(\mathbb{C}v) \subset \mathbb{C}v$ und $f((\mathbb{C}v)^\perp) \subset (\mathbb{C}v)^\perp$.

(c) Die Eigenräume von f zu verschiedenen Eigenwerten sind orthogonal zueinander.

Spektralsatz: Für jeden Endomorphismus f eines endlich-dimensionalen unitären Vektorraums V sind äquivalent:

(a) Der Endomorphismus f ist normal.

(b) Es existiert eine Orthonormalbasis von V bestehend aus Eigenvektoren von f.

Insbesondere ist jeder normale Endomorphismus diagonalisierbar.

Beispiel: Die Matrix $A := \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ kommutiert nicht mit $A^* = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$; also ist der Endomorphismus $L_A\colon \mathbb{C}^2 \to \mathbb{C}^2$ nicht normal. Tatsächlich ist er auch nicht diagonalisierbar.

Beispiel: Sei V der Vektorraum aller beliebig oft differenzierbaren 2π-periodischen Funktionen $\mathbb{R} \to \mathbb{C}$ mit dem Skalarprodukt $\langle f, g\rangle := \int_0^{2\pi} \overline{f(x)} g(x)\, dx$. Dann hat die lineare Abbildung

$$D\colon V \to V,\ f \mapsto \frac{df}{dx}$$

die Adjungierte $-D$. Somit ist D normal. Die Eigenwerte von D sind die Zahlen in für alle $n \in \mathbb{Z}$, jeweils mit der Multiplizität 1 und der Eigenfunktion $f_n(x) := e^{inx}$. Die Funktionen $f_n/\sqrt{2\pi}$ für alle $n \in \mathbb{Z}$ bilden eine Orthonormalbasis des Teilraums aller durch Polynome in e^{ix} und e^{-ix} ausdrückbaren Funktionen.

10.14 Klassifikation unitärer Endomorphismen

Satz: Für jeden unitären Endomorphismus f eines unitären Vektorraums V gilt:

(a) Alle Eigenwerte haben Absolutbetrag 1.

(b) Ist $\dim V < \infty$, so existiert eine Orthonormalbasis von V bestehend aus Eigenvektoren von f. Insbesondere ist f dann diagonalisierbar.

Beispiel: Die Norm $\|A\| := \sqrt{\sum_{k,\ell} |a_{k\ell}|^2}$ auf dem Raum der komplexen $n \times n$-Matrizen $A = (a_{k\ell})_{k,\ell}$ erfüllt die Ungleichung $\|AB\| \leqslant \|A\| \cdot \|B\|$ und nach Induktion folglich auch $\|A^m\| \leqslant \|A\|^m$ für alle $m \geqslant 1$. Für jedes A ist daher die Matrix-wertige Reihe

$$\exp(A) \;:=\; \sum_{m=0}^{\infty} \frac{A^m}{m!} \;=\; I_n + A + \frac{A^2}{2} + \frac{A^3}{6} + \ldots$$

absolut konvergent. Daraus ergeben sich direkt die Gleichungen

$$\exp(A^T) = (\exp A)^T \qquad \text{und} \qquad \exp(A^*) = (\exp A)^*.$$

Weiter gilt für jede invertierbare komplexe $n \times n$-Matrix B

$$\exp(B^{-1}AB) \;=\; B^{-1} \cdot \exp(A) \cdot B.$$

Für je zwei komplexe $n \times n$-Matrizen A und B mit $AB = BA$ gilt ausserdem

$$\exp(A + B) \;=\; \exp(A) \cdot \exp(B).$$

Insbesondere ist also $\exp(A)$ invertierbar mit $(\exp A)^{-1} = \exp(-A)$. Schliesslich gilt:

Proposition: Für jede hermitesche $n \times n$-Matrix A ist die Matrix $\exp(iA)$ unitär, und jede unitäre Matrix ist in dieser Form darstellbar.

11 Multilineare Algebra

11.1 Multilineare Abbildungen

Definition: Betrachte K-Vektorräume $V_1, \ldots, V_r$ und W. Eine Abbildung

$$\varphi\colon V_1 \times \ldots \times V_r \to W,\ (v_1, \ldots, v_r) \mapsto \varphi(v_1, \ldots, v_r),$$

die in jeder Variablen v_i separat linear ist, heisst *multilinear*. Die Menge aller solcher bezeichnen wir mit

$$\mathrm{Mult}_K(V_1, \ldots, V_r; W).$$

Proposition: Dies ist ein Unterraum des Raums aller Abbildungen $V_1 \times \ldots \times V_r \to W$.

Spezialfall: Für $r = 1$ ist $\mathrm{Mult}_K(V; W) = \mathrm{Hom}_K(V, W)$, vergleiche §5.9. Insbesondere ist $\mathrm{Mult}_K(V; K) = \mathrm{Hom}_K(V, K) = V^*$ der Dualraum von V, vergleiche §5.10.

Spezialfall: Für $r = 2$ heisst multilinear auch *bilinear*. Insbesondere ist $\mathrm{Mult}_K(V, V; K)$ der Raum aller Bilinearformen auf V, vergleiche §9.3.

Proposition: *(Funktorialität)* Lineare Abbildungen $f_i\colon V_i' \to V_i$ und $g\colon W \to W'$ induzieren eine lineare Abbildung

$$\begin{aligned} \mathrm{Mult}_K(V_1, \ldots, V_r; W) &\to \mathrm{Mult}_K(V_1', \ldots, V_r'; W'), \\ \varphi &\mapsto g \circ \varphi \circ (f_1 \times \ldots \times f_r). \end{aligned}$$

Proposition: Betrachte Basen B_i von V_i sowie C von W. Betrachte ein System von Koeffizienten $\alpha^c_{b_1,\ldots,b_r} \in K$ für alle $b_i \in B_i$ und $c \in C$ mit der Eigenschaft $\forall b_i \in B_i\colon \big|\{c \in C \mid \alpha^c_{b_1,\ldots,b_r} \neq 0\}\big| < \infty$. Dann existiert genau eine multilineare Abbildung $\varphi\colon V_1 \times \ldots \times V_r \to W$, so dass für alle $b_i \in B_i$ gilt:

$$\varphi(b_1, \ldots, b_r) = \sum_{c \in C}{}' \alpha^c_{b_1,\ldots,b_r} c.$$

Umgekehrt hat jede multilineare Abbildung $\varphi\colon V_1 \times \ldots \times V_r \to W$ diese Gestalt für eindeutige Koeffizienten $\alpha^c_{b_1,\ldots,b_r}$.

Proposition: Für beliebige $V_1, \ldots, V_r, W$ gilt, mit der Konvention $\infty \cdot 0 = 0$:

$$\dim_K \mathrm{Mult}_K(V_1, \ldots, V_r; W) = \Big(\prod_{i=1}^{r} \dim_K(V_i^*)\Big) \cdot \dim_K(W).$$

Folge:

(a) Es ist $\mathrm{Mult}_K(V_1, \ldots, V_r; W) \neq 0$ genau dann, wenn alle V_i, $W \neq 0$ sind.

(b) Es ist $\mathrm{Mult}_K(V_1, \ldots, V_r; W) \neq 0$ und endlich-dimensional genau dann, wenn alle V_i, $W \neq 0$ und endlich-dimensional sind.

11.2 Symmetrische und alternierende Abbildungen

Definition: Eine multilineare Abbildung $\varphi\colon V^r \to W$ heisst *symmetrisch*, wenn gilt

$$\forall v_1,\dots,v_r \in V\ \forall \sigma \in S_r:\ \varphi(v_{\sigma 1},\dots,v_{\sigma r}) = \varphi(v_1,\dots,v_r).$$

Sie heisst *alternierend*, wenn gilt

$$\forall v_1,\dots,v_r \in V\colon\ \big(\exists i \neq i'\colon\ v_i = v_{i'}\big) \longrightarrow \varphi(v_1,\dots,v_r) = 0.$$

Die Menge aller solcher bezeichnen wir mit

$$\mathrm{Sym}^r_K(V,W) \quad \text{bzw.} \quad \mathrm{Alt}^r_K(V,W).$$

Proposition: Dies sind Unterräume von $\mathrm{Mult}_K(V,\dots,V;W)$.

Bemerkung: Da jede Permutation ein Produkt von Transpositionen benachbarter Indizes ist, ist φ symmetrisch genau dann, wenn gilt

$$\forall v_1,\dots,v_r \in V\ \forall 2 \leqslant i \leqslant r:\ \varphi(v_1,\dots,v_{i-2},v_i,v_{i-1},v_{i+1},\dots,v_r) = \varphi(v_1,\dots,v_r).$$

Variante: Eine multilineare Abbildung $\varphi\colon V^r \to W$ heisst *antisymmetrisch*, wenn gilt

$$\forall v_1,\dots,v_r \in V\ \forall \sigma \in S_r:\ \varphi(v_{\sigma 1},\dots,v_{\sigma r}) = \mathrm{sgn}(\sigma)\cdot\varphi(v_1,\dots,v_r).$$

Proposition: (a) Es gilt immer alternierend $\Rightarrow$ antisymmetrisch.

(b) Ist $1+1 \neq 0$ in K, so gilt antisymmetrisch $\Leftrightarrow$ alternierend.

(c) Ist $1+1 = 0$ in K, so gilt antisymmetrisch $\Leftrightarrow$ symmetrisch.

Der Begriff „antisymmetrisch“ ist daher weniger wichtig als die übrigen.

Proposition: *(Funktorialität)* Lineare Abbildungen $f\colon V' \to V$ und $g\colon W \to W'$ induzieren lineare Abbildungen

$$\begin{aligned} \mathrm{Sym}^r_K(V,W) &\to \mathrm{Sym}^r_K(V',W'),\\ \mathrm{Alt}^r_K(V,W) &\to \mathrm{Alt}^r_K(V',W'),\\ \varphi &\mapsto g\circ\varphi\circ(f\times\ldots\times f). \end{aligned}$$

Proposition: Betrachte Basen B von V und C von W, sowie eine multilineare Abbildung $\varphi\colon V^r \to W$ mit Koeffizienten $\alpha^c_{b_1,\dots,b_r} \in K$ wie in §11.1. Dann ist φ symmetrisch genau dann, wenn gilt

$$\forall b_i \in B\ \forall c \in C\ \forall \sigma \in S_r\colon\ \alpha^c_{b_{\sigma 1},\dots,b_{\sigma r}} = \alpha^c_{b_1,\dots,b_r},$$

und φ ist alternierend genau dann, wenn gilt

$$\forall b_i \in B\ \forall c \in C\colon\ \begin{cases} \forall \sigma \in S_r\colon\ \alpha^c_{b_{\sigma 1},\dots,b_{\sigma r}} = \mathrm{sgn}(\sigma)\cdot\alpha^c_{b_1,\dots,b_r} & \text{und}\\ \big(\exists i \neq i'\colon\ b_i = b_{i'}\big) \longrightarrow \alpha^c_{b_1,\dots,b_r} = 0. \end{cases}$$

Im Spezialfall $r = 2$ bedeutet dies, dass für jedes $c \in C$ die Matrix $A_c := (\alpha^c_{b_1,b_2})_{b_1,b_2 \in B}$ symmetrisch ist, bzw. antisymmetrisch mit Diagonale Null.

Proposition:

$$\dim_K \mathrm{Sym}^r_K(V,W) \;=\; \binom{\dim_K(V^*)+r-1}{r} \cdot \dim_K(W),$$
$$\dim_K \mathrm{Alt}^r_K(V,W) \;=\; \binom{\dim_K(V^*)}{r} \cdot \dim_K(W).$$

Folge: Für alle $r > \dim_K(V)$ gilt $\mathrm{Alt}^r_K(V,W) = 0$.
Für $r = \dim_K(V)$ gilt $\dim_K \mathrm{Alt}^r_K(V,K) = 1$.

Beispiel: Die Determinante induziert eine von Null verschiedene alternierende multilineare Abbildung

$$(K^n)^n \to K,\ (v_1,\ldots,v_n) \mapsto \det((v_1,\ldots,v_n)).$$

Aus Dimensionsgründen bildet diese eine Basis von $\mathrm{Alt}^n_K(K^n,K)$.

Satz: Für jeden Endomorphismus f eines K-Vektorraums V der Dimension $n < \infty$ und jedes $\varphi \in \mathrm{Alt}^n_K(V,K)$ gilt

$$\varphi \circ (f \times \ldots \times f) \;=\; \det(f) \cdot \varphi.$$

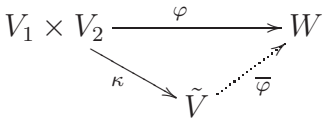
Bemerkung: Damit kann man die Determinante alternativ und basisfrei konstruieren.

11.3 Tensorprodukt

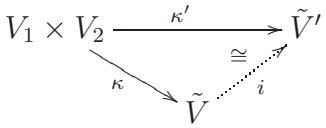
Betrachte zwei K-Vektorräume V_1 und V_2.

Definition: Ein *Tensorprodukt von V_1 und V_2 über K* besteht aus einem K-Vektorraum $\tilde{V}$ und einer bilinearen Abbildung $\kappa\colon V_1 \times V_2 \to \tilde{V}$ mit der *universellen Eigenschaft:*

Für jeden K-Vektorraum W und jede bilineare Abbildung $\varphi\colon V_1 \times V_2 \to W$ existiert genau eine lineare Abbildung $\overline{\varphi}\colon \tilde{V} \to W$ mit $\overline{\varphi} \circ \kappa = \varphi$, das heisst, so dass das folgende Diagramm kommutiert:



Proposition: Ein Tensorprodukt ist eindeutig bis auf eindeutige Isomorphie, mit anderen Worten: Ist sowohl $(\tilde{V}, \kappa)$ wie $(\tilde{V}', \kappa')$ ein Tensorprodukt von V_1 und V_2, so existiert ein eindeutiger Isomorphismus $i\colon \tilde{V} \xrightarrow{\sim} \tilde{V}'$ mit $i \circ \kappa = \kappa'$, das heisst, so dass das folgende Diagramm kommutiert:



Satz: Ein Tensorprodukt existiert immer.

Konvention: Wir fixieren ein für alle Mal ein Tensorprodukt $(\tilde{V}, \kappa)$ und bezeichnen den Vektorraum $\tilde{V}$ mit $V_1 \otimes_K V_2$ oder kurz $V_1 \otimes V_2$, sowie die Abbildung κ mit

$$V_1 \times V_2 \to V_1 \otimes_K V_2,\ (v_1, v_2) \mapsto v_1 \otimes v_2.$$

Danach vergessen wir die Notation $(\tilde{V}, \kappa)$.

Bemerkung: Wir kümmern uns also nicht darum, wie das Tensorprodukt genau konstruiert ist, sondern benutzen nur seine universelle Eigenschaft. Die Eindeutigkeit bis auf eindeutige (!) Isomorphie bewirkt, dass jedes Element einer zweiten Wahl von $(\tilde{V}, \kappa)$ einem eindeutigen Element der ersten Wahl entspricht, und dass diese Elemente jeweils dieselben Formeln erfüllen und dieselben sonstigen Eigenschaften besitzen.

Rechenregeln: Die Bilinearität von κ übersetzt sich in die folgenden Rechenregeln für alle $v_i, v_i' \in V_i$ und $\lambda \in K$:

$(v_1 + v_1') \otimes v_2 = v_1 \otimes v_2 + v_1' \otimes v_2$	$\lambda v_1 \otimes v_2 = \lambda(v_1 \otimes v_2)$
$v_1 \otimes (v_2 + v_2') = v_1 \otimes v_2 + v_1 \otimes v_2'$	$v_1 \otimes \lambda v_2 = \lambda(v_1 \otimes v_2)$

Proposition: *(Adjunktionsformel)* Es existieren eindeutige Isomorphismen

$$\begin{array}{ccccc} \mathrm{Hom}_K(V_1 \otimes_K V_2, W) & \cong & \mathrm{Mult}_K(V_1 \times V_2, W) & \cong & \mathrm{Hom}_K(V_1, \mathrm{Hom}_K(V_2, W)) \\ \cup & & \cup & & \cup \\ f & & \varphi & & \psi \end{array}$$

mit

$$f(v_1 \otimes v_2) \;=\; \varphi(v_1, v_2) \;=\; \psi(v_1)(v_2).$$

Proposition: *(Funktorialität)* Zu linearen Abbildungen $f_i \colon V_i \to V_i'$ existiert genau eine lineare Abbildung

$$f_1 \otimes f_2 \colon\; V_1 \otimes_K V_2 \to V_1' \otimes_K V_2' \quad \text{mit} \quad v_1 \otimes v_2 \mapsto f_1(v_1) \otimes f_2(v_2).$$

Proposition: Für alle linearen Abbildungen $V_i \xrightarrow{f_i} V_i' \xrightarrow{g_i} V_i''$ gilt

(a) $\mathrm{id}_{V_1} \otimes \mathrm{id}_{V_2} = \mathrm{id}_{V_1 \otimes V_2}$.
(b) $f_1 \otimes 0_{V_2} = 0_{V_1} \otimes f_2 = 0_{V_1 \otimes V_2}$.
(c) $(g_1 \otimes g_2) \circ (f_1 \otimes f_2) = (g_1 \circ f_1) \otimes (g_2 \circ f_2)$.
(d) Ist jeweils f_i ein Isomorphismus mit Inversem g_i, so ist $f_1 \otimes f_2$ ein Isomorphismus mit Inversem $g_1 \otimes g_2$.

Satz: Sei jeweils B_i eine Basis von V_i. Dann sind die $b_1 \otimes b_2$ für alle $(b_1, b_2) \in B_1 \times B_2$ verschieden, und $\big\{ b_1 \otimes b_2 \;\big|\; (b_1, b_2) \in B_1 \times B_2 \big\}$ ist eine Basis von $V_1 \otimes_K V_2$. Insbesondere gilt

$$\dim_K(V_1 \otimes_K V_2) = \dim_K(V_1) \cdot \dim_K(V_2).$$

Beispiel: Für alle natürlichen Zahlen m, n existiert ein natürlicher Isomorphismus

$$K^m \otimes_K K^n \xrightarrow{\sim} \mathrm{Mat}_{m \times n}(K) \quad \text{mit} \quad v \otimes w \mapsto v \cdot w^T.$$

Proposition: Für alle V und W existiert ein natürlicher injektiver Homomorphismus

$$V^* \otimes_K W \longrightarrow \mathrm{Hom}_K(V, W) \quad \text{mit} \quad \ell \otimes w \mapsto (v \mapsto \ell(v) \cdot w).$$

Sein Bild ist der Unterraum aller Homomorphismen von endlichem Rang. Insbesondere ist er ein Isomorphismus genau dann, wenn V oder W endlich-dimensional ist.

Definition: Ein Element von $V_1 \otimes_K V_2$ heisst ein *Tensor*.
Ein Element der Form $v_1 \otimes v_2$ heisst ein *reiner Tensor*.

Proposition: Die reinen Tensoren erzeugen $V_1 \otimes_K V_2$.

Proposition: Betrachte Vektoren $v_i, v_i' \in V_i$.

(a) Es ist $v_1 \otimes v_2 \neq 0$ genau dann, wenn $v_1, v_2 \neq 0$ sind.
(b) Im Fall (a) ist $v_1 \otimes v_2 = v_1' \otimes v_2'$ genau dann, wenn $\exists \lambda \in K^\times \colon (v_1', v_2') = (\lambda v_1, \lambda^{-1} v_2)$.
(c) Sind jeweils v_i, v_i' linear unabhängig, so ist $v_1 \otimes v_2 + v_1' \otimes v_2'$ kein reiner Tensor.

11.4 Höhere Tensorprodukte

Proposition: Es existieren eindeutige Isomorphismen, charakterisiert wie folgt:

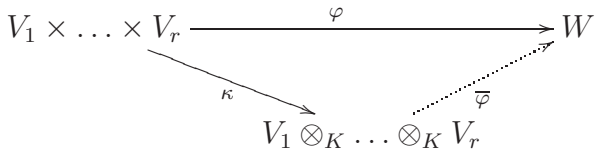
$$\begin{array}{rcll} V_1 \otimes_K K \xrightarrow{\sim} V_1 & \text{mit} & v_1 \otimes 1 \mapsto v_1 & \text{(Identität)} \\ V_1 \otimes_K V_2 \xrightarrow{\sim} V_2 \otimes_K V_1 & \text{mit} & v_1 \otimes v_2 \mapsto v_2 \otimes v_1 & \text{(Kommutativität)} \\ (V_1 \otimes V_2) \otimes V_3 \xrightarrow{\sim} V_1 \otimes (V_2 \otimes V_3) & \text{mit} & (v_1 \otimes v_2) \otimes v_3 \mapsto v_1 \otimes (v_2 \otimes v_3) & \text{(Assoziativität)} \end{array}$$

Damit lässt sich das Tensorprodukt einer beliebigen endlichen Folge von Vektorräumen $V_1 \otimes_K \ldots \otimes_K V_r$ ohne Klammern definieren. Dieses trägt eine natürliche multilineare Abbildung

$$\kappa \colon V_1 \times \ldots \times V_r \to V_1 \otimes_K \ldots \otimes_K V_r,\ (v_1, \ldots, v_r) \mapsto v_1 \otimes \ldots \otimes v_r.$$

Zusammen haben diese die *universelle Eigenschaft:*

Für jeden K-Vektorraum W und jede multilineare Abbildung $\varphi \colon V_1 \times \ldots \times V_r \to W$ existiert genau eine lineare Abbildung $\overline{\varphi} \colon V_1 \otimes_K \ldots \otimes_K V_r \to W$ mit $\overline{\varphi} \circ \kappa = \varphi$, das heisst, so dass das folgende Diagramm kommutiert:



Proposition:

$$\dim_K(V_1 \otimes_K \ldots \otimes_K V_r) \;=\; \prod_{i=1}^{r} \dim_K(V_i).$$

Definition: Für jede natürliche Zahl r ist die *r-te Tensorpotenz von V* definiert durch $V^{\otimes 0} := K$ beziehungsweise $V^{\otimes r} := V \otimes_K \ldots \otimes_K V$ mit r Faktoren für $r \geqslant 1$.

Definition: Der Raum $T^{r,s}(V) := V^{\otimes r} \otimes_K (V^*)^{\otimes s}$ heisst der Raum der *r-fach kovarianten und s-fach kontravarianten Tensoren*, oder kurz der *Tensoren vom Typ (r,s)*. (Nach einer anderen Konvention nennt man sie *vom Typ (s,r)*.)

Bemerkung: Ist $\dim_K(V) < \infty$, so liefert jede Basis von V die zugehörige duale Basis von V^* und somit eine Basis von $T^{r,s}(V)$. Insbesondere ist dann

$$\dim_K T^{r,s}(V) \;=\; (\dim_K V)^{r+s}.$$

Bemerkung: Je nach Situation kann eine $n \times n$-Matrix einen Tensor vom Typ $(0,2)$ oder $(1,1)$ oder $(2,0)$ darstellen für den Raum $V = K^n$. Für jeden dieser Typen wird der Basiswechsel mit der Basiswechselmatrix U durch eine andere Formel beschrieben, nämlich durch $A \mapsto U^T A U$ bzw. $U^{-1} A U$ bzw. $U^{-1} A \, (U^T)^{-1}$. Am besten vermeidet man diese Verwirrung, indem man so lange wie möglich bei den abstrakten Begriffen bleibt, wo V und V^* durch die Notation klar unterschieden werden.

Proposition: Für jeden eindimensionalen K-Vektorraum V existiert ein natürlicher Isomorphismus $V^* \otimes_K V \xrightarrow{\sim} K$ mit $\ell \otimes v \mapsto \ell(v)$.

Beispiel: Jede skalare physikalische Grösse liegt in einem gewissen eindimensionalen $\mathbb{R}$-Vektorraum, und die Wahl einer Grundeinheit entspricht der Wahl eines Basisvektors. Nur physikalische Grössen derselben Art, also Elemente desselben Vektorraums, sind miteinander vergleichbar. Beispiele:

Grösse	Vektorraum	Grundeinheit	Relation
Zeit	T	Sekunde, Stunde	$h = 3600s$
Länge	L	Meter, Zoll	$in = 0.0254m$
Masse	M	Kilogramm, Pfund	$lb = 0.45359237kg$

Zusammengesetzte skalare physikalische Grössen liegen auf natürliche Weise in gewissen Tensorräumen. Beispiele:

skalare Grösse	Vektorraum	Grundeinheit
Flächeninhalt	$L^{\otimes 2}$	m^2
Frequenz	T^*	$1/s$
Geschwindigkeit	$L \otimes T^*$	m/s
Beschleunigung	$L \otimes (T^*)^{\otimes 2}$	m/s^2
Kraft	$M \otimes L \otimes (T^*)^{\otimes 2}$	$N = kg\,m/s^2$
Energie	$M \otimes L^{\otimes 2} \otimes (T^*)^{\otimes 2}$	$J = kg\,m^2/s^2$

Klassische vektorielle physikalische Grössen liegen in euklidischen Vektorräumen. Zum Beispiel sei R der dreidimensionale Ortsraum. Einige weitere Grössen sind dann:

vektorielle Grösse	Vektorraum
Geschwindigkeit	$R \otimes T^*$
Beschleunigung	$R \otimes (T^*)^{\otimes 2}$
Impuls	$M \otimes R \otimes T^*$
Kraft	$M \otimes R \otimes (T^*)^{\otimes 2}$
Spannung	$M \otimes L \otimes (T^*)^{\otimes 2} \otimes (R^*)^{\otimes 2}$

Vergleiche dazu auch den Beitrag von Terence Tao:

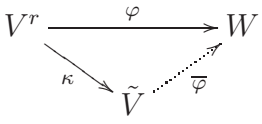
`http://terrytao.wordpress.com/2012/12/29/a-mathematical-formalisation-of-dimensional-analysis`

Quantenmechanische physikalische Zustände liegen in unitären Vektorräumen, die oft einen weniger direkten Zusammenhang mit dem klassischen Ortsraum haben.

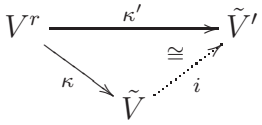
11.5 Symmetrische und alternierende Potenzen

Definition: Für jede natürliche Zahl r besteht eine *r-te symmetrische*, bzw. *alternierende, Potenz von V über K* aus einem K-Vektorraum $\tilde{V}$ und einer symmetrischen, bzw. alternierenden, multilinearen Abbildung $\kappa\colon V^r \to \tilde{V}$ mit der *universellen Eigenschaft:*

Für jeden K-Vektorraum W und jede symmetrische, bzw. alternierende, multilineare Abbildung $\varphi\colon V^r \to W$ existiert genau eine lineare Abbildung $\overline{\varphi}\colon \tilde{V} \to W$ mit $\overline{\varphi} \circ \kappa = \varphi$, das heisst, so dass das folgende Diagramm kommutiert:



Proposition: Eine symmetrische, bzw. alternierende, Potenz ist eindeutig bis auf eindeutige Isomorphie, mit anderen Worten: Ist sowohl $(\tilde{V}, \kappa)$ wie $(\tilde{V}', \kappa')$ eine solche Potenz von V, so existiert ein eindeutiger Isomorphismus $i\colon \tilde{V} \xrightarrow{\sim} \tilde{V}'$ mit $i \circ \kappa = \kappa'$, das heisst, so dass das folgende Diagramm kommutiert:



Satz: Ein symmetrische, bzw. alternierende, Potenz existiert immer.

Konvention: Wir fixieren ein für alle Mal eine r-te symmetrische Potenz und bezeichnen den zugehörigen Vektorraum mit $S^r_K V$ oder $S^r V$, sowie die zugehörige symmetrische multilineare Abbildung mit

$$V^r \to S^r_K V,\ (v_1, \ldots, v_r) \mapsto v_1 \cdots v_r.$$

Wir fixieren ein für alle Mal eine r-te alternierende Potenz und bezeichnen den zugehörigen Vektorraum mit $\Lambda^r_K V$ oder $\Lambda^r V$, sowie die zugehörige alternierende multilineare Abbildung mit

$$V^r \to \Lambda^r_K V,\ (v_1, \ldots, v_r) \mapsto v_1 \wedge \ldots \wedge v_r.$$

Rechenregeln: Die Multilinearität und Symmetrieeigenschaften dieser Abbildungen bedeuten gewisse Rechenregeln. Im Fall $r = 2$ sind die für alle $v, w, w' \in V$ und $\lambda \in K$:

$v \cdot (w + w') = v \cdot w + v \cdot w'$	$v \cdot \lambda w = \lambda(v \cdot w)$	$w \cdot v = v \cdot w$
$v \wedge (w + w') = v \wedge w + v \wedge w'$	$v \wedge \lambda w = \lambda(v \wedge w)$	$v \wedge v = 0$ $w \wedge v = -v \wedge w$

Proposition: *(Adjunktionsformel)* Es existieren eindeutige Isomorphismen

$$\begin{aligned}\mathrm{Hom}_K(S^r_K V, W) &\xrightarrow{\sim} \mathrm{Sym}^r_K(V,W), \quad f \mapsto \big((v_1,\ldots,v_r) \mapsto f(v_1 \cdots v_r)\big),\\ \mathrm{Hom}_K(\Lambda^r_K V, W) &\xrightarrow{\sim} \mathrm{Alt}^r_K(V,W), \quad f \mapsto \big((v_1,\ldots,v_r) \mapsto f(v_1 \wedge \ldots \wedge v_r)\big).\end{aligned}$$

Proposition: *(Funktorialität)* Zu jeder linearen Abbildung $f\colon V \to V'$ existieren eindeutige lineare Abbildungen

$$\begin{aligned}S^r f\colon\ & S^r_K V \to S^r_K V' \quad \text{mit} \qquad v_1 \cdots v_r \mapsto f_1(v_1)\cdots f_r(v_r),\\ \Lambda^r f\colon\ & \Lambda^r_K V \to \Lambda^r_K V' \quad \text{mit} \quad v_1 \wedge \ldots \wedge v_r \mapsto f_1(v_1) \wedge \ldots \wedge f_r(v_r).\end{aligned}$$

Proposition: Für alle linearen Abbildungen $V \xrightarrow{f} V' \xrightarrow{g} V''$ gilt

(a) $S^r\mathrm{id}_V = \mathrm{id}_{S^r V}$ und $\Lambda^r\mathrm{id}_V = \mathrm{id}_{\Lambda^r V}$.
(b) Ist $f = 0$, so ist auch $S^r f = 0$ und $\Lambda^r f = 0$.
(c) $S^r g \circ S^r f = S^r(g \circ f)$ und $\Lambda^r g \circ \Lambda^r f = \Lambda^r(g \circ f)$.
(d) Ist f ein Isomorphismus mit Inversem g, so ist $S^r f$ bzw. $\Lambda^r f$ ein Isomorphismus mit Inversem $S^r g$ bzw. $\Lambda^r g$.

Spezialfall: Es gilt

$$\begin{aligned}V^{\otimes 0} = S^0 V = \Lambda^0 V = K \quad &\text{und}\\ V^{\otimes 1} = S^1 V = \Lambda^1 V = V.&\end{aligned}$$

Satz: Sei B eine Basis von V, versehen mit einer Totalordnung $\preceq$ bzw. $\prec$. Dann bilden die Elemente

$$\begin{aligned}b_1 \cdots b_r \quad &\text{für alle } b_i \in B \text{ mit } b_1 \preceq \ldots \preceq b_r \text{ eine Basis von } S^r_K V, \text{ bzw.}\\ b_1 \wedge \ldots \wedge b_r \ &\text{für alle } b_i \in B \text{ mit } b_1 \prec \ldots \prec b_r \text{ eine Basis von } \Lambda^r_K V.\end{aligned}$$

Insbesondere gilt

$$\begin{aligned}\dim_K S^r_K V &= \binom{\dim_K(V) + r - 1}{r},\\ \dim_K \Lambda^r_K V &= \binom{\dim_K(V)}{r}.\end{aligned}$$

Folge: Für alle $r > \dim_K(V)$ gilt $\Lambda^r_K V = 0$.
Für $r = \dim_K(V)$ gilt $\dim_K \Lambda^r_K V = 1$.

Satz: Für jeden Endomorphismus f eines K-Vektorraums V der Dimension $n < \infty$ ist die Abbildung

$$\Lambda^n f\colon \Lambda^n_K V \to \Lambda^n_K V$$

gleich der Multiplikation mit $\det(f)$.

Bemerkung: Damit kann man die Determinante alternativ und basisfrei konstruieren.

11.6 Tensoralgebra, symmetrische, äussere Algebra

Proposition: Für alle $r, s \geqslant 0$ existieren eindeutige bilineare Abbildungen

$$\begin{aligned}
\otimes\colon\ & V^{\otimes r} \times V^{\otimes s} \to V^{\otimes(r+s)} \quad \text{mit} \quad (v_1 \otimes \ldots \otimes v_r) \otimes (v_{r+1} \otimes \ldots \otimes v_{r+s}) = v_1 \otimes \ldots \otimes v_{r+s},\\
\cdot\colon\ & S^r V \times S^s V \to S^{r+s} V \quad \text{mit} \quad (v_1 \cdots v_r) \cdot (v_{r+1} \cdots v_{r+s}) = v_1 \cdots v_{r+s},\\
\wedge\colon\ & \Lambda^r V \times \Lambda^s V \to \Lambda^{r+s} V \quad \text{mit} \quad (v_1 \wedge \ldots \wedge v_r) \wedge (v_{r+1} \wedge \ldots \wedge v_{r+s}) = v_1 \wedge \ldots \wedge v_{r+s}.
\end{aligned}$$

Definition: (Für die äussere direkte Summe $\boxplus$ siehe §5.4.)

Tensoralgebra	$TV := \boxplus_{r\geqslant 0} V^{\otimes r}$	$(\xi_r)_{r\geqslant 0} \otimes (\eta_s)_{s\geqslant 0} := \left(\sum_{r=0}^{t} \xi_r \otimes \eta_{t-r}\right)_{t\geqslant 0}$
symmetrische Algebra	$SV := \boxplus_{r\geqslant 0} S^r V$	$(\xi_r)_{r\geqslant 0} \cdot (\eta_s)_{s\geqslant 0} := \left(\sum_{r=0}^{t} \xi_r \cdot \eta_{t-r}\right)_{t\geqslant 0}$
äussere Algebra	$\Lambda V := \boxplus_{r\geqslant 0} \Lambda^r V$	$(\xi_r)_{r\geqslant 0} \wedge (\eta_s)_{s\geqslant 0} := \left(\sum_{r=0}^{t} \xi_r \wedge \eta_{t-r}\right)_{t\geqslant 0}$

Proposition: Mit der Addition des unterliegenden Vektorraums und der angegebenen Multiplikation sowie dem Einselement von $K = V^{\otimes 0} = S^0 V = \Lambda^0 V$ ist dies jeweils ein assoziativer unitärer graduierter Ring.

Die in dem jeweiligen Ring geltenden Rechenregeln ergeben sich aus denen in §11.3 und §11.5 sowie der obigen Definition.

Spezialfall: Für $\dim_K V = 0$ gilt $V^{\otimes r} = S^r V = \Lambda^r V = 0$ für alle $r > 0$, und daher $TV = SV = \Lambda V = K$.

Spezialfall: Sei $\dim_K V = 1$ mit Basis b. Für alle $r \geqslant 0$ gilt dann $\dim_K(V^{\otimes r}) = 1$ mit Basis $b^{\otimes r}$ und $\dim_K(S^r V) = 1$ mit Basis b^r, und es existieren eindeutige Isomorphismen

$$\begin{aligned}
& K[X] \xrightarrow{\sim} TV \xrightarrow{\sim} SV, \quad \textstyle\sum'_{i\geqslant 0} a_i X^i \mapsto \sum'_{i\geqslant 0} a_i b^{\otimes i} \mapsto \sum'_{i\geqslant 0} a_i b^i, \quad \text{bzw.}\\
& K \oplus Kb \xrightarrow{\sim} \Lambda V \quad \text{mit} \quad b \wedge b = 0.
\end{aligned}$$

Proposition: (a) Für $V \neq 0$ ist $\dim_K(TV) = \dim_K(SV) = \infty$.

(b) Für $\dim_K(V) < \infty$ ist $\dim_K(\Lambda V) = 2^{\dim_K(V)}$, andernfalls ist $\dim_K(\Lambda V) = \infty$.

Proposition: (a) Der Ring SV ist immer kommutativ.

(b) Für $\dim_K V \leqslant 1$ sind TV und ΛV kommutativ.

(c) Ist $1 + 1 = 0$ in K, so ist ΛV kommutativ.

(d) Für $\dim_K V \geqslant 2$ ist TV nicht kommutativ.

(e) Für $\dim_K V \geqslant 2$ und $1 + 1 \neq 0$ in K ist ΛV nicht kommutativ.

Proposition: Für alle $r, s \geqslant 0$ und alle $\xi \in \Lambda^r V$ und $\eta \in \Lambda^s V$ gilt

$$\eta \wedge \xi \;=\; (-1)^{rs} \cdot \xi \wedge \eta.$$

Insbesondere kommutiert für gerades r jedes Element von $\Lambda^r V$ mit ganz ΛV.

Proposition: Ist $\dim_K(V) < \infty$, so existieren für alle $r \geqslant 0$ natürliche Isomorphismen

$$\Lambda^r(V^*) \xrightarrow[\sim]{\varphi_r} \mathrm{Alt}^r(V, K) \xrightarrow[\sim]{} (\Lambda^r V)^*.$$

Dabei ist der zweite ein Spezialfall der Adjunktionsformel, und der erste entsteht aus der in $(\ell_1, \ldots, \ell_r)$ und $(v_1, \ldots, v_r)$ separat alternierenden Multilinearform

$$(V^*)^r \times V^r \longrightarrow K, \quad (\ell_1, \ldots, \ell_r, v_1, \ldots, v_r) \mapsto \sum_{\sigma \in S_r} \mathrm{sgn}(\sigma) \cdot \ell_1(v_{\sigma 1}) \cdots \ell_r(v_{\sigma r}).$$

Proposition: Für alle $r, s \geqslant 0$ kommutiert das folgende Diagramm

$$\begin{array}{ccccc} \Lambda^r(V^*) & \times & \Lambda^s(V^*) & \xrightarrow{\wedge} & \Lambda^{r+s}(V^*) \\ \wr\downarrow{\scriptstyle \varphi_r} & & \wr\downarrow{\scriptstyle \varphi_s} & & \wr\downarrow{\scriptstyle \varphi_{r+s}} \\ \mathrm{Alt}^r(V, K) & \times & \mathrm{Alt}^s(V, K) & \xrightarrow{\wedge} & \mathrm{Alt}^{r+s}(V, K) \end{array}$$

wobei die untere Abbildung $\wedge$ definiert ist durch die Formel

$$(\varphi \wedge \psi)(v_1, \ldots, v_{r+s}) \;:=\; \sum_{\sigma} \mathrm{sgn}(\sigma) \cdot \varphi(v_{\sigma 1}, \ldots, v_{\sigma r}) \cdot \psi(v_{\sigma(r+1)}, \ldots, v_{\sigma(r+s)})$$

und die Summe sich über alle $\sigma \in S_r$ mit $\sigma 1 < \ldots < \sigma r$ und $\sigma(r+1) < \ldots < \sigma(r+s)$ erstreckt.

Bemerkung: In der Vorlesung Analysis II wurden Differentialformen als alternierende Multilinearformen eingeführt und ihr äusseres Produkt durch die obige alternierende Summe. Diese etwas künstlich wirkende Formel wird nach Übertragung auf die äussere Potenz $\Lambda^r(V^*)$ viel einfacher und natürlicher.

11.7 Vektorprodukt im $\mathbb{R}^3$

Für jeden K-Vektorraum V der Dimension $n < \infty$ induziert das äussere Produkt einen natürlichen Isomorphismus

$$\Lambda^{n-1}V \xrightarrow{\sim} \mathrm{Hom}_K(V, \Lambda^n V),\ \xi \mapsto (v \mapsto \xi \wedge v).$$

Die Wahl eines Isomorphismus $\Lambda^n V \cong K$ liefert dann einen Isomorphismus

$$\Lambda^{n-1}V \xrightarrow{\sim} \mathrm{Hom}_K(V, K) = V^*.$$

Ist weiter V ein euklidischer Vektorraum, so induziert das Skalarprodukt wie in §9.11 einen Isomorphismus $V^* \cong V$, insgesamt also einen Isomorphismus

$$\Lambda^{n-1}V \xrightarrow{\sim} V.$$

Im Spezialfall $n = 3$ liefert dies eine alternierende bilineare Abbildung

$$V \times V \to \Lambda^2 V \xrightarrow{\sim} V,\ (v, w) \mapsto v \wedge w \mapsto v \times w.$$

Für $V = \mathbb{R}^3$ mit dem Isomorphismus $\det\colon \Lambda^3(\mathbb{R}^3) \xrightarrow{\sim} \mathbb{R}$ und dem Standard-Skalarprodukt erhält man so das *Vektorprodukt*

$$\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 \quad \text{mit} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} := \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

Das Vektorprodukt hat viele Anwendungen in der dreidimensionalen Geometrie, Mechanik, Elektrodynamik, usw. Der natürliche allgemeine Begriff dahinter ist jedoch immer das äussere Produkt, welches eben in beliebiger Dimension existiert.

Grundeigenschaften: Nach Konstruktion ist das Vektorprodukt bilinear und alternierend. Weiter gilt für alle $u, v, w \in \mathbb{R}^3$:

(a) $\langle u, v \times w\rangle = \langle u \times v, w\rangle = \det(u, v, w)$.
(b) $u \times (v \times w) = \langle u, w\rangle \cdot v - \langle u, v\rangle \cdot w$ *(Grassmann-Identität).*
(c) $u \times (v \times w) + v \times (w \times u) + w \times (u \times v) = 0$ *(Jacobi-Identität).*
(d) $v \times w = 0$ genau dann, wenn v und w linear abhängig sind.

Geometrisch ist das Vektorprodukt charakterisiert durch die Eigenschaften:

(e) $v \times w$ ist orthogonal zu v und w.
(f) Sind v und w linear unabhängig, so ist $(v, w, v \times w)$ ein *Rechtssystem*, das heisst, es gilt $\det(v, w, v \times w) > 0$.
(g) $|v \times w|$ ist der Flächeninhalt des von v und w aufgespannten Parallelogramms, also $|v \times w| = |v| \cdot |w| \cdot |\sin \vartheta|$ wenn $\langle v, w\rangle = |v| \cdot |w| \cdot \cos \vartheta$.

11.8 Körpererweiterung

Sei K ein Unterkörper eines Körpers L, das heisst eine Teilmenge, welche mit den von L induzierten Rechenoperationen selbst einen Körper bildet. Dann ist jeder L-Vektorraum mit derselben Addition und der auf K eingeschränkten skalaren Multiplikation auch ein K-Vektorraum. Insbesondere wird L selbst zu einem K-Vektorraum.

Proposition: Für jeden K-Vektorraum V existiert genau eine Struktur als L-Vektorraum auf $V \otimes_K L$, deren additive Gruppe die von $V \otimes_K L$ ist und für die gilt:

$$\forall x, y \in L \; \forall v \in V : x \cdot (v \otimes y) = v \otimes xy.$$

Definition: Der L-Vektorraum $V_L := V \otimes_K L$ heisst die *Basiserweiterung von V bezüglich $K \subset L$.*

Proposition: Für jede Basis B des K-Vektorraums V bilden die Elemente $b \otimes 1$ von V_L für alle $b \in B$ eine Basis B_L des L-Vektorraums V_L. Insbesondere gilt

$$\dim_L(V_L) \;=\; \dim_K(V).$$

Beispiel: Für jedes $n \geqslant 0$ existiert ein natürlicher Isomorphismus von L-Vektorräumen

$$K^n \otimes_K L \xrightarrow{\sim} L^n \quad \text{mit} \quad v \otimes x \mapsto xv.$$

Definition: Im Fall $\mathbb{R} \subset \mathbb{C}$ heisst $V_{\mathbb{C}}$ die *Komplexifizierung* des reellen Vektorraums V.

Definition: Der *komplex Konjugierte* eines $\mathbb{C}$-Vektorraums $(W, +, \cdot\,, 0_W)$ ist der $\mathbb{C}$-Vektorraum $(W, +, \bar{\cdot}, 0_W)$, bei dem die skalare Multiplikation $\cdot$ ersetzt wurde durch

$$\bar{\cdot} : \; \mathbb{C} \times W \to W, \; (z, w) \mapsto z \bar{\cdot} w := \bar{z} \cdot w.$$

So wie wir üblicherweise $(W, +, \cdot\,, 0_W)$ mit W abkürzen, schreiben wir für $(W, +, \bar{\cdot}, 0_W)$ nur kurz $\overline{W}$.

Beispiel: Für jeden $\mathbb{C}$-Unterraum $W \subset \mathbb{C}^n$ existiert ein natürlicher Isomorphismus von $\mathbb{C}$-Vektorräumen

$$\overline{W} \xrightarrow{\sim} \{\overline{w} \mid w \in W\} \subset \mathbb{C}^n, \quad w \mapsto \overline{w}.$$

Bemerkung: Für jeden $\mathbb{C}$-Vektorraum W gilt $\overline{(\overline{W})} = W$.

Bemerkung: Jede Basis von W ist auch eine Basis von $\overline{W}$.

Proposition: (Vergleiche §9.11 und §10.7.) Für jeden endlich-dimensionalen unitären Vektorraum W existiert ein natürlicher Isomorphismus von $\mathbb{C}$-Vektorräumen

$$\delta : \overline{W} \xrightarrow{\sim} W^* := \mathrm{Hom}_{\mathbb{C}}(W, \mathbb{C}), \quad v \mapsto \delta(v) := \langle v, \; \rangle.$$

Proposition: Für jeden $\mathbb{C}$-Vektorraum W existiert ein natürlicher Isomorphismus von $\mathbb{C}$-Vektorräumen

$$W \otimes_{\mathbb{R}} \mathbb{C} \xrightarrow{\sim} W \boxplus \overline{W}, \; w \otimes z \mapsto (zw, z \bar{\cdot} w).$$