LINEARE ALGEBRA

GERTRUD DESCH

30. September 2014

Das griechische Alphabet.

alpha	A	α	ny	N	ν
beta	B	β	xi	Ξ	ξ
gamma	Γ	γ	omikron	O	o
delta	Δ	δ	pi	Π	π
epsilon	E	ϵ, ε	rho	P	ρ
zeta	Z	ζ	sigma	Σ	σ
eta	H	η	tau	T	τ
theta	Θ	θ, ϑ	ypsilon	Y	υ
iota	I	ι	phi	Φ	ϕ, φ
kappa	K	κ	chi	X	χ
lambda	Λ	λ	psi	Ψ	ψ
my	M	μ	omega	Ω	ω

gertrud.desch@uni-graz.at
Telefon (0316) 380 5177
Karl-Franzens-Universität Graz
Institut für Mathematik und Wissenschaftliches Rechnen
Heinrichstrasse 36, oberster = 4. Stock, Zimmer 505,
durch die Glastür gleich beim Lift, dann vorletzte Tür links.

Inhaltsverzeichnis

0. Einleitung 2
0.1. Mathematik verstehen 2
0.2. Der mathematische Lösungsprozess 4
0.3. Ein paar allgemeine Tipps für die Arbeit 6
1. Analytische Geometrie und lineare Gleichungssysteme 10
1.1. Analytische Geometrie in $\mathbb{R}^2$ und $\mathbb{R}^3$ 10
2. Gruppen, Ringe, Körper 23
2.1. Mengen und Funktionen 23
2.2. Gruppen 29
2.3. Ringe, Restklassen 36
2.4. Körper, komplexe Zahlen 40
3. Vektorräume 46
3.1. Definitionen und Beispiele, Unterräume 46
3.2. Lineare Hülle und Erzeugendensysteme 48
3.3. Erste Schritte in der Matrizenrechnung 53
3.4. Eindeutigkeit der Dimension 63
3.5. Summenräume und Dimensionsformel 66
3.6. Äußere direkte Summe 70
4. Lineare Abbildungen 71
4.1. Definitionen und Beispiele linearer Abbildungen 71
4.2. Lineare Abbildungen in $\mathbb{K}^n$, Matrizen 77
4.3. Inverse Matrix 85
4.4. Koordinaten 90
4.5. Faktorräume 97
5. Determinanten 102
5.1. Definition der Determinante 102
5.2. Multiplikationssatz für Determinanten 107
5.3. Existenz der Determinante 109
5.4. Entwicklungssatz von Laplace 113
6. Vektorräume mit innerem Produkt 118
6.1. Der euklidische Raum 118
6.2. Norm und inneres Produkt in abstrakten Vektorräumen 130
6.3. Orthogonale Projektion 139
6.4. Orthonormalsysteme 148
6.5. Dualraum, duale Abbildung, adjungierte Matrix 152
7. Eigenwerte 158
7.1. Polynome 158
7.2. Eigenvektoren und Eigenwerte 165
7.3. Diagonalisierung 172
7.4. Hauptraumzerlegung 178
7.5. Matrizen in Polynome eingesetzt 190
8. Normale Matrizen 194
8.1. Orthogonale und unitäre Matrizen 194
8.2. Normale Matrizen 197
8.3. Selbstadjungierte Matrizen und Definitheit 198
8.4. Quadriken 205

0. Einleitung

Willkommen in der Linearen Algebra, einem mathematischen Kerngebiet, das seine Finger in fast alle Teilgebiete der Mathematik ausstreckt. Mit dieser Vorlesung möchte ich Ihnen Wissen und Verständnis für Vektoren und lineare Gleichungssysteme vermitteln. Vor allem möchte ich Ihnen aber auch ein Bild mitgeben, wie man Mathematik betreibt, wie eine mathematische Theorie gebaut ist, wie man sie beweist, und wie man sie zu Papier bringt.

0.1. **Mathematik verstehen.** Mathematik kann und muss auf vielen verschiedenen Ebenen verstanden werden, und daraus folgt auch, wie man sie lernen muss: Mit einem wachen und flexiblen Geist, der an den Text aus vielen verschiedenen Blickwinkeln herangeht. Die schlechte Nachricht zuerst: Und mit viel, viel Arbeit und Mühe. Um eine halbe Seite Mathematik zu verstehen, kann man manchmal einen Tag lang sitzen. Bitte lassen Sie sich nicht entmutigen! Gerade dann, wenn man den Widerstand eines Faches spürt, findet ein intensiver Lernprozess statt.

0.1.1. *Mathematik als Werkzeug.* Sie kennen Mathematik wahrscheinlich vor allem als ein Kompendium von Standardregeln, die man auf Standardprobleme anwendet, um eine richtige Lösung zu erhalten. Mathematik muss für den Benutzer, die Benutzerin ein zuverlässiges Werkzeug sein. Zu lernen gibt es verschiedene Methoden, zum Beispiel, wie man ein Gleichungssystem löst, wie man Matrizen miteinander multipliziert, wie man eine Determinante ausrechnet und was sie bedeutet. Diese Dinge trainiert man durch wiederholtes Üben, nicht zuletzt in den Proseminaren zur Vorlesung.

0.1.2. *Fachsprache.* Mathematische Fachliteratur, auch dieses Skriptum, hat ihre eigene, sphinxenartig steinerne Sprache: Texte sind in Definitionen, Sätze und Beweise aufgebrochen. Die Sprache ist monoton, es wiederholen sich immer wieder die Standardphrasen, an die Sie sich gewöhnen werden. Alles ist auf Unmissverständlichkeit ausgerichtet. Die kleinen Worte (dann, dann und nur dann, es gibt, es gibt genau ein, für alle, ...) sind sehr wichtig und dürfen nicht überlesen oder beliebig gegen andere ausgetauscht werden. Diese Sprache lernen Sie gleichzeitig mit dem Nachvollziehen der Beweise, wenn Sie mathematische Texte, zum Beispiel dieses Skriptum, genau durcharbeiten und sich immer wieder fragen: Wie würde ich denselben Text formulieren? Warum verwendet die Autorin genau diese umständliche Formulierung? Hätte ich wirklich mit meiner Formulierung dasselbe gesagt, was im Skriptum steht, oder besteht ein feiner Unterschied? Enthält meine Formulierung alles, was ein Leser, eine Leserin wissen muss, wenn er/sie nicht bei mir rückfragen können? Versuchen Sie gleich von Anfang an, alles was Sie an mathematischen Inhalten ausdrücken wollen, korrekt auszudrücken. Jedes Zeichen, das Sie setzen, jedes kleine Wort muss seine Bedeutung haben und bewusst gesetzt sein. Vermeiden Sie langatmiges Herumreden. Schreiben Sie Ihre mathematischen Gedanken auf, dann werden Sie ganz von selbst unnötige Phrasen weglassen.

0.1.3. *Nachprüfbarkeit.* Wir wollen Sie in dieser Lehrveranstaltung davon überzeugen, dass Sie selbst nachprüfen können, ob unsere Theorien stimmen.[1] Bitte glauben Sie mir nichts außer den Definitionen. Wenn ich eine Behauptung aufstelle, schulde ich Ihnen einen Beweis. Das Wesentliche eines Beweises ist, dass er von Behauptungen ausgeht, die bereits auf sicherem Boden stehen, und in kleinen Schritten

[1] Schon deshalb, weil in diesem Skriptum immer wieder Tippfehler gefunden werden, und wenn ich einen korrigiert habe, wachsen zehn neue nach, wie die Köpfe der Hydra selig.

immer mehr Einsichten gewinnt, bis man sieht, dass die gewünschte Aussage unvermeidlich gelten muss. Wenn Sie Beweise verstehen wollen, brauchen Sie mehrere Durchgänge.

(1) Zunächst fragen Sie sich: Was soll bewiesen werden? Kenne ich die Definitionen der Objekte, von denen die Rede ist? (Bei Bedarf nachschlagen.) Welche Voraussetzungen werden gemacht?
(2) Dann ein sorgfältiges Durcharbeiten von vorne nach hinten. Fragen Sie sich bei jedem Schritt: Warum darf denn die Autorin das behaupten? Wissen wir das schon, und woher?
(3) Wenn Sie sich überzeugt haben, dass der Beweis korrekt ist, sollten Sie ihn nach folgenden Gesichtspunkten wiederholt durchlesen: Was sind die entscheidenden Ideen, die Wendepunkte im Beweis, und was ergibt sich einfach automatisch durch Nachrechnen.
(4) An welchen Stellen werden die Voraussetzungen gebraucht? Werden wirklich alle Voraussetzungen benötigt?

Lesen Sie die Definitionen sehr genau und lernen Sie sie auswendig, dass Sie nicht immer nachschlagen müssen, worüber wir reden. Merken Sie sich aber die Beweise nicht durch Auswendiglernen, sondern indem Sie sich die großen Linien des Gedankengangs, die grundlegenden Ideen, merken, und üben Sie, bei geschlossenem Skriptum den Beweis in eigenen Worten aufzufüllen und nachzuvollziehen. Nehmen Sie sich die Mühe und schreiben Sie Ihren Beweis exakt aus, oder erzählen Sie ihn einer/m KollegIn. In der Mathematik hat man nur das verstanden, was man auch exakt aussprechen kann. Fällt Ihnen ein anderer, kürzerer oder eleganterer Beweis ein? Hat die Verfasserin Fehler gemacht?[2] Arbeiten Sie bitte in kleinen Schritten, Beweis für Beweis. Es ist schwer, gerade am Anfang, Übersicht über ein ganzes Kapitel oder auch nur mehrere Seiten zu behalten.

0.1.4. *Theorie mit Höhepunkten.* Mathematik baut strukturierte Theorien. Es gibt zentrale Begriffe und Kernsätze, die Einsichten, um deretwillen die Theorie entwickelt wurde. Und es gibt viele Hilfssätze und Hilfsbegriffe, die helfen, sich zu den Beweisen der Kernsätze hochzuarbeiten. Wenn Sie einen Abschnitt durchgearbeitet haben, fragen Sie sich: Was sind die wesentlichen Begriffe und Aussagen des Kapitels, was ist die Pointe? Wie greifen die Aussagen des Kapitels ineinander? Wie baut das Kapitel auf vorigen Abschnitten auf?

0.1.5. *Abstraktion.* Mathematische Fachliteratur ist meist sehr abstrakt gehalten. Das heißt, es wird über die betrachteten Objekte so wenig wie möglich vorausgesetzt. Warum soll man für die reellen und komplexen Zahlen zweimal dasselbe beweisen, wenn man in Wirklichkeit nur ein paar Eigenschaften der Grundrechnungsarten braucht, die für beide Zahlenmengen gleich sind. Also reden wir von Körpern. Warum soll man eine Ebene oder den dreidimensionalen Raum getrennt betrachten, und gelten nicht auch dieselben Prinzipien im sechsdimensionalen Raum? Selbst Funktionen lassen sich als Vektoren auffassen, und auf diese Weise kann man Ideen der Geometrie auf Aufgaben übertragen, in denen die gesuchten Objekte Funktionen sind (zum Beispiel das Geschwindigkeitsfeld in einer strömenden Flüssigkeit). Auch wenn Sie zu Beginn eher den Eindruck haben werden, dass die Abstraktion eine Last ist, und sie Ihnen vielleicht sogar wie eine aufgeblasene akademische Seifenblase vorkommen mag: Wenn Sie im Studium fortschreiten, werden Sie immer mehr den Nutzen der Abstraktion schätzen lernen. Mit ihr kann man komplizierte Aufgaben behandeln, als wären sie einfach, weil sie einfach sind, wenn man sich auf das Wesentliche beschränkt und verwirrende Details weglassen kann. Für den

[2]Das kommt leider vor. Bitte um ein e-Mail, wenn Sie einen Fehler finden!

Anfang ist wichtig, dass Sie sich daran gewöhnen, zwischen den Abstraktionsebenen hin und her zu schalten. Wenn Sie einen Satz über einen allgemeinen Vektorraum lesen, fragen Sie sich: Was bedeutet dieser Satz in einer Ebene oder im dreidimensionalen Raum? Versuchen Sie sich, den Sachverhalt durch eine Zeichnung klarer zu machen. Verbirgt sich hinter dem mathematischen Formalismus vielleicht eine sehr einleuchtende Einsicht, die unsere Vorstellung von Raum und Ebene anspricht? Wenn Ihnen Ihre Intuition einen Sachverhalt nahelegt, der im dreidimensionalen Raum gilt, können Sie auch versuchen, ob Sie diesen Sachverhalt im allgemeinen Vektorraum beweisen können. Gilt er immer, oder brauchen Sie vielleicht Zusatzbedingungen, die nicht jeder Vektorraum erfüllt?

0.2. **Der mathematische Lösungsprozess.**

Ich bin manchmal gefragt worden, wie ich in einem so trockenen Beruf überleben kann. Die Antwort ist: Mathematik ist eine kreative Tätigkeit, weit entfernt von trocken. Mathematik ist für mich Wahrnehmen von Strukturen. Sehen ist keine maschinelle Tätigkeit, sondern ein Vorgang, der große Teile der Seele mit einbezieht, auch solche, die sich dem reinen Verstand entziehen.

Ich gehe jetzt bei meiner Beschreibung der schöpferischen Arbeit davon aus, dass Sie allein sind, und dass in der Ihnen zugänglichen Literatur die Lösung nicht zu finden ist, oder dass Sie darauf aus sind, das Beispiel allein zu lösen. Aber Sie wissen ja: Lesen und Diskussion mit KollegInnen machen Freude und sind unverzichtbare Quellen auf dem Weg der Wissenschaft.

Man darf sich den mathematischen Lösungsprozess nicht vorstellen, wie er leider oft in den Schulen eingepaukt wird, nämlich dass für eine oftmals geübte Musteraufgabe einfach das richtige Rechenverfahren eingeschaltet wird. Meist stehen wir ja vor Fragen, für die wir (zumindest selbst) noch kein Standardmuster zur Verfügung haben. Ich werde jetzt versuchen, den Lösungsprozess in Phasen aufzuteilen. Er läuft aber meistens nicht so linear ab. Immer wieder muss man von einer Phase in eine frühere zurückspringen. Einsicht ist ein Kreisprozess, eine Art Volleyballspiel zwischen den kreativen und den ordnenden Kräften im Kopf.

0.2.1. *Aufgabenstellung klären:* Oft ist dieser Abschnitt eine Routineaufgabe: Man überprüft, wie die Objekte definiert sind, und was die Aussage bedeutet, die man beweisen will. Manchmal hat man aber nur eine vage Vorstellung, dass ein bestimmter Sachverhalt gelten müsste, und muss selbst erst suchen, in welcher Form man ihn in einen mathematischen Satz gießen kann, wie man seine Objekte definieren muss, und welche Voraussetzungen passend sein könnten. In diesem Fall ist schon die Klärung der Aufgabenstellung ein Stück der Suchphase.

0.2.2. *Suchphase:* Das ist der eigentlich schöpferische Teil, für den es natürlich kein Patentrezept gibt. Wichtigste Spielregel ist, dass sich die Assoziationen frei entfalten dürfen. Alle Routinearbeiten werden für später aufgeschoben, jetzt sind nur die Aspekte interessant, die neue Ideen brauchen.

Man lässt alle Ideen an sich heran, und frägt sich dann: Hilft mir das? Hier sind einige Beispiele von Fragen, die man sich stellen kann, wenn man noch gar keine Vorstellung von einer möglichen Lösung hat:

- Was weiss ich über die Objekte, von denen die Rede ist? Welche Sätze kenne ich über sie?
- Käme ich mit zusätzlichen Voraussetzungen weiter, oder wenn ich in einer einfacheren Situation wäre, etwa in der Ebene statt im abstrakten Vektorraum?
- Kann ich eine Zeichnung machen? Was sagt mir meine Intuition?

- Kenne ich einen Satz, der ähnlich klingt wie der, den ich beweisen möchte? Wie wird er bewiesen? Wie unterscheidet er sich von meinem Problem?
- Kann ich mein Problem umformulieren, so dass es in eine mir bekannte Theorie fällt?
- Was muss noch gelten, wenn meine Behauptung wirklich stimmt? Was muss gelten, wenn sie nicht stimmt? Kann ich indirekt vorgehen?
- Stimmt meine Behauptung überhaupt? Gibt es ein Gegenbeispiel?
- Ich glaube, ich könnte die Aussage XXX beweisen. Würde mir die bei meinem Problem helfen?

Irgendwann fügen sich dann die vielen Gedanken zu einem Bild zusammen, in dem sich ein Weg abzeichnet. Versuchen Sie jetzt, diesen Weg zu gehen. Wo bleibt man hängen, welche Hürden müssen noch genommen werden? Am Ende schließt sich vielleicht ein zusammenhängender Weg von den Annahmen zur Aussage, oder vielleicht steht auch ein Gegenbeispiel da. Vielleicht kommt man auch zu keinem Ergebnis. Das ist in der Wissenschaft ganz normal. Wenn man alles beweisen könnte, was man will, wären Beweise ja sinnlos. Die vielen Irrwege sind ein wichtiger Teil der Erkenntnis.

Schmierpapier und Bleistifte sind die wertvollsten Ressourcen für schöpferische MathematikerInnen. Vielleicht müssen Sie sich, wie ich musste, erst daran gewöhnen, auf Blättern zu schreiben, auf denen schon viel durchgestrichen ist. Skizzieren Sie ihre Gedanken ruhig in allen Ecken und verbinden Sie sie mit Pfeilen quer über das Blatt. Radieren Sie nicht, streichen Sie einfach durch. Geniessen Sie es, auf einem Stück Papier ein Chaos zu inszenieren.

Wenn Sie sich festfahren und hängen bleiben, unterbrechen Sie und machen Sie eine Runde durch den Stadtpark oder sonst etwas Lockerndes, das Ihnen gut tut. In solchen Pausen tauchen oft schlagartig überraschende Einsichten auf. Schöpferische Arbeit ist ein Wechselspiel zwischen höchster Konzentration und Entspannung.

0.2.3. *Ordnungsphase:* Wenn die Lösung der Aufgabe gefunden ist, wird sie dokumentiert. Sie wird noch einmal systematisch ausformuliert. Ich habe (unpublizierte) Skizzen aus meiner mathematischen Sturm- und Drangzeit, die ich heute selbst nicht mehr verstehe, weil ich verabsäumt habe, sie nachvollziehbar hinzuschreiben. Oft merke ich auch beim sorgfältigen Hinschreiben, dass meine Gedanken noch Lücken haben, die zu schließen sind.

Um gut zu schreiben, müssen Sie nur einen Satz befolgen: Stellen Sie sich bitte immer, wenn Sie schreiben, einen Menschen vor, für den Sie schreiben.

Schreiben Sie alles, was diese Person zum Verständnis Ihrer Gedanken braucht. Verwirren Sie die LeserInnen nicht mit unnötigen Worten. Ordnen Sie den Text so, dass sich der Inhalt für die LeserInnen erschließt, ohne dass sie im Text ständig vor und rückwärts springen müssen.

Schreiben Sie einen Beweis möglichst so auf, dass Sie die Leser vom Bekannten zum Unbekannten führen. In der Suchphase geht es oft genau umgekehrt: “Ich möchte beweisen, dass XXX gilt, also muss ich zeigen, dass YYY gilt, also muss ich zeigen ...”. Das ist zwar korrekt, aber anstrengend zu lesen, weil jede noch nicht bewiesene Aussage im Kopf des Lesers sozusagen mit einem roten Warnschild “Weiss ich aber noch nicht” gespeichert wird, und ein Gefühl von Unsicherheit verursacht. Kennzeichnen Sie deutlich Behauptungen, die noch zu zeigen sind, und orientieren Sie die LeserInnen über Ihre Strategie. (z.B., “Wir zeigen zunächst, dass die Behauptung XXX in endlichdimensionalen Vektorräumen gilt. Der unendlich dimensionale Fall wird dann daraus gefolgert.” oder “Wir führen den Beweis indirekt. Sei also die Behauptung XXX falsch ...”) Wenn Sie für Ihren Beweis

Resultate brauchen, die schon bekannt sind, geben Sie Zitate, wo man diese Resultate mit Beweis nachschlagen kann. Ganz einfache Schritte können Sie auslassen, wenn Sie von der Leserschaft ausreichend Vorkenntnisse erwarten. (z.B., "Beweis als Übung.") Wenn Ihnen aber andere AutorInnen solche Stellen vorsetzen, überprüfen Sie bitte, ob der Beweis wirklich so leicht ist. Manchmal besteht die einfache Übung im Aufdecken eines fundamentalen Denkfehlers!

Am Ende der Ordnungsphase liegt das Ergebnis Ihrer Arbeit vor Ihnen, jederzeit nachvollziehbar für Sie und andere. In diesen schön aufgeschriebenen Resultaten sieht man oft nicht mehr die Mühe und auch nicht die Motivationen der Suchphase. Daher ist der Stil der mathematischen Literatur gut für die Dokumentation, als Lernende/r hätte man manchmal gern mehr Einblick auch in die Irrwege der schöpferischen Phase. Aber welche AutorInnen lassen sich schon gerne in die Karten gucken.

0.3. **Ein paar allgemeine Tipps für die Arbeit.**

0.3.1. *Offen sein.* Das ist vielleicht der beste und wichtigste Rat, den ich Ihnen mitgeben kann. Seien Sie offen und neugierig dem Studium gegenüber. Freuen Sie sich, dass Sie jetzt Mathematik betreiben können wie die "Großen", und dass Sie in eine viel weitere, vielfältigere Gedankenwelt eingetreten sind, als Sie es sich wahrscheinlich erträumt hätten.

Sie müssen und sollen nichts wegwerfen, was Sie aus der Schule mitgebracht haben. Richtige Mathematik bleibt richtig. Aber klammern Sie sich auch nicht ängstlich an den Methoden fest, die sie in der Schule gelernt haben, probieren Sie aus, was wir hier anders machen, und machen Sie sich selbst ein Bild, wie Sie gerne arbeiten. Lassen Sie sich ruhig auf die ungewohnte Form der Mathematik ein! Sie sind gut in der Schulmathematik, sonst hätten Sie ja nicht dieses Fach gewählt. Sie werden auch gut in der Hochschulmathematik sein. Sie können nur dazugewinnen, und was Sie schon können, kann Ihnen niemand nehmen.

0.3.2. *Vorlesung.* Vorweg, der Besuch der Vorlesung ist nicht verpflichtend, und auch wenn es für mich angenehm ist, wenn ich pünktlich anfangen kann, wird Sie niemand fragen oder gar böse ansehen, wenn Sie einmal erst später dazukommen oder früher gehen müssen.

Die Vorlesung sollte eigentlich Vordenkung heißen. Ich möchte vor Ihnen den Stoff entstehen lassen, mit der richtigen Sprache, der richtigen Argumentation, und — unabsichtlich — den gelegentlichen Fehlern, die mit der Entwicklung einer Theorie einhergehen. Sie sollen Gelegenheit haben, simultan mitzudenken, und sofort zu fragen, wenn Ihnen ein Gedankengang unklar ist. Trauen Sie sich! Fragen sind willkommen und zeigen, dass Sie bei der Sache sind. Ich erkläre gern einen Sachverhalt ein zweites Mal mit anderen Worten, und wenn es sein muss, auch öfter.

Das Tempo der Vorlesung können Sie mitbestimmen. Bitte bremsen Sie mich ruhig, wenn ich für Sie zu schnell vorgehe, und treiben Sie mich an, wenn es langweilig wird. Und weisen Sie mich auch hin, wenn meine Stimme zu leise oder meine Handschrift, sagen wir wertfrei und politically correct, kreativ wird. In einer Vorlesung sind Sie im Hörsaal keinerlei Prüfungen ausgesetzt. Wenn ich Fragen in die Hörerschaft richte, so nur, um Sie zum Mitdenken zu animieren, oder um ein Bild zu bekommen, wieviel Sie verstehen.

Niemand kann von Ihnen erwarten, über den Stoff in allen Details auf dem Laufenden zu sein, obwohl ein gewisses Maß an Mitlernen Ihnen später viel Arbeit spart und Sie von der Vorlesung mehr haben. Damit Sie verstehen können, was vorgetragen wird, sollten Sie die Definitionen geläufig haben, und eine Vorstellung

haben, um welche Kerninhalte es zur Zeit geht. Die Beweise müssen Sie nicht von Stunde zu Stunde parat haben.

Es gibt zwei hervorragende Möglichkeiten, mit denen Sie für sich den Wert der Vorlesung zunichte machen können. Selbstverständlich verbietet Ihnen niemand, gelegentlich ein Wort mit Ihren NachbarInnen zu tauschen, solange das nicht permanent geschieht, und so leise, dass nicht andere ZuhörerInnen gestört werden. Aber immer wenn Sie Ihre Aufmerksamkeit von der Mathematik abziehen, geht Ihnen der Fluss der Argumentation verloren. Übrigens müssen in einer Mathematikvorlesung die Tafeln sehr oft gelöscht werden, und die Pausen, die dabei entstehen, sind auch für Sie Entspannungs- und Gesprächspausen, und für mich eine Erholungsphase für die Stimme.

Die zweite Möglichkeit der Ressourcenverschwendung ist das kalligrafische Mitschreiben. Der ganze Stoff der Lehrveranstaltung liegt detailliert ausgearbeitet in diesem Skriptum vor Ihnen. Sie brauchen also als eigene Mitschrift nur einige wenige Notizen, die Sie vielleicht sogar in dieses Skriptum dazuschreiben können, wenn Sie es einseitig ausdrucken. Wenn Sie eine Mitschrift auf Schönschreiben, mit sauberen Unterstreichungen, eingerahmten Sätzen und mehreren Farben führen, verschwenden Sie alle Konzentration und Zeit auf die äußere Form. Was Ihnen aber die Vorlesung über das Skriptum hinaus bieten kann und soll, ist die Erfahrung und die Übung im Mitdenken.

0.3.3. *Proseminar.* Die Proseminare dienen dazu, dass Sie selbst versuchen, eigenständig Aufgaben zu lösen, die mit dem Stoff der Vorlesung Hand in Hand gehen. Darunter finden sich Routineaufgaben, in denen die Anwendung der gelernten Rechenverfahren geübt wird, und leichte Beweisaufgaben, damit Sie den Umgang mit den Begriffen und das mathematische Argumentieren üben. Immer wieder dazwischen kommt auch einmal eine schwierigere Aufgabe. Wenn Sie zu denen gehören, denen das Fach leicht fällt, werden Sie Spaß daran haben, wenn man Sie ein wenig herausfordert.

Es besteht immanenter Prüfungscharakter, das heißt, Sie müssen bei der Lehrveranstaltung anwesend sein, denn die Leistungen, die Sie im Hörsaal erbringen, tragen zumindest teilweise zur Gesamtnote bei. Es ist durchaus nicht unerwünscht, wenn Sie die Hausübungsbeispiele gemeinsam mit KollegInnen vorbereiten, aber diskutieren Sie sie in diesem Fall so weit aus, bis alle TeilnehmerInnen das Beispiel auch verstanden haben. Wenn Sie an die Tafel kommen und etwas vortragen, das Sie gar nicht verstehen, findet das die/der Lehrende schnell heraus, und Sie müssen mit Punkteabzug rechnen. Sonst gilt: Natürlich haben Sie das Recht auf Fehler. Es kommt einfach vor, dass man ein Beispiel fehlerhaft vorbereitet, leider geht mir das auch in der Vorlesung manchmal so. Wenn der Übungsleiter, die Leiterin, in Ihrer Lösung einen Fehler aufdeckt, wird er/sie vielleicht versuchen, Sie auf den richtigen Gedanken hinzulenken. Denken Sie mit und lassen Sie sich ruhig führen. Sie lernen dabei gleich die unschätzbare Fähigkeit, ein bisschen zu improvisieren.

Ich möchte Ihnen empfehlen, Proseminar und Vorlesung möglichst zugleich zu besuchen. In allen Fragen, welche die Proseminare betreffen, wenden Sie sich bitte an den Leiter / die Leiterin der Übungsgruppe, der Sie zugeteilt sind. Ich selbst bin für das Proseminar nicht zuständig, und habe keinen Einfluss auf die ÜbungsleiterInnen. Sie arbeiten autonom und eigenverantwortlich.

0.3.4. *Tutorium.* Es wird auch ein Tutorium angeboten. Dieses ist formlos und wird, damit für Sie die Schwellenangst möglichst gering ist, von Studierenden gehalten. Es soll Ihnen helfen, Schwierigkeiten zu überbrücken, die in Vorlesung oder Proseminar für Sie aufgetreten sind. Sie können mit Ihren Fragen zu Vorlesung und

Proseminar in das Tutorium kommen, oder auch einfach zuhören, was sonst im Tutorium geschieht.

0.3.5. *Fachbücher.* Diese Vorlesung folgt weitgehend dem Taschenbuch Lineare Algebra von Gerd Fischer, Vieweg Verlag. Es ist ein gut ausgereiftes und sorgfältig geschriebenes Lehrbuch, das inzwischen seine siebzehnte Auflage erlebt. Ich empfehle Ihnen auch, in der Bibliothek des Institutes für Mathematik an der KFU am Standort für Lineare Algebra zu schmökern. Vieles wird jetzt noch weit über Ihrem Horizont liegen, aber es gibt auch sehr gut erklärte Anfängerlehrbücher. Mir selbst gefallen zum Beispiel die Bücher von Klaus Jänich und von Gilbert Strang. Blättern Sie auch ruhig in englischsprachiger Fachliteratur. Die Schwierigkeit der Fremdsprache wird oft wett gemacht, indem englisch schreibende VerfasserInnen typischerweise mehr Wert auf Verständlichkeit und gute Erklärung legen.

0.3.6. *Fragen kommen.* Wenn Sie mit dem Verständnis anstehen, oder auch allgemeine Fragen oder Wünsche zur Vorlesung haben, scheuen Sie sich nicht, fangen Sie mich nach der Vorlesung ab, oder klopfen Sie bei mir an. Das gilt natürlich auch für Fragen, auf die Sie später beim Lernen auf die Prüfung stossen. Sie sind willkommen, ich gebe Ihnen gerne Auskunft. Das Schlimmste, was passieren kann, ist dass ich gerade nicht da bin oder keine Zeit habe. Wenn Sie mich ganz sicher erreichen wollen, machen Sie am besten einen Termin über e-Mail aus.

0.3.7. *Gemeinsam lernen.* Es ist ein guter Tipp, sich mit KollegInnen zusammen zu tun und gemeinsam zu lernen. Erstens kann man sich erzählen, was in den Lehrveranstaltungen geschehen ist, die jemand versäumt hat. Es ist auch gut, beim Lernen und Verstehen des Stoffes miteinander zu diskutieren. Vier oder sechs Augen finden schneller den springenden Punkt als zwei. Wenn Sie sich gegenseitig beim Lernen auf die Prüfung unterstützen, lassen Sie sich von der Kollegin oder dem Kollegen den Stoff genau und unmissverständlich erzählen, damit Unklarheiten aufgedeckt werden können, und auch die Fähigkeit zum exakten sprachlichen Ausdruck trainiert wird. Geben Sie bitte als Zuhörende/r nicht Ruhe, bis alles sauber und unmissverständlich ausformuliert ist. “Ich weiss eh schon, was Du meinst” hilft dem Lernpartner nicht.

0.3.8. *Prüfungen und Noten.* Ich werde, die Vorlesungen Lineare Algebra 1 und 2 schriftlich prüfen, um mit der TU einheitlich vorzugehen und auch wegen der großen Studierendenzahlen . Es wird für jede der beiden Vorlesungen 6 Termine geben, der erste davon gleich nach Ende der Lehrveranstaltungen, die anderen verstreut über ein Jahr. Mustertests, an denen Sie sehen können, wie meine Prüfungen aufgebaut sind, werden rechtzeitig vor den ersten Prüfungsterminen auf meiner Homepage angeboten werden. Bitte unterschätzen Sie nicht die Schwierigkeit der Prüfungen an der Hochschule. Normalerweise reichen nicht bloss wenige Tage aus, sondern man braucht mehrere Wochen, um den umfangreichen und schwierigen Stoff zu bewältigen.

Gar nicht wenige von Ihnen sind anfangs ziemlich enttäuscht, wenn Sie auf der Höheren Schule stets gewohnt waren, in Mathematik zu glänzen, und nun auf der Universität vielleicht nur im Mittelfeld abschneiden, oder sich überhaupt sehr schwer tun. Lassen Sie sich nicht entmutigen! Waren Sie in Ihrer Klasse eine/r der wenigen, für die ernsthaft ein Mathematikstudium in Frage gekommen ist, so finden Sie sich jetzt unter lauter KollegInnen, die auch Mathematik gewählt haben. Das ist ein ganz anderer Konkurrenzdruck. Es ist absolut ok, nicht der/die Beste zu sein.

Die ersten Schritte des Mathematikstudiums sind oft von einer Art Kulturschock begleitet. Wenn man dann zu begreifen beginnt, wo der Kern des mathematischen Denkens liegt, öffnet sich das spröde Fach. Ich wünsche Ihnen, dass Sie nicht nur immer tiefer in das Verständnis der Mathematik vordringen, sondern dabei auch etwas von der Begeisterung und Faszination erleben, mit der Mathematik für mich und viele verbunden ist. Viel Erfolg und Freude im Studium!

Ein herzliches Dankeschön an alle Studierenden und KollegInnen, die mir mit ihren Hinweisen geholfen haben, in diesem Skriptum Fehler zu verbessern und Erklärungen besser verständlich zu machen. Ich möchte mich ganz besonders bei den Herren Marko Golob, Florian Kapus, Alexander Perl, Manfred Scheucher und Stefan Steiner bedanken[3], die das gesamte Skriptum mit größter Sorgfalt durchgearbeitet haben, und mir jeweils detaillierte und sehr hilfreiche Listen von Korrekturen und Verbesserungsvorschlägen zusammengestellt haben.

Graz, 30. September 2014, Gertrud Desch

[3]Viele, viele weitere Tippfehlerhinweise sind bei mir angekommen. Wenn Ihr Name nicht mit aufgezählt ist, nehmen Sie bitte trotzdem meinen herzlichen Dank!

1. Analytische Geometrie und lineare Gleichungssysteme

1.1. Analytische Geometrie in $\mathbb{R}^2$ und $\mathbb{R}^3$.

1.1.1. *Gerade und Ebene in* $\mathbb{R}^2$ *und* $\mathbb{R}^3$.

Vereinbarung 1.1.1.

(1) Die Menge der natürlichen Zahlen ist $\mathbb{N} = \{1, 2, 3, \cdots\}$.[4]
(2) Die Menge der ganzen Zahlen ist $\mathbb{Z} = \{0, 1, -1, 2, -2, \cdots\}$.
(3) Die reellen Zahlen sind alle Zahlen der Zahlengeraden, also 0 sowie alle positiven und alle negativen Zahlen, die eine Länge bezeichnen können. Die Menge der reellen Zahlen schreiben wir $\mathbb{R}$.

Beachten Sie bitte, dass die obige Vereinbarung viel zu vage für eine mathematische Definition wäre. Wir gehen davon aus, dass Sie mit den Zahlen und den Grundrechnungsarten vertraut sind. Natürlich gibt es einen mathematisch exakten Aufbau des Zahlensystems. Wenn man für die natürlichen Zahlen einige grundlegende Axiome voraussetzt, kann man die anderen Zahlenmengen, die Grundrechnungsarten darin und ihre Eigenschaften streng formal konstruieren. Speziell die Konstruktion der reellen Zahlen, wenn man einmal die Bruchzahlen hat, ist ziemlich genial und trickreich. Das würde uns aber hier zu weit von unserem Thema abbringen.

Definition 1.1.2.

(1) Sei $n \in \mathbb{N}$. Mit $\mathbb{R}^n$ bezeichnen wir die Menge aller geordneten n-Tupel
$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$
mit $\alpha_i \in \mathbb{R}$ für $i = 1, \cdots, n$.
(2) Wir führen auf $\mathbb{R}^n$ Grundrechnungsarten ein: Die Addition und die Subtraktion, und die Multiplikation mit einer reellen Zahl:
$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \pm \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} := \begin{pmatrix} \alpha_1 \pm \beta_1 \\ \alpha_2 \pm \beta_2 \\ \vdots \\ \alpha_n \pm \beta_n \end{pmatrix} \qquad \gamma \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} := \begin{pmatrix} \gamma\alpha_1 \\ \gamma\alpha_2 \\ \vdots \\ \gamma\alpha_n \end{pmatrix}.$$
(3) Der Nullvektor ist der Spaltenvektor
$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$
(4) Die Elemente von $\mathbb{R}^n$ bezeichnen wir als n-dimensionale Spaltenvektoren. Zum Unterschied von Vektoren bezeichnen wir die einzelnen reellen Zahlen als Skalare. Die einzelnen Einträge eines Spaltenvektors (z.B. α_1 oder α_2 usw.) bezeichnen wir als die Komponenten des Vektors.

Schreibweise 1.1.3. Wir schreiben normalerweise $\mathbb{R}^n$ durch Spaltenvektoren. Dagegen würden wir
$$(\alpha_1, \alpha_2, \cdots, \alpha_n)$$
als Zeilenvektor bezeichnen. In der linearen Algebra macht es einen Unterschied, ob man Zeilen oder Spalten schreibt.

Vereinbarung 1.1.4.

[4] Achtung, manche Autoren zählen auch 0 zu den natürlichen Zahlen.

(1) Vektoren in $\mathbb{R}^2$ oder $\mathbb{R}^3$ veranschaulichen wir durch Pfeile, der Vektor $\begin{pmatrix}\alpha_1\\ \alpha_2\end{pmatrix} \in \mathbb{R}^2$ entspricht zum Beispiel einem Pfeil, der α_1 Einheiten von links nach rechts, und α_2 Einheiten von unten nach oben zeigt. Der Ansatzpunkt des Pfeiles ist gleichgültig, parallele, gleich lange Pfeile werden mit demselben Vektor identifiziert.
(2) Fixieren wir (willkürlich) einen Punkt des $\mathbb{R}^n$ als Nullpunkt 0 ("Koordinatenursprung"), so können wir jeden anderen Punkt P des $\mathbb{R}^n$ durch einen Vektor beschreiben, der von 0 nach P zeigt. Wir nennen diesen Vektor den Ortsvektor von P.
(3) Jeder Vektor $a \in \mathbb{R}^n$ läßt sich mit einer Parallelverschiebung identifizieren, welche jeden Punkt $x \in \mathbb{R}^n$ nach $x + a$ verschiebt.
(4) Die Summe $a + b$ kann man grafisch darstellen indem man den Vektor b an der Spitze von a ansetzt. Der Vektor vom Ausgangspunkt von a an die Spitze von b ist $a + b$. Wenn man $a + b$ auch konstruiert, indem man a an der Spitze von b ansetzt, erhält man ein Parallelogramm nach Art eines Kräfteparallelogramms.
(5) Das Produkt λa mit positivem λ ist ein Vektor in die gleiche Richtung wie a, aber mit der λ-fachen Länge. Das Produkt $-\lambda a$ hat genau entgegengesetzte Richtung zu a, und wieder λ-fache Länge wie a.

Definition 1.1.5. Seien a und $b \neq 0$ zwei Vektoren des $\mathbb{R}^n$. Die Gerade durch a in Richtung b ist definiert als

$$\{a + \lambda b \mid \lambda \in \mathbb{R}\}.$$

In Definition 1.1.5 war es wichtig, einen Richtungsvektor $b \neq 0$ zu haben. Für $b = 0$ würde das Gebilde auf den einen Punkt a zusammenfallen. Wie definiert man eine Ebene? Es reicht nicht, einfach 2 Richtungsvektoren zu nehmen, die nicht Null sind. Sie dürfen auch nicht Vielfache voneinander sein, sonst erhalten wir eine Gerade.

Definition 1.1.6. Seien $a, b_1, b_2 \in \mathbb{R}^n$, es sei weder b_1 ein Vielfaches von b_2, noch umgekehrt. Die Ebene durch a, deren Richtungen von b_1, b_2 aufgespannt werden, ist

$$\{a + \lambda_1 b_1 + \lambda_2 b_2 \mid \lambda_1, \lambda_2 \in \mathbb{R}\}.$$

Beispiel 1.1.7. Gegeben seien die 3 Punkte mit Ortsvektoren

$$x = \begin{pmatrix}2\\1\\4\end{pmatrix}, \quad y = \begin{pmatrix}0\\2\\6\end{pmatrix}, \quad z = \begin{pmatrix}4\\-3\\0\end{pmatrix}.$$

Legen Sie eine Ebene durch die drei Punkte.

Lösung: Wir können uns den Ausgangspunkt in der Ebene aussuchen, z.B. den Punkt x aus der Angabe, und als Richtungsvektoren kommen alle Vektoren in Frage, die zwei Punkte der Ebene verbinden, z.B. $y - x$ und $z - x$:

$$\{\begin{pmatrix}2\\1\\4\end{pmatrix} + \lambda_1 \begin{pmatrix}-2\\1\\2\end{pmatrix} + \lambda_2 \begin{pmatrix}2\\-4\\-4\end{pmatrix}\}.$$

Wir könnten genausogut einen anderen Punkt der Ebene als Ausgangspunkt nehmen und die Richtungsvektoren mit Konstanten multiplizieren, wir erhalten dieselbe Ebene:

$$\{\begin{pmatrix}0\\2\\6\end{pmatrix} + \lambda_1 \begin{pmatrix}-2\\1\\2\end{pmatrix} + \lambda_2 \begin{pmatrix}-1\\2\\2\end{pmatrix}\}.$$

□

1.1.2. *Geometrische Aufgaben, welche auf lineare Gleichungssysteme führen.*

Beispiel 1.1.8. Bestimmen Sie den Durchschnitt der Geraden

$$\begin{aligned} G &= \{\begin{pmatrix}-1\\-3\end{pmatrix} + \lambda \begin{pmatrix}2\\3\end{pmatrix} \mid \lambda \in \mathbb{R}\}, \\ H &= \{\begin{pmatrix}1\\-2\end{pmatrix} + \lambda \begin{pmatrix}1\\2\end{pmatrix} \mid \lambda \in \mathbb{R}\}. \end{aligned}$$

Lösung: Um nicht denselben Buchstaben λ in zwei verschiedenen Bedeutungen in der Rechnung zu haben, parametrisieren wir die Gerade H mit μ. Wie der Parameter heißt, ist schließlich gleichgültig. Wir suchen also einen Punkt x, der zugleich in zwei Mengen liegt:

$$x \in \{\begin{pmatrix}-1\\-3\end{pmatrix} + \lambda \begin{pmatrix}2\\3\end{pmatrix} \mid \lambda \in \mathbb{R}\} \cap \{\begin{pmatrix}1\\-2\end{pmatrix} + \mu \begin{pmatrix}1\\2\end{pmatrix} \mid \mu \in \mathbb{R}\}\,.$$

Der Vektor x muss sich also auf zwei verschiedene Arten schreiben lassen, und wir erhalten

$$\begin{aligned} x = \begin{pmatrix}-1\\-3\end{pmatrix} + \lambda \begin{pmatrix}2\\3\end{pmatrix} &= \begin{pmatrix}1\\-2\end{pmatrix} + \mu \begin{pmatrix}1\\2\end{pmatrix}, \\ \lambda \begin{pmatrix}2\\3\end{pmatrix} - \mu \begin{pmatrix}1\\2\end{pmatrix} &= -\begin{pmatrix}-1\\-3\end{pmatrix} + \begin{pmatrix}1\\-2\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix}. \end{aligned}$$

Schreiben wir die einzelnen Komponenten auf, erhalten wir zwei Gleichungen in zwei Unbekannten:

$$\begin{array}{rcrcl} 2\lambda & - & \mu & = & 2 \\ 3\lambda & - & 2\mu & = & 1 \end{array}.$$

Wir ziehen zweimal die erste Gleichung von der zweiten ab, mit der Absicht, aus der zweiten Gleichung die Unbekannte μ hinauszuwerfen. Die erste Gleichung behalten wir unverändert:

$$\begin{array}{rcrcl} 2\lambda & - & \mu & = & 2 \\ -\lambda & + & 0\mu & = & -3 \end{array}.$$

Nun addieren wir zweimal die zweite Gleichung zur ersten und eliminieren damit λ aus der ersten Zeile:

$$\begin{array}{rcrcl} 0\lambda & - & \mu & = & -4 \\ -\lambda & + & 0\mu & = & -3 \end{array}.$$

Nun können wir direkt ablesen:

$$\lambda = 3\,, \quad \mu = 4\,.$$

Wir setzen $\lambda = 3$ in die Gleichung von G ein und erhalten

$$x = \begin{pmatrix}-1\\-3\end{pmatrix} + 3 \begin{pmatrix}2\\3\end{pmatrix} = \begin{pmatrix}5\\6\end{pmatrix}.$$

Als Probe können wir auch μ in die Gleichung von H einsetzen und müssen denselben Vektor x als Schnittpunkt erhalten:

$$x = \begin{pmatrix}1\\-2\end{pmatrix} + 4 \begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}5\\6\end{pmatrix}.$$

Die Probe zeigt also, dass wir richtig gerechnet haben. Der gesuchte Schnittpunkt ist

$$x = \begin{pmatrix}5\\6\end{pmatrix}.$$

□

Bemerkung 1.1.9. Haben Sie Geduld mit dem Verfahren, mit dem wir oben das Gleichungssystem gelöst haben. Es sieht auf ersten Blick bei einem Gleichungssystem in 2 Unbekannten ein wenig klobig aus. Seine Qualität liegt darin, innerhalb der Gleichungen Ordnung zu behalten. Man kann damit ebensogut Gleichungssysteme in 10 Unbekannten lösen, ohne die Übersicht zu verlieren.

Beispiel 1.1.10. Gegeben seien die folgende Ebene E und die Punkte x, y. Liegen $x \in E$, $y \in E$?

$$x = \begin{pmatrix}5\\-4\\7\end{pmatrix}, \quad y = \begin{pmatrix}2\\5\\8\end{pmatrix}, \quad E = \{\begin{pmatrix}1\\1\\2\end{pmatrix} + \lambda \begin{pmatrix}2\\1\\3\end{pmatrix} + \mu \begin{pmatrix}1\\4\\2\end{pmatrix} \mid \lambda, \mu \in \mathbb{R}\}.$$

Lösung: x liegt genau dann in E, wenn es sich in der Form

$$x = \begin{pmatrix} 5 \\ -4 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}$$

schreiben lässt. Wenn wir diese Gleichung wieder in einzelne Komponenten auflösen, erhalten wir ein System von drei Gleichungen in zwei Unbekannten:

$$\begin{array}{rcrcr} 2\lambda & + & \mu & = & 4 \\ \lambda & + & 4\mu & = & -5 \\ 3\lambda & + & 2\mu & = & 5 \end{array}$$

Wir ziehen die zweite Zeile zweimal von der ersten und dreimal von der dritten ab, die zweite Zeile behalten wir:

$$\begin{array}{rcrcr} 0\lambda & - & 7\mu & = & 14 \\ \lambda & + & 4\mu & = & -5 \\ 0\lambda & - & 10\mu & = & 20 \end{array}$$

Wir dividieren die letzte Zeile durch 10, das ändert nichts an der Lösung:

$$\begin{array}{rcrcr} 0\lambda & - & 7\mu & = & 14 \\ \lambda & + & 4\mu & = & -5 \\ 0\lambda & - & \mu & = & 2 \end{array}$$

Nun ziehen wir die dritte Zeile 7 mal von der ersten ab. Die zweite und dritte behalten wir unverändert.

$$\begin{array}{rcrcr} 0\lambda & + & 0\mu & = & 0 \\ \lambda & + & 4\mu & = & -5 \\ 0\lambda & - & \mu & = & 2 \end{array}$$

Die erste Zeile sagt: $0\lambda + 0\mu = 0$. Das gilt immer, daher macht es keinen Unterschied, ob diese Zeile im Gleichungssystem steht oder nicht. Wir können sie weglassen. Es bleibt

$$\begin{array}{rcrcr} \lambda & + & 4\mu & = & -5 \\ 0\lambda & - & \mu & = & 2 \end{array}$$

Wir addieren noch die letzte Zeile viermal zur ersten und erhalten

$$\begin{array}{rcrcr} \lambda & + & 0\mu & = & 3 \\ 0\lambda & - & \mu & = & 2 \end{array}$$

Also ist $\lambda = 3$ und $\mu = -2$. Als Probe setzen wir λ und μ in die Gleichung der Ebene ein und erhalten tatsächlich x:

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \\ 7 \end{pmatrix} .$$

Der Vektor x liegt also in der Ebene E.

Nun versuchen wir, ob auch y in E liegt, und setzen an:

$$y = \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} .$$

Wenn wir diese Gleichung wieder in einzelne Komponenten auflösen, erhalten wir ein System von drei Gleichungen in zwei Unbekannten:

$$\begin{array}{rcrcr} 2\lambda & + & \mu & = & 1 \\ \lambda & + & 4\mu & = & 4 \\ 3\lambda & + & 2\mu & = & 6 \end{array}$$

Wir ziehen die zweite Zeile wieder zweimal von der ersten und dreimal von der dritten ab, die zweite Zeile behalten wir:

$$\begin{array}{rcrcr} 0\lambda & - & 7\mu & = & -7 \\ \lambda & + & 4\mu & = & 4 \\ 0\lambda & - & 10\mu & = & -6 \end{array}$$

Wir dividieren die erste Zeile durch 7, das ändert nichts an der Lösung:

$$\begin{array}{rcrcr} 0\lambda & - & \mu & = & -1 \\ \lambda & + & 4\mu & = & 4 \\ 0\lambda & - & 10\mu & = & -6 \end{array}$$

Nun ziehen wir die erste Zeile 10 mal von der dritten ab:

$$\begin{array}{rcrcr} 0\lambda & - & 0\mu & = & -1 \\ \lambda & + & 4\mu & = & 4 \\ 0\lambda & - & 0\mu & = & 4 \end{array}$$

Die dritte Zeile sagt nun: $0\lambda + 0\mu = 4$. Das kann nie erfüllt sein, und daher hat das Gleichungssystem keine Lösung. Der Vektor y läßt sich also nicht in der Parameterform von E schreiben, er liegt nicht in E. □

Bemerkung 1.1.11. Im obigen Beispiel haben wir gesehen, was geschieht, wenn in der Elimination links eine Zeile aus lauter Nullen entsteht. Steht auf der rechten Seite auch Null, dann kann man die Zeile ignorieren und einfach mit den anderen Zeilen weiterrechnen. Steht aber auf der rechten Seite eine Zahl ungleich Null, so gibt es keine Lösung, und damit auch keinen Grund, weiterzurechnen.

Beispiel 1.1.12. Bestimmen Sie den Durchschnitt der Ebene E mit der Geraden G:

$$\begin{aligned} E &= \{\begin{pmatrix} 7 \\ 14 \\ -11 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 8 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R}\}, \\ G &= \{\begin{pmatrix} 11 \\ 2 \\ -19 \end{pmatrix} + \lambda \begin{pmatrix} 5 \\ 0 \\ -10 \end{pmatrix} \mid \lambda \in \mathbb{R}\}. \end{aligned}$$

Lösung: Wir suchen wieder einen Punkt x, der beide Parameterformen erfüllt, also

$$x = \begin{pmatrix} 7 \\ 14 \\ -11 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 8 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix} = \begin{pmatrix} 11 \\ 2 \\ -19 \end{pmatrix} + \nu \begin{pmatrix} 5 \\ 0 \\ -10 \end{pmatrix}.$$

(Beachten Sie, dass wir wieder den Parameter aus G umbenannt haben, um nicht einen Buchstaben für zwei verschiedene Zahlen zu verwenden.) Wir sortieren die Terme und erhalten

$$\lambda \begin{pmatrix} 2 \\ 8 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix} - \nu \begin{pmatrix} 5 \\ 0 \\ -10 \end{pmatrix} = \begin{pmatrix} 11 \\ 2 \\ -19 \end{pmatrix} - \begin{pmatrix} 7 \\ 14 \\ -11 \end{pmatrix} = \begin{pmatrix} 4 \\ -12 \\ -8 \end{pmatrix}.$$

Wir werden der Bequemlichkeit wegen in den einzelnen Gleichungen die Unbekannten weglassen. Statt

$$\begin{array}{rcrcrcr} 2\lambda & + & 3\mu & - & 5\nu & = & 4 \\ 8\lambda & + & 9\mu & - & 0\nu & = & -12 \\ 4\lambda & + & 0\mu & + & 10\nu & = & -8 \end{array}$$

schreiben wir nur ein Schema, das die Zahlenwerte enthält. Nach wie vor bedeutet aber jede Zeile eine Gleichung, und jede Spalte links gehört zu einer Unbekannten:

$$\begin{array}{rrrr} 2 & 3 & -5* & 4 \\ 8 & 9 & 0 & -12 \\ 4 & 0 & 10 & -8 \end{array}$$

Wir beginnen nun mit der Elimination. Zunächst werfen wir ν aus allen Gleichungen ausser der ersten, d.h., wir behalten die erste Zeile unverändert, die zweite müssen wir diesmal nicht verändern, und zur dritten zählen wir die erste zweimal dazu. Die erste Zeile, mit der wir arbeiten, heißt unsere Pivotzeile, die dritte Spalte, in der wir die Nullen erzeugen, heißt die Pivotspalte, und das Element -5 ist das Pivotelement. Wir markieren die Pivotelemente mit Sternchen. Wir führen nun den Pivotschritt durch:

$$\begin{array}{rrrr} 2 & 3 & -5* & 4 \\ 8* & 9 & 0 & -12 \\ 8 & 6 & 0 & 0 \end{array}$$

Die erste Zeile, die bereits einmal Pivotzeile war, lassen wir bis auf weiteres in Ruhe. Wir ziehen die zweite Zeile von der dritten ab, um λ aus der dritten Zeile zu eliminieren. Nun ist die zweite Zeile Pivotzeile, die erste Spalte Pivotspalte, und das Pivotelement ist 8.

$$\begin{array}{rrrr} 2 & 3 & -5* & 4 \\ 8* & 9 & 0 & -12 \\ 0 & -3* & 0 & 12 \end{array}$$

Das Element -3 in der letzten Zeile wird nun Pivotelement, dazu gibt es aber keine Zeilenumformung mehr, denn alle Zeilen, die man umformen könnte, wurden schon als Pivotzeilen verwendet. Damit ist die Eliminationsphase fertig.

Wir beginnen mit der Rücksubstitutionsphase. Wir fangen nun mit dem letzten Pivotelement an (dritte Zeile, zweite Spalte) und eliminieren alle anderen Zahlen aus dieser Spalte. Dazu addieren wir die dritte Zeile einmal zur ersten und dreimal zur zweiten Zeile:

$$\begin{array}{cccc} 2 & 0 & -5* & 16 \\ 8* & 0 & 0 & 24 \\ 0 & -3* & 0 & 12 \end{array}$$

Die vorletzte Pivotzeile war die zweite, mit dem Pivotelement 8. In der dritten Zeile gibt es nichts zu eliminieren. Von der ersten Zeile ziehen wir ein Viertel der zweiten Zeile ab.

$$\begin{array}{cccc} 0 & 0 & -5* & 10 \\ 8* & 0 & 0 & 24 \\ 0 & -3* & 0 & 12 \end{array}$$

Nun steht die Lösung vor uns:

$$-5\nu = 10\,, \quad 8\lambda = 24\,, -3\mu = 12\,,$$

das heißt

$$\lambda = 3\,, \quad \mu = -4\,, \quad \nu = -2\,.$$

Setzen wir $\lambda = 3$, $\mu = -4$ in die Ebenengleichung ein, so erhalten wir den Schnittpunkt

$$x = \begin{pmatrix} 7 \\ 14 \\ -11 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 8 \\ 4 \end{pmatrix} - 4 \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\,.$$

Wir können, als Probe, ebenso $\nu = -2$ in die Geradengleichung einsetzen, und müssen denselben Schnittpunkt erhalten:

$$x = \begin{pmatrix} 11 \\ 2 \\ -19 \end{pmatrix} - 2 \begin{pmatrix} 5 \\ 0 \\ -10 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}\,.$$

□

Beispiel 1.1.13. Bestimmen Sie den Durchschnitt der Ebenen

$$\begin{aligned} E &= \{ \begin{pmatrix} 5 \\ -1 \\ 9 \end{pmatrix} + \lambda \begin{pmatrix} 4 \\ -4 \\ 5 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \}, \\ F &= \{ \begin{pmatrix} 7 \\ 7 \\ 6 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \}. \end{aligned}$$

Lösung: Wir suchen wieder Punkte, die zugleich in beiden Parameterformen geschrieben werden können:

$$\begin{pmatrix} 5 \\ -1 \\ 9 \end{pmatrix} + \lambda \begin{pmatrix} 4 \\ -4 \\ 5 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \\ 6 \end{pmatrix} + \nu \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} + \kappa \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}\,,$$

$$\lambda \begin{pmatrix} 4 \\ -4 \\ 5 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} - \nu \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} - \kappa \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \\ 6 \end{pmatrix} - \begin{pmatrix} 5 \\ -1 \\ 9 \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ -3 \end{pmatrix}\,.$$

Damit erhalten wir das folgende Schema, wobei jede Zeile wieder für eine Gleichung steht, und die Spalten nacheinander die Koeffizienten für $\lambda, \mu, \nu, \kappa$ und die rechte Seite beinhalten:

$$\begin{array}{ccccc} 4 & 4 & -2 & -3 & 2 \\ -4 & -2 & -2 & -1* & 8 \\ 5 & 3 & -3 & -2 & -3 \end{array}$$

Wir führen nun Pivotschritte durch, bis keine Zeile als Pivotzeile mehr übrig bleibt. Die Pivotelemente werden jeweils durch Sternchen gekennzeichnet:

$$\begin{array}{ccccc} 16 & 10 & 4 & 0 & -22 \\ -4 & -2 & -2 & -1* & 8 \\ 13 & 7 & 1* & 0 & -19 \end{array}$$

$$\begin{array}{ccccc} -36 & -18 & 0 & 0 & 54 \\ -4 & -2 & -2 & -1* & 8 \\ 13 & 7 & 1* & 0 & -19 \end{array}$$

Wir dividieren die erste Zeile durch 18, und erhalten als letztes Pivotelement z.B. -1.

$$\begin{array}{ccccc} -2 & -1* & 0 & 0 & 3 \\ -4 & -2 & -2 & -1* & 8 \\ 13 & 7 & 1* & 0 & -19 \end{array}$$

Damit ist die Elimination abgeschlossen, alle Zeilen waren Pivotzeilen. Wir beginnen mit der Rücksubstitution, zuerst mit dem letzten Pivotelement:

$$\begin{array}{ccccc} -2 & -1* & 0 & 0 & 3 \\ 0 & 0 & -2 & -1* & 2 \\ -1 & 0 & 1* & 0 & 2 \end{array}$$

Die vorletzte Pivotzeile war Zeile Nummer 3. Wir setzen mit ihr die Rücksubstitution fort:

$$\begin{array}{ccccc} -2 & -1* & 0 & 0 & 3 \\ -2 & 0 & 0 & -1* & 6 \\ -1 & 0 & 1* & 0 & 2 \end{array}$$

Nun ist auch die Rücksubstitution fertig. Trotzdem können wir keine eindeutige Lösung ablesen, denn die erste Spalte war nie Pivotspalte. Unsere Gleichungen bedeuten:

$$\begin{array}{rcrcrcl} -2\lambda & - & \mu & & & = & 3 \\ -2\lambda & & & - & \kappa & = & 6 \\ -\lambda & + & \nu & & & = & 2 \end{array}$$

Für jeden Wert von λ können wir Werte von μ, ν, κ berechnen, sodass sich eine Lösung ergibt:

$$\mu = -3 - 2\lambda\,, \quad \nu = 2 + \lambda\,, \quad \kappa = -6 - 2\lambda\,.$$

Wir setzen in die erste Ebenengleichung ein und erhalten

$$\begin{aligned} E \cap F &= \left\{ \begin{pmatrix} 5 \\ -1 \\ 9 \end{pmatrix} + \lambda \begin{pmatrix} 4 \\ -4 \\ 5 \end{pmatrix} + (-3-2\lambda) \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} 5 \\ -1 \\ 9 \end{pmatrix} - 3 \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 4 \\ -4 \\ 5 \end{pmatrix} - 2\lambda \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} -7 \\ 5 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -4 \\ 0 \\ -1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}. \end{aligned}$$

Das ist die Parameterform einer Geraden. Die beiden Ebenen schneiden einander also in einer Geraden. □

1.1.3. *Gaußsches Eliminationsverfahren.*

Wir wollen nun ein bisschen systematischer aufarbeiten, was wir vorhin beobachtet haben. Zunächst befassen wir uns mit linearen Gleichungssystemen. Die geometrischen Aspekte kommen im nächsten Unterabschnitt.

Definition 1.1.14.

(1) Ein lineares Gleichungssystem von m Gleichungen in n Unbekannten hat die Gestalt

$$\begin{array}{ccccccccc} \alpha_{1,1}\xi_1 & + & \alpha_{1,2}\xi_2 & + & \cdots & + & \alpha_{1,n}\xi_n & = & \beta_1 \\ \alpha_{2,1}\xi_1 & + & \alpha_{2,2}\xi_2 & + & \cdots & + & \alpha_{2,n}\xi_n & = & \beta_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ \alpha_{m,1}\xi_1 & + & \alpha_{m,2}\xi_2 & + & \cdots & + & \alpha_{m,n}\xi_n & = & \beta_m \end{array} \tag{1.1.1}$$

Dabei sind für $i = 1, \cdots, m$ und $j = 1, \cdots, n$ die $\alpha_{i,j}$ und β_i gegebene reelle Zahlen[5]. Gesucht sind die reellen Zahlen $\xi_1, \cdots, \xi_n$.

[5]Später werden wir sehen, dass es genausogut Elemente eines beliebigen Körpers sein können.

(2) Eine partikuläre Lösung des Systems (1.1.1) ist ein Vektor

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix},$$

dessen Komponenten alle Gleichungen des Systems (1.1.1) gleichzeitig erfüllen.
(3) Die allgemeine Lösung oder Lösungsmenge des Systems (1.1.1) ist die Menge aller partikulären Lösungen.
(4) Wenn alle β_i gleich Null sind, heißt das System ein homogenes Gleichungssystem, sonst heißt es inhomogen.

Bemerkung 1.1.15. System (1.1.1) lässt sich kurz auch so anschreiben:

$$(\forall i = 1, \cdots, m) \ \sum_{j=1}^{n} \alpha_{i,j}\xi_j = \beta_i \,.$$

Beobachtung 1.1.16. Ein homogenes lineares Gleichungsystem hat immer mindestens eine partikuläre Lösung, den Nullvektor.

Beweis. Sind alle $\xi_j = 0$, so ist für alle $i = 1, \cdots, m$

$$\sum_{j=1}^{n} \alpha_{i,j}\xi_j = \sum_{j=1}^{n} \alpha_{i,j}0 = 0 \,.$$

□

Definition 1.1.17. Eine $(m \times n)$-Matrix über $\mathbb{R}$ ist ein rechteckiges Schema, in dem in m Zeilen und n Spalten reelle Zahlen angeordnet sind:

$$(\alpha_{i,j})_{i=1\cdots m, j=1\cdots n} \ = \ \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}.$$

Die Menge aller $(m \times n)$-Matrizen über $\mathbb{R}$ bezeichnen wir mit $\mathbb{R}^{m\times n}$.

Definition 1.1.18. (1) Die Systemmatrix (Koeffizientenmatrix) des linearen Gleichungssystems (1.1.1) ist die Matrix

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}.$$

(2) Die rechte Seite des linearen Gleichungssystems (1.1.1) ist der m-dimensionale Spaltenvektor

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}.$$

(3) Die erweiterte Matrix des linearen Gleichungssystems (1.1.1) ist die Matrix

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} & \beta_m \end{pmatrix}.$$

Wir fassen nun zusammen, wie wir im vorigen Unterabschnitt lineare Gleichungssysteme gelöst haben. Ein Algorithmus ist ein Rechenverfahren, das so präzis formuliert ist, dass man es nur rein mechanisch nachvollziehen muss.

Algorithmus 1.1.19 (Gaußsche Elimination)**.** Zu lösen ist ein System von m linearen Gleichungen in n Unbekannten. Wir gehen von der erweiterten Matrix des Systems aus.

Eliminationsphase:

1) Gibt es eine Zeile i in der erweiterten Matrix, die noch nicht Pivotzeile war, und auf der linken Seite ein Element $\alpha_{i,j} \neq 0$ enthält? Wenn ja, wähle dieses Element als Pivotelement. (Zeile i heißt dann Pivotzeile, Spalte j heißt Pivotspalte.) Wenn nein, dann gehe zu Schritt 4.
2) Dividiere die Pivotzeile durch das Pivotelement. Von jeder anderen Zeile k, welche noch nicht Pivotzeile war, subtrahiere das $\alpha_{k,j}/\alpha_{i,j}$-Fache der Pivotzeile. Es entsteht eine Null in der Pivotspalte.
3) Gehe zu Schritt 1.
4) Alle Zeilen, die nur aus Nullen bestehen, werden gestrichen.
5) Gibt es eine Zeile, die links nur Nullen besitzt, aber rechts ein Element ungleich Null enthält? Wenn ja, dann hat das System überhaupt keine Lösung. Beende die Rechnung. Wenn nein, gehe weiter zu Schritt 6.

Rücksubstitutionsphase:

6) Gibt es noch Zeilen, welche nicht für die Rücksubstition verwendet wurden? Wenn ja, wähle aus ihnen jene Zeile i, die als letzte Pivotzeile war und gehe zu Schritt 7. Wenn nein, gehe zu Schritt 10.
7) Wähle das ehemalige Pivotelement $\alpha_{i,j} = 1$ wieder als Pivotelement.
8) Von alle Zeilen k, die vor der Zeile Nummer i Pivotzeilen waren, subtrahiere das $\alpha_{k,j}/\alpha_{i,j}$-Fache der Pivotzeile.
9) Gehe zu Schritt 6.

Freie Parameter und Auflösung:

10) Gibt es Spalten, die nie Pivotspalten waren? Setze die entsprechenden Variablen ξ_j mit allgemeinen Zahlen (=Buchstaben) als freie Parameter an.
11) Aus jeder Zeile i, deren Pivotelement $\alpha_{i,j}$ war, läßt sich der Wert von ξ_j jetzt ablesen. Ende des Verfahrens.

Bemerkung 1.1.20. Von diesem Verfahren gibt es verschiedene Varianten. Eine ist, die Rücksubstitutionsphase nicht separat zu führen, sondern gleich in der Eliminationsphase auch in den Zeilen, die schon Pivotzeilen waren, zu eliminieren. Sie werden aber sehen, dass in den Pivotschritten der Rücksubstitutionsphase so viele Nullen auftreten, dass die Rechnung sehr kurz wird und das Verfahren mit getrennten Phasen etwas bequemer ist.

Beispiel 1.1.21.

$$\begin{array}{rcrcrcrcr} 2x & + & y & + & 3z & + & u & = & 7 \\ 6x & - & 2y & - & 11z & + & 3u & = & 1 \\ -6x & & & + & 6z & - & 2u & = & 4 \\ 2x & + & 3y & + & 17z & + & 4u & = & 45 \end{array}$$

Lösung: Zunächst stellen wir die erweiterte Matrix auf. Beachten Sie, dass in der dritten Gleichung y nicht vorkommt. Wir müssen aber an der Stelle für y in der Matrix eine Null schreiben, sodass alle Zeilen gleich lang sind.

$$\begin{pmatrix} 2 & 1 & 3 & 1 & 7 \\ 6 & -2 & -11 & 3 & 1 \\ -6 & 0 & 6 & -2 & 4 \\ 2 & 3 & 17 & 4 & 45 \end{pmatrix}$$

Wir formen nun die erweiterte Matrix mit Eliminationsschritten um. Pivotelemente kennzeichnen wir jeweils mit Stern $*$, die Nummer über dem Stern sagt, das wievielte Pivotelement hier steht.

$$\begin{pmatrix} 2 & 1*^{1} & 3 & 1 & 7 \\ 6 & -2 & -11 & 3 & 1 \\ -6 & 0 & 6 & -2 & 4 \\ 2 & 3 & 17 & 4 & 45 \end{pmatrix} \to \begin{pmatrix} 2 & 1*^{1} & 3 & 1 & 7 \\ 10 & 0 & -5 & 5*^{2} & 15 \\ -6 & 0 & 6 & -2 & 4 \\ -4 & 0 & 8 & 1 & 24 \end{pmatrix}$$

$$\to \begin{pmatrix} 2 & 1*^{1} & 3 & 1 & 7 \\ 2 & 0 & -1 & 1*^{2} & 3 \\ -2*^{3} & 0 & 4 & 0 & 10 \\ -6 & 0 & 9 & 0 & 21 \end{pmatrix} \to \begin{pmatrix} 2 & 1*^{1} & 3 & 1 & 7 \\ 2 & 0 & -1 & 1*^{2} & 3 \\ 1*^{3} & 0 & -2 & 0 & -5 \\ 0 & 0 & -3*^{4} & 0 & -9 \end{pmatrix}$$

$$\to \begin{pmatrix} 2 & 1*^{1} & 3 & 1 & 7 \\ 2 & 0 & -1 & 1*^{2} & 3 \\ 1*^{3} & 0 & -2 & 0 & -5 \\ 0 & 0 & 1*^{4} & 0 & 3 \end{pmatrix}$$

Für das letzte Pivotelement bestand der Pivotschritt nur in der Division der Pivotzeile durch das Pivotelement.

Mit dem vierten Pivotelement leiten wir nun die Rücksubstitutionsphase ein:

$$\begin{pmatrix} 2 & 1*^{1} & 3 & 1 & 7 \\ 2 & 0 & -1 & 1*^{2} & 3 \\ 1*^{3} & 0 & -2 & 0 & -5 \\ 0 & 0 & 1*^{4} & 0 & 3 \end{pmatrix} \to \begin{pmatrix} 2 & 1*^{1} & 0 & 1 & -2 \\ 2 & 0 & 0 & 1*^{2} & 6 \\ 1*^{3} & 0 & 0 & 0 & 1 \\ 0 & 0 & 1*^{4} & 0 & 3 \end{pmatrix}$$

$$\to \begin{pmatrix} 0 & 1*^{1} & 0 & 1 & -4 \\ 0 & 0 & 0 & 1*^{2} & 4 \\ 1*^{3} & 0 & 0 & 0 & 1 \\ 0 & 0 & 1*^{4} & 0 & 3 \end{pmatrix} \to \begin{pmatrix} 0 & 1*^{1} & 0 & 1 & -8 \\ 0 & 0 & 0 & 1*^{2} & 4 \\ 1*^{3} & 0 & 0 & 0 & 1 \\ 0 & 0 & 1*^{4} & 0 & 3 \end{pmatrix}$$

Nun können wir die Lösung direkt ablesen:

$$x = 1,\ y = -8,\ z = 3, u = 4\,.$$

□

Bemerkung 1.1.22. Wir haben zwischen die Matrizen in den Pivotschritten Pfeile und bewusst keine Gleichheitszeichen geschrieben, denn die Matrizen gehen auseinander hervor, sie sind aber nicht gleich.

Beispiel 1.1.23. Lösen Sie das System mit der erweiterten Matrix:

$$\begin{pmatrix} 10 & 2 & 14 & 4 & 6 \\ 1 & 3 & 19 & 0 & 5 \\ 11 & -2 & -11 & 5 & 0 \\ 14 & 0 & 2 & 6 & 4 \end{pmatrix}$$

Lösung: Wir beginnen mit den Eliminationsschritten.

$$\begin{pmatrix} 10 & 2*^{1} & 14 & 4 & 6 \\ 1 & 3 & 19 & 0 & 5 \\ 11 & -2 & -11 & 5 & 0 \\ 14 & 0 & 2 & 6 & 4 \end{pmatrix} \to \begin{pmatrix} 5 & 1*^{1} & 7 & 2 & 3 \\ -14 & 0 & -2 & -6 & -4 \\ 21 & 0 & 3 & 9 & 6 \\ 14 & 0 & 2*^{2} & 6 & 4 \end{pmatrix}$$

$$\to \begin{pmatrix} 5 & 1*^{1} & 7 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 1*^{2} & 3 & 2 \end{pmatrix}$$

Die beiden Nullzeilen ignorieren wir in Zukunft. Es bleibt ein Rücksubstitutionsschritt:

$$\begin{pmatrix} 5 & 1*^{1} & 7 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 1*^{2} & 3 & 2 \end{pmatrix} \to \begin{pmatrix} -44 & 1*^{1} & 0 & -19 & -11 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 1*^{2} & 3 & 2 \end{pmatrix}$$

Die zweite und dritte Spalte waren Pivotspalten, die erste und vierte Variable werden frei angesetzt:

$$\xi_1 = \lambda,\ \xi_4 = \mu\,.$$

Aus der ersten Zeile erhalten wir nun ξ_2, und aus der vierten Zeile erhalten wir ξ_3:

$$\begin{aligned} \xi_2 &= -11 + 44\lambda + 19\mu\,, \\ \xi_2 &= 2 - 7\lambda - 3\mu\,. \end{aligned}$$

Fassen wir die allgemeine Lösung zu einem Vektor zusammen, so erhalten wir als Lösungsmenge

$$\left\{ \begin{pmatrix} 0 \\ -11 \\ 2 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 44 \\ -7 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 19 \\ -3 \\ 1 \end{pmatrix} \;\Big|\; \lambda, \mu \in \mathbb{R} \right\}.$$

□

Beobachtung 1.1.24. Wir hatten bei der Lösung des letzten Systems 4 Gleichungen in 4 Unbekannten. Zwei Gleichungen wurden zu Nullzeilen, es sind zwei Pivotzeilen geblieben. Wir werden später sagen, der Rang der Systemmatrix ist zwei. Zu zwei Pivotzeilen gehören natürlich zwei Pivotspalten. Von vier Variablen lassen sich also die 2 Variablen, die zu den Pivotspalten gehören, ausrechnen. Die anderen 2 Variablen werden als freie Parameter angesetzt.

1.1.4. *Affine und lineare Unterräume in $\mathbb{R}^n$.*

Die Lösungsmenge des letzten Beispiels 1.1.23 hat die Gestalt

$$\{p + \lambda v_1 + \mu v_2 \mid \lambda, \mu \in \mathbb{R}\} \tag{1.1.2}$$

mit geeigneten Vektoren p, v_1, v_2. Das sieht genauso aus wie die Parameterform einer Ebene, nur dass sich unsere Lösungsmenge in einem vierdimensionalen Raum befindet. In diesem Unterabschnitt werden wir kursorisch geometrische Begriffe einführen, unter die sich Gerade, Ebene und auch höherdimensionale Gebilde systematisch einordnen lassen.

Wir vergessen zunächst den Vektor p aus Gleichung 1.1.2 und betrachten nur den Ausdruck $\lambda v_1 + \mu v_2$.

Definition 1.1.25. Seien $v_1, \cdots, v_k$ Vektoren in $\mathbb{R}^n$.

(1) Sind $\lambda_1, \cdots, \lambda_k \in \mathbb{R}$, so heißt der Vektor $\lambda_1 v_1 + \cdots + \lambda_k v_k$ eine Linearkombination von $v_1, \cdots, v_k$.
(2) Die lineare Hülle der Menge $\{v_1, \cdots, v_k\}$ ist die Menge aller Linearkombinationen dieser Vektoren:
$$\mathcal{L}(\{v_1, \cdots, v_k\}) := \{\lambda_1 v_1 + \cdots + \lambda_k v_k \mid \lambda_1, \cdots, \lambda_k \in \mathbb{R}\}\,.$$
(3) Die lineare Hülle der leeren Menge wird definiert als der Nullraum, das ist die Menge, die nur aus dem Nullvektor besteht:
$$\mathcal{L}(\emptyset) := \{0\}\,.$$

Definition 1.1.26. Sei U eine Teilmenge von $\mathbb{R}^n$. Wir sagen, U ist ein Unterraum von $\mathbb{R}^n$, wenn gilt

(1) U ist nicht leer.
(2) Sind $u, v \in U$ und $\lambda, \mu \in \mathbb{R}$, so ist auch $\lambda u + \mu v \in U$.

Im $\mathbb{R}^3$ ist der Nullraum ein Unterraum (der einzige Punkt, der einen Unterraum bildet). Geraden und Ebenen sind genau dann Unterräume, wenn sie durch den Nullpunkt gehen. Auch $\mathbb{R}^3$ selbst ist ein Unterraum des $\mathbb{R}^3$.

Bemerkung 1.1.27. Jede lineare Hülle einer Vektormenge $\{v_1, \cdots, v_k\} \subset \mathbb{R}^n$ ist ein Unterraum des $\mathbb{R}^n$. Wir sagen auch, $\mathcal{L}(\{v_1, \cdots, v_k\})$ ist der von $v_1, \cdots, v_n$ erzeugte oder aufgespannte Unterraum.

Beweis. Man sieht sofort, dass man den Nullvektor durch $\lambda_1 = \cdots = \lambda_n = 0$ als Linearkombination erzeugen kann. Also ist $\mathcal{L}(\{v_1, \cdots, v_k\})$ nicht leer. Seien nun $u, w \in \mathcal{L}(\{v_1, \cdots, v_k\})$ und $\alpha, \beta \in \mathbb{R}$. Wir wollen zeigen, dass auch $\alpha u + \beta w$ in $\mathcal{L}(\{v_1, \cdots, v_k\})$ liegt. Wegen der Definition von $\mathcal{L}(\{v_1, \cdots, v_k\})$ lassen sich u, w in der folgenden Weise schreiben:

$$u = \lambda_1 v_1 + \cdots + \lambda_k v_k, \quad w = \mu_1 v_1 + \cdots + \mu_k v_k \,.$$

Klammern mit Vektoren lassen sich ausmultiplizieren wie Klammern reeller Zahlen (Distributivgesetz), Klammern in Summen von Vektoren und die Reihenfolge der Addition lassen sich beliebig verschieben (Assoziativ- und Kommutativgesetz). Daher können wir die folgende Rechnung ausführen:

$$\begin{aligned} & \alpha u + \beta w \\ = \; & \alpha(\lambda_1 v_1 + \cdots + \lambda_k v_k) + \beta(\mu_1 v_1 + \cdots + \mu_k v_k) \\ = \; & \alpha\lambda_1 v_1 + \cdots + \alpha\lambda_k v_k + \beta\mu_1 v_1 + \cdots + \beta\mu_k v_k \\ = \; & (\alpha\lambda_1 + \beta\mu_1) v_1 + \cdots + (\alpha\lambda_k + \beta\mu_k) v_k \,. \end{aligned}$$

Weil ja auch die Klammern $(\alpha\lambda_i + \beta\mu_i)$ wieder reelle Zahlen ergeben, steht am Ende wieder eine Linearkombination von $v_1, \cdots, v_k$. □

Bemerkung 1.1.28. Bitte gewöhnen Sie sich möglichst schnell an den Gebrauch des Summenzeichens: Seien $p \leq q$ ganze Zahlen, und $v_p, \cdots, v_q$ Zahlen oder Vektoren, dann ist

$$\sum_{i=p}^{q} v_i := v_p + v_{p+1} + \cdots + v_q \,.$$

D.h.: Zähle die Objekte v_i von $i = p$ bis $i = q$ auf, und addiere sie alle.
Der Summationsindex i kann jeden Namen haben, er ist nur ein Platzhalter für die Zahlen $p, \cdots, q$, man könnte genausogut t oder jeden anderen Buchstaben schreiben.

Zum Beispiel schreibt sich der obige Beweis mit Summenzeichen so:

$$\alpha u + \beta w = \alpha \sum_{i=1}^{k} \lambda_i v_i + \beta \sum_{i=1}^{k} \mu_i v_i = \sum_{i=1}^{k} \alpha\lambda_i v_i + \sum_{i=1}^{k} \beta\mu_i v_i = \sum_{i=1}^{k} (\alpha\lambda_i + \beta\mu_i) v_i \,.$$

Wir sehen also, dass jede endliche Menge von Vektoren einen Unterraum erzeugt (eben ihre lineare Hülle). Andererseits lässt sich jeder Unterraum des $\mathbb{R}^n$ als lineare Hülle einer endlichen Vektormenge schreiben. Auch wenn uns das anschaulich einleuchtet, ist das nicht ganz selbstverständlich. Ist es wirklich ausgeschlossen, dass eine Menge im $\mathbb{R}^3$ existiert, die die Bedingungen eines Unterraumes erfüllt aber nicht als lineare Hülle von ein paar Vektoren aufgespannt wird?

Die Intuition sagt uns auch: Eine Ebene ist zweidimensional, und dass soll sagen: Wenn wir eine Ebene aufspannen wollen, die durch den Nullpunkt geht, brauchen wir immer zwei Vektoren, die nicht parallel sind. Es gelingt nie mit einem Vektor, und von drei Vektoren ist immer einer überflüssig. Ebenso braucht man für eine Gerade durch den Nullpunkt immer genau einen Vektor. Obwohl es also verschiedene Vektorsysteme gibt, die dieselbe Ebene aufspannen, kann man aus jedem System alle bis auf zwei Vektoren weglassen, und mit den restlichen zwei die Ebene aufspannen. Was für die Intuition klar ist, braucht später einen Beweis.

Im Augenblick geht es nur darum, anschauliche Bilder für die Theorie zu schaffen, die wir in dieser Vorlesung axiomatisch entwickeln werden. Dann werden wir den Begriff der Dimension auf feste, sichere Füße stellen und nicht mehr an die Anschauung appellieren müssen.

Um auch Geraden und Ebenen, die nicht durch den Nullpunkt gehen, in unser Begriffsrepertoire aufzunehmen, müssen wir einen Begriff schaffen, der weiter gefasst ist als der Unterraum. Die folgende Definition passt perfekt auf die Parameterform von Ebenen und Geraden. Gleichzeitig passt Sie auch zur Lösungsmenge von linearen Gleichungssystemen in Analogie zu Gleichung (1.1.2).

Definition 1.1.29. Sei W eine Teilmenge des $\mathbb{R}^n$, $p \in \mathbb{R}^n$ und U ein Unterraum des $\mathbb{R}^n$ so, dass gilt

$$W = \{p + u \mid u \in U\}.$$

Dann heißt W ein affiner Unterraum von $\mathbb{R}^n$ mit Tangentialraum U.

Bemerkung 1.1.30. Achtung, die Benennung dieser Objekte kann verwirren: Jeder Unterraum ist ein affiner Unterraum, aber nicht jeder affine Unterraum ist ein Unterraum!

Bemerkung 1.1.31. Der Tangentialraum umfasst also alle Richtungsvektoren, die von Punkten aus W zu Punkten aus W führen.

Vereinbarung 1.1.32. Seien $A, B \subset \mathbb{R}^n$ und $c \in \mathbb{R}^n$. Wir verwenden folgende Kurzschreibweisen:

$$\begin{aligned} A + B &= \{a + b \mid a \in A, b \in B\}, \\ c + B &= \{c + b \mid b \in B\}. \end{aligned}$$

Ein affiner Unterraum des $\mathbb{R}^n$ ist demnach eine Menge der Form $p + U$, wobei p ein Vektor und U ein Unterraum des $\mathbb{R}^n$ ist.

Wir wissen zum Beispiel, dass wir als Ausgangsform für die Parameterdarstellung einer Gerade oder Ebene jeden beliebigen Punkt der Geraden oder Ebene heranziehen können. Das wird etwas abstrakter und allgemeiner von der folgenden Bemerkung wiedergegeben:

Bemerkung 1.1.33. Sei W ein affiner Unterraum des $\mathbb{R}^n$ mit Tangentialraum U. Sei $q \in \mathbb{R}^n$. Es gilt:
q liegt genau dann in W, wenn $W = q + U$.

Beweis. Da wir eine Äquivalenz von zwei Aussagen beweisen, müssen wir zwei Richtungen beweisen.
Beweis von $q \in W \Rightarrow W = q + U$: Jedenfalls lässt sich W in der Weise $W = p + U$ mit einem geeigneten Vektor $p \in \mathbb{R}^n$ schreiben. Ist $q \in W$, so gibt es ein $v \in U$ mit $q = p + v$. Ist nun $u \in U$, so ist $v + u \in U$, weil U ein Unterraum ist. Dann ist aber $q + u = p + (v + u) \in p + U = W$. Also ist $q + U \subset W$. Ist umgekehrt $y \in W$, so kann man $y = p + u$ mit einem $u \in U$ schreiben. Nun ist auch $u - v \in U$, und $y = (p + v) + (u - v) = q + (u - v) \in q + U$. Also ist $W \subset q + U$.
Beweis von $W = q + U \Rightarrow q \in W$: Es ist $q = q + 0 \in q + U$, weil jeder Unterraum den Nullvektor enthält. □

Die folgende wichtige Beobachtung beweisen wir jetzt noch nicht, weil wir später viel bequemere Schreibweisen zur Verfügung haben werden:

Bemerkung 1.1.34. Sei $M \subset \mathbb{R}^n$ die Lösungsmenge eines linearen Gleichungssystems in n Unbekannten, und sei $M \neq \emptyset$. Sei $x \in M$ eine beliebige partikuläre Lösung des Gleichungssystems. Dann lässt sich M als affiner Unterraum des $\mathbb{R}^n$ in der Form $x + U$ schreiben, dabei ist U die Lösungsmenge des homogenen Gleichungssystems, das man erhält, indem man die rechte Seite durch den Nullvektor ersetzt.

2. Gruppen, Ringe, Körper

2.1. Mengen und Funktionen.

Bevor wir mit der axiomatischen Arbeit beginnen, müssen wir uns noch auf die Sprache der Mengen und Funktionen einigen.

2.1.1. *Mengen.*

Will man die Mengenlehre axiomatisch auf ein solides Fundament stellen, muss man sehr diffizile logische Probleme lösen. Mit unserer naiven Vorstellung: "Eine Menge ist eine Zusammenfassung irgendwelcher Element zu einem Ganzen" laufen wir sehr schnell in Widersprüche:

Beispiel 2.1.1 (Russelsche Antinomie)**.** Sei M die Menge aller Mengen, die sich selbst nicht als Element enthalten.

Die harmlos scheinende Definition verfängt sich in Widersprüche, wenn wir entscheiden sollen, ob M selbst ein Element von M ist:

> Sei $M \in M$. Die Menge M enthält nur Elemente, die sich selbst nicht als Element enthalten. Weil aber $M \in M$ vorausgesetzt wurde, kann also M kein Element von M sein.
> Sei $M \notin M$. Nach Definition von M muss also M zu den Elementen von M gehören.

Wie man es dreht, es kommt immer ein Widerspruch heraus. Die Definition von M liefert keine sinnvolle Menge!

Trotz dieser Gefahren können wir ruhig für diese Vorlesung und für weite Bereiche der Mathematik mit unserer naiven Vorstellung fortfahren. Wir wenden keine so starken Mittel der Mengenlehre an, wir definieren keine so komplizierten Mengen, dass wir auf die Feinheiten der Axiomatik achten müssen. Dieses Kapitel soll nur die Schreibweisen festlegen, mit denen wir arbeiten. Wir gehen auf die Problematik der Mengentheorie nicht ein.

Vereinbarung 2.1.2. Seien $a, b, c, \cdots$ irgendwelche Objekte, und sei $A(x)$ eine Aussage, die man über ein Objekt x machen kann, und die dann wahr oder falsch sein kann. Wir führen folgende Mengenschreibweisen ein:

(1) $\{a, b, c, \cdots\}$ ist die Menge, die die Objekte a, b, c usw. enthält (aufzählende Mengendeklaration).
(2) $\{x \mid A(x)\}$ ist die Menge aller Objekte x, für die die Aussage $A(x)$ wahr ist (beschreibende Mengendeklaration).

Die in einer Menge M enthaltenen Objekte nennen wir die Elemente von M. Ist x ein Element von M, schreiben wir $x \in M$. Die Menge, die gar kein Element enthält, heißt die leere Menge $\emptyset$.

Hier haben wir uns nicht getraut, von einer Definition zu reden. Mathematisch ist das alles sehr verschwommen. Was sind "irgendwelche Objekte"?

Vereinbarung 2.1.3. Seien A und B zwei Mengen. Wir sagen, A ist eine Teilmenge von B ($A \subset B$, B ist eine Obermenge von A, $B \supset A$), wenn jedes Element aus A auch in B enthalten ist.
A ist eine echte Teilmenge von B, wenn A eine Teilmenge von B ist und $A \neq B$ gilt. Wir schreiben auch $A \subsetneq B$.

Achtung, manche Autoren schreiben $A \subset B$ nur für echte Teilmengen, und im allgemeinen Fall $A \subseteq B$.

Vereinbarung 2.1.4. Seien $A, B, A_1, A_2, \cdots$ Mengen. Wir verwenden folgende Bezeichnungen:

$A \cap B$	Durchschnitt von A und B: Die Menge aller Elemente, die sowohl in A als auch in B liegen.
$A \cup B$	Vereinigung von A und B: Die Menge aller Elemente, die in A oder B oder beiden liegen.
$A \setminus B$	Differenzmenge. Die Menge aller Elemente von A, die nicht in B liegen.
$\bigcap_{i=1}^{\infty} A_i$	Durchschnitt aller A_i. Die Menge aller Elemente, die in jedem A_i enthalten sind.
$\bigcup_{i=1}^{\infty} A_i$	Vereinigung aller A_i. Die Menge aller Elemente, die mindestens in einem A_i enthalten sind.
$A \times B$	Produktmenge. Menge aller geordneten Paare (a, b) mit $a \in A$ und $b \in B$.

2.1.2. *Der Begriff der Funktion.*
Wir beschreiben den Begriff der Funktion zunächst auf eine Weise, die der Anschauung gut entspricht. Wir werden gleich eine saubere Definition nachtragen.

Vereinbarung 2.1.5. Seien A und B zwei nichtleere Mengen. Eine Funktion (= Abbildung) von A nach B ist eine Vorschrift, die zu jedem $a \in A$ genau ein $b \in B$ bestimmt.
Mit der Notation $f : A \to B$ drücken wir aus, dass f eine Funktion von A nach B ist. Mit $f : a \mapsto b$ oder mit $f(a) = b$ ("$f(a)$ ist b", "f bildet a auf b ab") drücken wir aus, dass f einem bestimmten Element $a \in A$ das Element $b \in B$ zuweist.
Die Menge A heißt der Definitionsbereich von f, die Menge B heißt Wertevorrat von f.

Das ist wieder eine sehr verwaschene Begriffserklärung: "Eine Vorschrift, die ..." — Für die Anschauung verständlich, aber mathematisch keine Definition. Wir wissen, dass der Begriff der Menge schon diffizil und gefährlich ist. Werden wir bei der Definition von Funktionen wieder Grundlagenprobleme bekommen, kann man "Vorschriften" definieren, die zwingend auf Widersprüche führen? Die gute Nachricht ist: Hat man einmal den Mengenbegriff auf die sichere Seite gerettet, kann man eine Funktion präzise definieren. Wir definieren die Funktion einfach als die Menge der Paare $(a, f(a))$:

Definition 2.1.6. Seien A und B nichtleere Mengen. Eine Funktion f ist eine Teilmenge der Produktmenge $A \times B$ mit der Eigenschaft:
Für jedes $a \in A$ gibt es genau ein $b \in B$, sodass das Paar (a, b) in f liegt.
Sind $a \in A$, $b \in B$, mit $(a, b) \in f$, so schreiben wir auch $b = f(a)$.
Will man speziell auf die Definition der Funktion durch Paare hinweisen, nennt man die Menge $\{(a, f(a)) \mid a \in A\}$ auch Graph der Funktion.

Für unsere praktische Arbeit genügt aber die anschauliche Vorstellung aus Vereinbarung 2.1.5.

Bemerkung 2.1.7. Seien A und B Mengen, und $f : A \to B$ eine Funktion. Beachten Sie, dass f nicht den vollen Wertevorrat B ausschöpfen muss, es darf "uberflüssige" Elemente in B geben. Dagegen muss sich f auf jedes Element in A anwenden lassen.

Bemerkung 2.1.8. Seien A und B Mengen, und $f : A \to B$ eine Funktion. Sei $a \in A$. Unterscheiden Sie genau:

f ist eine Funktion, also axiomatisch eine Menge geordneter Paare, anschaulich eine Abbildungsvorschrift.
$f(a)$ ist ein Element von B, jenes Element, auf das a abgebildet wird.

Beispiel 2.1.9. Hier finden Sie einige Beispiele von Funktionen:

(1) Jeder/m Studierenden wird genau eine Matrikelnummer zugeordnet, niemand von ihnen hat keine oder mehrere Nummern. Wir haben daher eine Funktion von der Menge der Studierenden auf die 7-stelligen natürlichen Zahlen. Umgekehrt findet sich zu jeder Matrikelnummer ein/e Studierende/r. Trotzdem ist das keine Funktion von der Menge der 7-stelligen natürlichen Zahlen auf die Menge der Studierenden, weil es Zahlen gibt, die nicht als Matrikelnummer vorkommen.
(2) Die Abbildung $x \mapsto \sqrt{x}$ ist eine Funktion von $[0, \infty)$ nach $[0, \infty)$.
(3) Die Abbildung $x \mapsto x^2$ ist eine Abbildung von $\mathbb{R}$ nach $\mathbb{R}$, denn jede reelle Zahl hat ihr eindeutiges Quadrat. Es macht nichts, dass der Wertevorrat nicht ausgeschöpft wird (negative Zahlen sind keine Quadrate), und auch nichts, dass jede positive Zahl zweimal als Wert (nämlich als Quadrat einer positiven und einer negativen Zahl) vorkommt.
(4) Seien $\alpha_1, \cdots, \alpha_n \in \mathbb{R}$. Die Formel

$$f : \begin{cases} \mathbb{R}^n & \to \mathbb{R} \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} & \mapsto \sum_{i=1}^n \alpha_i \xi_i \end{cases}$$

definiert eine Funktion.

Funktionen von Intervallen reeller Zahlen in Intervalle lassen sich grafisch als Kurven darstellen. Jedes Paar $(x, f(x))$ des Graphen von f wird als Koordinatenpaar eines Punktes der Ebene dargestellt. Der gesamte Graph ergibt bei (einigermaßen "gutmütigen") Funktionen eine Kurve. Funktionen von $\mathbb{R}^2$ nach $\mathbb{R}$ kann man als Flächen im dreidimensionalen Raum darstellen.

Eine Funktion, die auf einer endlichen Menge A definiert ist, lässt sich auch als Tabelle schreiben. Eine Funktion einer endlichen Menge A nach $\mathbb{R}$ kann man als Stabdiagramm darstellen. Funktionen von einer endlichen Menge A in eine andere endliche Menge B lässt sich als Pfeildiagramm schreiben, indem man die Elemente von A und von B als Punkte darstellt, und jeweils einen Pfeil von a nach $f(a)$ zieht.

Definition 2.1.10. Seien A und B Mengen und $f : A \to B$ eine Funktion. Seien $U \subset A$ und $V \subset B$ Teilmengen (die auch leer sein dürfen). Wir definieren

(1) Das Bild von U besteht aus allen Elementen von B, die als Werte $f(u)$ mit einem $u \in U$ vorkommen: $f(U) := \{f(u) \mid u \in U\}$.
(2) Das Urbild von V besteht aus allen Elementen von A, die durch f in Elemente von V abgebildet werden: $f^{-1}(V) = \{a \in A \mid f(a) \in V\}$.

Bemerkung 2.1.11. Bitte beachten Sie unbedingt, dass es zur Definition des Urbildes einer *Menge* unter einer Funktion f nicht notwendig ist, dass f eine Umkehrfunktion besitzt, auch wenn es durch die Schreibweise suggeriert wird.
Wenn die Funktion den Wertevorrat nicht voll ausschöpft, dann gibt es auch nichtleere Mengen, deren Urbild die leere Menge ist. Dagegen gibt es nie eine nichtleere Menge, deren Bild leer ist.

Beispiel 2.1.12.

(1) Sei $f : \mathbb{R} \to \mathbb{R}$ gegeben durch $f(x) = x^2$. Sei $0 < a < b < \infty$. Dann sind Bild und Urbild des Intervalles (a, b) gegeben durch

$$f((a,b)) = (a^2, b^2), \quad f^{-1}((a,b)) = (-\sqrt{b}, -\sqrt{a}) \cup (\sqrt{a}, \sqrt{b}) .$$

(2) Seien $\alpha_{i,j} \in \mathbb{R}$ für $i = 1, \cdots, m$, $j = 1, \cdots, n$. Wir betrachten die Funktion

$$f : \begin{cases} \mathbb{R}^n & \to \mathbb{R}^m \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} & \mapsto \begin{pmatrix} \alpha_{1,1}\xi_1 + \cdots + \alpha_{1,n}\xi_n \\ \vdots \\ \alpha_{m,1}\xi_1 + \cdots + \alpha_{m,n}\xi_n \end{pmatrix} \end{cases} .$$

Das Urbild des Nullraumes ist die Lösungsmenge des homogenen linearen Gleichungssystems

$$\begin{array}{ccccccc} \alpha_{1,1}\xi_1 & + & \cdots & + & \alpha_{1,n}\xi_n & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ \alpha_{m,1}\xi_1 & + & \cdots & + & \alpha_{m,n}\xi_n & = & 0 \end{array}$$

Die Lösungsmenge kann aus einem Vektor (dem Nullvektor) oder einem mehrdimensionalen Unterraum des $\mathbb{R}^n$ bestehen.

Definition 2.1.13. Seien A und B Mengen, und $f : A \to B$ eine Funktion. Sei C eine Teilmenge von A. Die Einschränkung von f auf C ist die Funktion

$$f|_C : \begin{cases} C & \to B \\ x & \mapsto f(x) \end{cases} .$$

Die Einschränkung hat also dieselbe Abbildungsvorschrift wie f, aber der Definitionsbereich ist auf die Teilmenge C zurückgestutzt.

2.1.3. *Hintereinanderausführung, Injektivität, Surjektivität, Umkehrfunktion.*

Definition 2.1.14. Seien A, B, C Mengen, und $f : A \to B$, $g : B \to C$ Funktionen. Die Hintereinanderausführung (Komposition) $g \circ f$ ("g nach f") ist die Funktion

$$g \circ f : \begin{cases} A & \to C \\ a & \mapsto g(f(a)) \end{cases} .$$

Bemerkung 2.1.15. Achten Sie darauf, dass die Hintereinanderausführung $g \circ f$ nur existiert, wenn der Wertevorrat von f im Definitionsbereich von g als Teilmenge enthalten ist.

Satz 2.1.16. *Die Hintereinanderausführung von Funktionen ist assoziativ im folgenden Sinn:*
Seien A, B, C, D Mengen, und $f : A \to B$, $g : B \to C$, $h : C \to D$ Funktionen. Dann ist $(h \circ g) \circ f = h \circ (g \circ f)$.

Beweis. Es ist $g \circ f$ eine Funktion von A nach C und h eine Funktion von C nach D, daher existiert $h \circ (g \circ f)$ und ist eine Funktion von A nach D. Dabei ist für $a \in A$

$$[h \circ (g \circ f)](a) = h([g \circ f](a)) = h(g(f(a)) \,.$$

Es ist f eine Funktion von A nach B, und $h \circ g$ eine Funktion von B nach D. Daher existiert $(h \circ g) \circ f$ und es ist

$$[(h \circ g) \circ f](a) = (h \circ g)(f(a)) = h(g(f(a)) \,.$$

Beide Formeln geben dieselbe Funktion. □

Bemerkung 2.1.17. Die Reihenfolge der Hintereinanderausführung darf aber im Allgemeinen nicht vertauscht werden.

Beispiel 1: Seien $f, g : \mathbb{R} \to \mathbb{R}$ gegeben durch $f(x) = x^2$, $g(x) = x + 1$. Es ist

$$[g \circ f](x) = g(x^2) = x^2 + 1, \quad [f \circ g](x) = f(x + 1) = (x + 1)^2 \,.$$

Beispiel 2: Seien $f : \mathbb{R} \to \mathbb{R}^2$ und $g : \mathbb{R}^2 \to \mathbb{R}^3$ Funktionen. Es existiert dann $g \circ f : \mathbb{R} \to \mathbb{R}^3$, aber nicht $f \circ g$. Denn ist $x \in \mathbb{R}^2$, so ist $g(x) \in \mathbb{R}^3$, aber $f(g(x))$ ist nicht definiert, weil $f(y)$ nur für $y \in \mathbb{R}$ existiert. □

Definition 2.1.18. Sei M eine beliebige nichtleere Menge, so ist die Identität auf M definiert durch

$$\mathrm{id}_M : \begin{cases} M & \to M \\ x & \mapsto x \end{cases} .$$

Bemerkung 2.1.19. Seien A und B Mengen und $f : A \to B$ eine Funktion. Dann gilt

$$f = \mathrm{id}_B \circ f = f \circ \mathrm{id}_A \,.$$

Definition 2.1.20. Seien A und B Mengen und $f : A \to B$ eine Funktion. Sei $g : B \to A$ eine Funktion mit der Eigenschaft

$$g \circ f = \mathrm{id}_A, \quad f \circ g = \mathrm{id}_B \,.$$

Dann heißt g Umkehrfunktion (inverse Funktion) zu f, und wir schreiben $g = f^{-1}$.

Das bedeutet: Die Funktion und die Umkehrfunktion heben einander auf. Führt man erst f, und anschließend f^{-1} aus, kommt man zum Ausgangswert zurück. Ebenso, wenn man erst f^{-1} und dann f ausführt.

Die Schreibweise $g = f^{-1}$ wird durch die folgende Bemerkung gerechtfertigt:

Bemerkung 2.1.21. Sei $f : A \to B$ eine Funktion. Dann besitzt f höchstens eine Umkehrfunktion.

Beweis. Seien $g : B \to A$ und $h : B \to A$ Umkehrfunktionen von f. Dann ist

$$g = g \circ \mathrm{id}_B = g \circ (f \circ h) = (g \circ f) \circ h = \mathrm{id}_A \circ h = h \,.$$

□

Bemerkung 2.1.22. Seien A und B Mengen, und $f : A \to B$ eine Funktion. Dann gilt:
Die Funktion f besitzt genau dann eine Umkehrfunktion, wenn für jedes $y \in B$ die Gleichung $f(x) = y$ genau eine Lösung $x \in A$ hat. Diese Lösung ist dann $f^{-1}(y)$.

Beweis. Wir nehmen zunächst an, dass für jedes $y \in B$ genau ein x existiert, sodass $f(x) = y$. Wir definieren die Funktion

$$g : \begin{cases} B & \to A \\ y & \mapsto x \text{ mit } f(x) = y \end{cases} .$$

Wir zeigen, dass g die Umkehrfunktion von f ist. Sei $z \in A$. Setze $y = f(z)$. Dann ist $g(f(z)) = g(y)$ die Lösung x von $f(x) = y = f(z)$, davon gibt es nach Voraussetzung genau eine, und das ist natürlich z. Also ist $g \circ f = \mathrm{id}_A$. Sei nun $y \in B$ beliebig. Es ist $g(y)$ die eindeutige Lösung x der Gleichung $f(x) = y$, und damit ist natürlich $f(g(y)) = f(x) = y$. Also ist $f \circ g = \mathrm{id}_B$.
Nun nehmen wir an, dass f eine Umkehrfunktion f^{-1} besitzt. Es ist $x = f^{-1}(y)$ eine Lösung von $f(x) = y$, denn

$$f(x) = f(f^{-1}(y)) = [f \circ f^{-1}](y) = \mathrm{id}_B\, y = y\,.$$

Ist z irgendeine Lösung von $f(z) = y$, so ist

$$z = \mathrm{id}_A(z) = [f^{-1} \circ f](z) = f^{-1}(f(z)) = f^{-1}(y)\,.$$

Daher ist $x = f^{-1}(y)$ die einzige Lösung der Gleichung $f(x) = y$. □

Beispiel 2.1.23. Sei $f : \mathbb{R} \to \mathbb{R}$ gegeben durch $f(x) = x^3 + 1$. Gibt es eine Umkehrfunktion?

Lösung: Ist $y \in \mathbb{R}$, so muss $x = f^{-1}(y)$ die Gleichung $x^3 + 1 = y$ lösen. Wir formen diese Gleichung äquivalent um:

$$x^3 + 1 = y \iff x^3 = y - 1 \iff x = \sqrt[3]{y-1}\,.$$

Wir haben also zu jedem $y \in \mathbb{R}$ genau eine Lösung gefunden. Also ist $f^{-1}(y) = \sqrt[3]{y-1}$. □

Beispiel 2.1.24. Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

$$f(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} := \xi_1 + \xi_2\,.$$

Gibt es eine Umkehrfunktion?

Lösung: Sei $y \in \mathbb{R}$. Wir untersuchen, ob die Gleichung $f(x) = y$ eindeutig lösbar ist. Nun sind sowohl der Vektor $\begin{pmatrix} y \\ 0 \end{pmatrix}$ als auch $\begin{pmatrix} 0 \\ y \end{pmatrix}$ Lösungen. Daher ist die Lösung nicht eindeutig, und es gibt keine Umkehrfunktion. □

Definition 2.1.25. Seien A und B Mengen, und $f : A \to B$ eine Funktion. Wir sagen

(1) Die Funktion $f : A \to B$ heißt surjektiv, wenn $f(A) = B$, d.h., wenn alle Elemente von B als Werte $f(a)$ mit geeigneten $a \in A$ vorkommen.
(2) Die Funktion $f : A \to B$ heißt injektiv, wenn für alle $x \neq y \in A$ gilt: $f(x) \neq f(y)$, wenn also verschiedene Elemente aus A auf verschiedene Elemente aus B abgebildet werden.
(3) Die Funktion $f : A \to B$ heißt bijektiv, wenn sie sowohl injektiv als auch surjektiv ist.

Die folgende Bemerkung ist einfach dasselbe mit anderen Worten:

Bemerkung 2.1.26. Seien A und B Mengen, und $f : A \to B$ eine Funktion. Es gilt:

(1) $f : A \to B$ ist genau dann surjektiv, wenn für jedes $y \in B$ die Gleichung $f(x) = y$ mindestens eine Lösung $x \in A$ besitzt.

(2) $f : A \to B$ ist genau dann injektiv, wenn für jedes $y \in B$ die Gleichung $f(x) = y$ höchstens eine Lösung $x \in A$ besitzt.
(3) $f : A \to B$ ist genau dann bijektiv, wenn für jedes $y \in B$ die Gleichung $f(x) = y$ genau eine Lösung $x \in A$ besitzt.

Bemerkung 2.1.27. Seien A und B Mengen, und $f : A \to B$ eine Funktion. Dann gilt:
Die Funktion f ist genau dann bijektiv, wenn sie eine Umkehrfunktion besitzt.

Beweis. Bemerkungen 2.1.26 und 2.1.22. □

Bemerkung 2.1.28. Seien A, B und C Mengen, und $f : A \to B$, $g : B \to C$ Funktionen. Es gilt:

(1) Sind $f : A \to B$ und $g : B \to C$ surjektiv, so ist auch $g \circ f : A \to C$ surjektiv.
(2) Sind $f : A \to B$ und $g : B \to C$ injektiv, so ist auch $g \circ f : A \to C$ injektiv.
(3) Sind $f : A \to B$ und $g : B \to C$ bijektiv, so ist auch $g \circ f : A \to C$ bijektiv. In diesem Fall ist $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Beweis. Seien f, g surjektiv. Sei $z \in C$. Dann gibt es wegen der Surjektivität von g eine Lösung $y \in B$ von $g(y) = z$, und wegen der Surjektivität von f eine Lösung $x \in A$ von $f(x) = y$. Es ist dann $[g \circ f](x) = g(f(x)) = g(y) = z$, also hat $[g \circ f](x) = y$ eine Lösung.
Seien nun f und g injektiv, und seien $x_1 \neq x_2 \in A$. Wegen Injektivität von f ist $f(x_1) \neq f(x_2)$, und wegen der Injektivität von g ist $g(f(x_1)) \neq g(f(x_2))$. Also ist $[g \circ f](x_1) \neq [g \circ f](x_2)$.
Der dritte Punkt folgt unmittelbar aus den beiden ersten. Es gilt auch wegen der Assoziativität der Hintereinanderausführung (Satz 2.1.16)

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ \mathrm{id}_B \circ f = f^{-1} \circ f = \mathrm{id}_A \,,$$
$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ \mathrm{id}_B \circ g^{-1} = g \circ g^{-1} = \mathrm{id}_C \,.$$

Also ist $f^{-1} \circ g^{-1}$ die Umkehrfunktion zu $g \circ f$. □

Bemerkung 2.1.29. Beachten Sie die Umkehr der Reihenfolge: $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

2.2. **Gruppen.**

2.2.1. *Definition der Gruppe und Beispiele.*

Definition 2.2.1. Sei M eine nichtleere Menge und $\circ$ eine Abbildung:

$$\circ : \begin{cases} M \times M & \to M \\ (a, b) & \mapsto a \circ b \end{cases}$$

Dann heißt $\circ$ eine innere Verknüpfung auf M. Das Paar $(M, \circ)$ heißt dann eine algebraische Struktur.

Bemerkung 2.2.2.

(1) Das Symbol $\circ$ steht in Definition 2.2.1 für irgendeine Verknüpfung und muss nicht die Hintereinanderausführung von Funktionen bedeuten.
(2) Wir reden von einer inneren Verknüpfung, weil alles innerhalb einer Menge passiert: Die Elemente $a, b \in M$ werden verknüpft, und es entsteht wieder ein Element $a \circ b$ aus M.

Beispiel 2.2.3.

(1) Die Addition und die Multiplikation auf der Menge der natürlichen Zahlen sind eine innere Verknüpfungen.

(2) Die Division ergibt auf der Menge der ganzen Zahlen keine innere Verknüpfung, denn auch wenn a, b ganze Zahlen sind, muss a/b keine ganze Zahl sein, die Verknüpfung führt also aus $\mathbb{Z}$ heraus. Aber selbst auf der Menge $\mathbb{Q}$ der rationalen Zahlen (Brüche von ganzen Zahlen) ergibt sich keine innere Verknüpfung, weil $a/0$ nicht definiert ist.

(3) Verknüpfungen endlicher Mengen lassen sich mit einer Verknüpfungstafel darstellen:

$\circ$	a	b	$\cdots$	z
a	$a \circ a$	$a \circ b$	$\cdots$	$a \circ z$
b	$b \circ a$	$b \circ b$	$\cdots$	$b \circ z$
$\vdots$	$\vdots$	$\vdots$		$\vdots$
z	$z \circ a$	$z \circ b$		$z \circ z$

(4) Die Menge $\{0, 1, -1\}$ bildet mit der Multiplikation eine algebraische Struktur. Wir zeigen die Verknüpfungstafel:

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

(5) Sei Z eine endliche oder unendliche Menge (wir denken an die Buchstaben eines Alphabets oder sonstigen Zeichensatzes). Ein Wort über Z ist eine endliche Folge von Elementen aus Z. Sind $a = (a_1, a_2, \cdots, a_m)$ und $b = (b_1, \cdots, b_n)$ Wörter über Z, so definieren wir die Verkettung $\circ$ durch

$$a \circ b := (a_1, \cdots, a_m, b_1, \cdots, b_n)\,.$$

Die Menge der Wörter über Z mit der Verkettung ist eine algebraische Struktur (die sogenannte freie Halbgruppe).

Definition 2.2.4. Sei M eine Menge mit einer inneren Verknüpfung $\circ$.

- Die Verknüpfung $\circ$ heißt assoziativ, wenn gilt:
$$(\forall x, y, z \in M)\ x \circ (y \circ z) = (x \circ y) \circ z.$$
- Die Verknüpfung $\circ$ heißt kommutativ, wenn gilt:
$$(\forall x, y \in M)\ x \circ y = y \circ x.$$
- Ein Element $e \in M$ heißt ein neutrales Element, wenn gilt
$$(\forall x \in M)\ e \circ x = x \circ e = x.$$
- Gibt es ein neutrales Element $e \in M$, und ist $x \in M$, so heißt ein Element $y \in M$ inverses Element zu x falls gilt
$$x \circ y = y \circ x = e.$$

Beobachtung 2.2.5. Bei assoziativen Verknüpfungen kann man Klammern beliebig umgruppieren. Daher müssen bei assoziativen Verknüpfungen keine Klammern geschrieben werden.

Definition 2.2.6. Sei G eine Menge mit einer inneren Verknüpfung $\circ$. Das Paar $(G, \circ)$ heißt eine Gruppe, wenn gilt:

- $\circ$ ist assoziativ.

- $(\exists e \in G)(\forall x \in G)\ e \circ x = x$.
- $(\forall x \in G)(\exists y \in G)\ y \circ x = e$.

Ist $\circ$ zusätzlich kommutativ, so heißt $(G, \circ)$ eine kommutative Gruppe oder abelsche Gruppe.

Satz 2.2.7. *Sei $(G, \circ)$ eine Gruppe. Dann gilt*

(1) *e ist ein neutrales Element, d.h., es gilt auch $(\forall x \in G) x \circ e = x$.*
(2) *Sei $x \in G$ und sei $y \in G$ so dass $y \circ x = e$. Dann ist auch $x \circ y = e$, und damit ist y ein inverses Element zu x und x ein inverses Element zu y.*
(3) *Es gibt genau ein neutrales Element in G.*
(4) *Zu jedem $x \in G$ gibt es genau ein inverses Element in G.*
(5) *Sind $x, y, z \in G$, so gelten folgende Kürzungsregeln:*

$$\begin{aligned} x \circ z = y \circ z \quad &\Rightarrow \quad x = y, \\ z \circ x = z \circ y \quad &\Rightarrow \quad x = y. \end{aligned}$$

Beweis. Wir zeigen zuerst (2): Sei $x \in G$, ferner sei $y \in G$ mit $y \circ x = e$ und $z \in G$ mit $z \circ y = e$. Es ist dann:

$$x \circ y = e \circ (x \circ y) = (z \circ y) \circ (x \circ y) = z \circ (y \circ x) \circ y = z \circ (e \circ y) = z \circ y = e.$$

Wir zeigen (1): Wieder sei $x \in G$ und y so dass $y \circ x = x \circ y = e$. Es ist

$$x \circ e = x \circ (y \circ x) = (x \circ y) \circ x = e \circ x = x.$$

Wir zeigen (3): Seien e, f zwei neutrale Elemente in G. Dann gilt

$$f = e \circ f = e.$$

Wir zeigen (4): Seien y, z beide inverse Elemente zu $x \in G$. Dann gilt

$$z = z \circ e = z \circ (x \circ y) = (z \circ x) \circ y = e \circ y = y.$$

Letztlich zeigen wir (5). Seien $x, y, z \in G$ und t das inverse Element zu z.

$$\begin{aligned} x \circ z = y \circ z \quad &\Rightarrow \quad (x \circ z) \circ t = (y \circ z) \circ t \ \Rightarrow\ x \circ (z \circ t) = y \circ (z \circ t) \\ &\Rightarrow \quad x \circ e = y \circ e \ \Rightarrow\ x = y. \end{aligned}$$

Die zweite Kürzungsregel folgt ebenso. □

Definition 2.2.8. Sei $(G, \circ)$ eine Gruppe mit neutralem Element e, sei $x \in G$. Das eindeutig bestimmte Element $y \in G$ mit $x \circ y = y \circ x = e$ bezeichnen wir als x^{-1} das inverse Element zu x.
Wird die Verknüpfung mit einem Additionszeichen geschrieben, bezeichnen wir auch oft mit 0 das neutrale Element und mit $-x$ das inverse Element zu x.

Gruppen lassen sich auch durch die eindeutige Lösbarkeit von Gleichungen charakterisieren:

Satz 2.2.9. *Sei G eine nichtleere Menge mit einer assoziativen Verknüpfung $\circ$. Die folgenden Aussagen sind äquivalent:*

1) *Für alle $a, b \in G$ besitzen die beiden folgenden Gleichungen jeweils eine eindeutige Lösung $x \in G$ bzw. $y \in G$:*

$$\begin{aligned} a \circ x &= b, \\ y \circ a &= b. \end{aligned}$$

2) *Für alle $a, b \in G$ besitzen die beiden folgenden Gleichungen jeweils mindestens eine Lösung $x \in G$ bzw. $y \in G$:*

$$\begin{aligned} a \circ x &= b, \\ y \circ a &= b. \end{aligned}$$

3) $(G, \circ)$ *ist eine Gruppe.*

Beweis. Beispiel eines Ringschlusses!

(1) $\Rightarrow$ (2): Trivial!

(2) $\Rightarrow$ (3): Wir suchen zuerst ein neutrales Element, dazu wählen wir ein beliebiges $a \in G$ und lösen die Gleichung $e \circ a = a$. Wir zeigen nun, dass für alle $x \in G$ gilt $e \circ x = x$. Ist $x \in G$, so gibt es $y \in G$ so, dass $a \circ y = x$. Nun ist

$$e \circ x = e \circ (a \circ y) = (e \circ a) \circ y = a \circ y = x.$$

Nun suchen wir zu x ein inverses Element, das gibt es aber, weil $y \circ x = e$ eine Lösung besitzt.

(3) $\Rightarrow$ (1). Wir lösen die Gleichung $a \circ x = b$: Eine Lösung ist $a^{-1} \circ b$, denn es ist

$$a \circ (a^{-1} \circ b) = (a \circ a^{-1}) \circ b = e \circ b = b.$$

Es kann aber nur eine Lösung geben: Ist $a \circ x = b = a \circ y$, so ist wegen der Kürzungsregel $x = y$. □

Beispiel 2.2.10.

(1) $(\mathbb{Z}, +)$ und $(\mathbb{R}, +)$, auch $(\mathbb{R}^n, +)$ sind kommutative Gruppen.
(2) $(\mathbb{N}, +)$ ist keine Gruppe, weil es für kein $x \in \mathbb{N}$ ein inverses Element innerhalb von $\mathbb{N}$ gibt.
(3) $(\mathbb{Z}, -)$ ist keine Gruppe, weil die Subtraktion nicht assoziativ ist: $1 - (1 - 1) = 1$, aber $(1 - 1) - 1 = -1$.
(4) $((0, \infty), \cdot)$ ist eine abelsche Gruppe.
(5) $(\mathbb{R}, \cdot)$ ist keine Gruppe, weil es zu 0 kein inverses Element bezüglich der Multiplikation gibt.

Viele Beispiele von Gruppen ergeben sich für Mengen von bijektiven Abbildungen:

Satz 2.2.11. *Sei M eine nichtleere Menge, und $\mathcal{S}(M)$ die Menge der bijektiven Abbildungen von M nach M. Dann ist $\mathcal{S}(M)$ mit der Hintereinanderausführung eine Gruppe.*

Beweis. Die Identität $\mathrm{id}_M : M \to M$ ist jedenfalls bijektiv und daher in $\mathcal{S}(M)$. Die Identität ist das neutrale Element. Sind $f, g \in \mathcal{S}(M)$, also bijektive Abbildungen von M nach M, so ist auch die Hintereinanderausführung $g \circ f$ bijektiv, also in $\mathcal{S}(M)$. Das inverse Element zu f ist die Umkehrabbildung f^{-1}, die natürlich ebenfalls bijektiv ist. Die Assoziativität der Hintereinanderausführung wurde schon in Satz 2.1.16 gezeigt. □

Schreibweise 2.2.12. Die Gruppe der bijektiven Abbildungen der Menge $\{1 \cdots n\}$ auf sich heißt die Permutationengruppe oder Symmetrische Gruppe $\mathcal{S}_n$.

Bemerkung 2.2.13. Die Permutationengruppe $\mathcal{S}_3$ ist nicht kommutativ.

Gegenbeispiel: Betrachten Sie die Permutationen

$$f : \begin{cases} \{1,2,3\} & \to \{1,2,3\} \\ 1 & \mapsto 2 \\ 2 & \mapsto 1 \\ 3 & \mapsto 3 \end{cases}, \quad g : \begin{cases} \{1,2,3\} & \to \{1,2,3\} \\ 1 & \mapsto 1 \\ 2 & \mapsto 3 \\ 3 & \mapsto 2 \end{cases}.$$

Wir erhalten als Hintereinanderausführungen, wie Sie leicht nachrechnen können (versuchen Sie es):

$$g \circ f : \begin{cases} \{1,2,3\} & \to \{1,2,3\} \\ 1 & \mapsto 3 \\ 2 & \mapsto 1 \\ 3 & \mapsto 2 \end{cases}, \quad f \circ g : \begin{cases} \{1,2,3\} & \to \{1,2,3\} \\ 1 & \mapsto 2 \\ 2 & \mapsto 3 \\ 3 & \mapsto 1 \end{cases}.$$

□

Beispiel 2.2.14. Seien A, B, C, D die Ecken eines Rechtecks, keines Quadrates, in der Ebene, durchnummeriert so dass A links unten liegt und die Reihung dem Gegenuhrzeigersinn folgt. Eine Bewegung (Kongruenzabbildung) in der Ebene (oder allgemein im $\mathbb{R}^n$ ist eine Abbildung, die alle Längen — und damit auch alle Winkel — gleich läßt. Die Bewegungen, die das Rechteck wieder in sich selbst überführen, heißen die Deckbewegungen des Rechtecks. Das sind

id Die Identität auf $\mathbb{R}^2$,

$$(A, B, C, D) \mapsto (A, B, C, D)\,,$$

s_w die Spiegelung um die waagrechte Symmetrieachse des Rechtecks,

$$(A, B, C, D) \mapsto (D, C, B, A)\,,$$

s_s die Spiegelung um die senkrechte Symmetrieachse des Rechtecks,

$$(A, B, C, D) \mapsto (B, A, D, C)\,,$$

d die Drehung um den Mittelpunkt des Rechtecks um 180°,

$$(A, B, C, D) \mapsto (C, D, A, B)\,.$$

Dann bilden die Deckbewegungen des Rechtecks eine kommutative Gruppe.

Beweis. Wir stellen uns erst die Gruppentafel auf: (Versuchen Sie das selbst!)

$\circ$	id	s_w	s_s	d
id	id	s_w	s_s	d
s_w	s_w	id	d	s_s
s_s	s_s	d	id	s_w
d	d	s_s	s_w	id

Die Assoziativität gilt allgemein für die Hintereinanderausführung von Abbildungen. Ebenso gilt für Abbildungen immer, dass die Indentität als neutrales Element bei der Hintereinanderausführung auftritt. Sehen wir die Gruppentafel an, so bemerken wir, dass jedes Element zu sich selbst invers ist. Da die Gruppentafel, gespiegelt an der Diagonalen von links oben nach rechts unten, sich nicht ändert, ist die Verknüpfung kommutativ. □

2.2.2. *Homomorphismen.*

Beispiel 2.2.15. Wir betrachten die Gruppe G der Bewegungen in der Ebene, welche den Nullpunkt festlassen, mit der Hinteranderausführung als Gruppenverknüpfung. Manche von ihnen (Spiegelungen) kehren die Orientierung der Objekte um, manche (Drehungen) lassen die Orientierung unverändert. Wir definieren eine Abbildung

$$f : \begin{cases} G & \to \{1, -1\} \\ & \\ \phi & \mapsto +1 \text{ wenn } \phi \text{ die Orientierung nicht umkehrt} \\ \phi & \mapsto -1 \text{ wenn } \phi \text{ die Orientierung umkehrt} \end{cases}.$$

Betrachten wir die Menge $\{1, -1\}$ mit der Multiplikation als Gruppe, so erhalten wir für alle $\phi, \psi \in G$:

$$f(\phi \circ \psi) = f(\phi) \cdot f(\psi)\,.$$

Wir erhalten also eine Abbildung zwischen zwei Gruppen, die mit den Gruppenoperationen verträglich ist. Es ist gleichgültig, ob erst die Verknüpfung und dann die Abbildung, oder erst die Abbildung und dann die Verknüpfung ausgeführt wird.

Beispiel 2.2.16. Wir betrachten die Gruppe $\mathbb{Z}, +$ und die Gruppe[6] $\mathbb{Z}/2\mathbb{Z} := \{[0], [1]\}$ mit der Gruppentafel

$\oplus$	[0]	[1]
[0]	[0]	[1]
[1]	[1]	[0]

Wir betrachten die Abbildung

$$\pi : \begin{cases} \mathbb{Z} & \to \mathbb{Z}/2\mathbb{Z} \\ & \\ x & \mapsto [0] \text{ wenn } x \text{ gerade ist} \\ x & \mapsto [1] \text{ wenn } x \text{ ungerade ist} \end{cases}.$$

Dann gilt wieder für $x, y \in \mathbb{Z}$:

$$\pi(x + y) = \pi(x) \oplus \pi(y)\,.$$

Beispiel 2.2.17. Wir betrachten die Gruppe $G = \{d_\phi \mid \phi \in \{0, 90, 180, 270\}\}$ der Drehungen um die Vielfachen von 90° (wobei eine Drehung um 360° dasselbe ist wie eine Drehung um 0°, mit der Hintereinanderausführung. Wir betrachten andererseits die Gruppe $H := \{1, i, -1, -i\}$ mit der Multiplikation, dabei ist i die imaginäre Zahl[7] mit der Eigenschaft $i^2 = -1$. Wir stellen die beiden Gruppentafeln auf. (Dabei schreiben wir in der Gruppentafel der Drehungen der Einfachheit halber nur die Drehwinkel ein.)

$\circ$	0	90	180	270
0	0	90	180	270
90	90	180	270	0
180	180	270	0	90
270	270	0	90	180

$\cdot$	1	i	-1	-i
1	1	i	-1	-i
i	i	-1	-i	1
-1	-1	-i	1	i
-i	-i	1	i	-1

[6]Wir werden diese Gruppe bald als eine Restklassengruppe besser verstehen können.

[7]Wir werden das besser verstehen, wenn wir die komplexen Zahlen diskutieren. Jedenfalls lässt sich ein solches Objekt konstruieren, aber natürlich ist es keine reelle Zahl.

Vergleichen Sie die Gruppentafeln. Wenn Sie in der ersten Tafel systematisch folgende Ersetzung durchführen:

$$f : \begin{cases} G & \to H \\ & \\ d_0 & \mapsto 1 \\ d_{90} & \mapsto i \\ d_{180} & \mapsto -1 \\ d_{270} & \mapsto -i \end{cases},$$

dann erhalten Sie genau die Gruppentafel von H. Bis auf die Namen der Elemente haben die Gruppen G und H die gleichen Eigenschaften. Wir haben hier eine bijektive Abbildung f mit der Eigenschaft $f(x \circ y) = f(x) \cdot f(y)$.

Definition 2.2.18. Seien $(G, \circ)$ und $(H, \odot)$ zwei Gruppen, und $f : G \to H$ eine Abbildung. Dann heißt f ein Gruppenhomomorphismus, wenn gilt

$$(\forall x, y \in G)\ f(x) \odot f(y) = f(x \circ y).$$

Ist f außerdem bijektiv, dann heißt f ein Gruppenisomorphismus. Homomorphismen einer Gruppe in sich selbst heißen Endomorphismen, und Isomorphismen einer Gruppe auf sich selbst heißen Automorphismen.
Gibt es einen Isomorphismus von $(G, \circ)$ nach $(H, \odot)$, so heißen die Gruppen $(G, \circ)$ und $(H, \odot)$ zueinander isomorph. (Wir schreiben dann $(G, \circ) \simeq (H, \odot)$.)

Satz 2.2.19.

(1) *Die Hintereinanderausführung von zwei Gruppenhomomorphismen ist ein Gruppenhomomorphismus.*
(2) *Die Umkehrabbildung eines Gruppenisomorphismus ist ein Homomorphismus (und folglich wieder ein Isomorphismus).*
(3) *Sind $(G, \cdot)$, $(H, \odot)$ und $(K, *)$ drei Gruppen, so gilt:*
   (a) $(G, \cdot) \simeq (G, \cdot)$.
   (b) *Ist $(G, \cdot) \simeq (H, \odot)$ und $(H, \odot) \simeq (K, *)$, dann ist $(G, \cdot) \simeq (K, *)$.*
   (c) *Ist $(G, \cdot) \simeq (H, \odot)$, so ist auch $(H, \odot) \simeq (G, \cdot)$.*

Beweis. Im ganzen Beweis bezeichnen $(G, \cdot)$, $(H, \odot)$ und $(K, *)$ drei Gruppen.
Wir zeigen 1.). Seien $u : G \to H$ und $v : H \to K$ zwei Gruppenhomomorphismen. Für alle Elemente $a, b \in G$ gilt dann

$$\begin{aligned} [v \circ u](a) * [v \circ u](b) &= v(u(a)) * v(u(b)) = v(u(a) \odot u(b)) \\ &= v(u(a \cdot b)) = [v \circ u](a \cdot b). \end{aligned}$$

Wir zeigen 2.) Sei $u : G \to H$ ein Isomorphismus. Weil u ein Isomorphismus ist, gilt für alle $c, d \in H$:

$$c \odot d = u(u^{-1}(c)) \odot u(u^{-1}(d)) = u(u^{-1}(c) \cdot u^{-1}(d)).$$

Weil u injektiv ist, folgt daraus

$$u^{-1}(c \odot d) = u^{-1}(c) \cdot u^{-1}(d).$$

Die Behauptung (3) folgt unmittelbar aus (1) und (2). (Überzeugen Sie sich selbst!) □

Die folgende Bemerkung ist nur ein Beispiel dafür, dass isomorphe Gruppen alle Eigenschaften gemeinsam haben, die sich nur mit Hilfe der Verknüpfung ausdrücken lassen.

Bemerkung 2.2.20. Sind zwei Gruppen $(G, \circ)$ und $(H, \odot)$ isomorph, und ist eine davon kommutativ, so ist auch die andere kommutativ.

Beweis. Seien $(G, \circ)$ und $(H, \odot)$ isomorph durch den Isomorphismus $f : G \to H$, und sei $(H, \odot)$ kommutativ. Seien $x, y \in G$. Weil $(H, \odot)$ kommutativ und $f : G \to H$ ein Homomorphismus ist, gilt

$$f(x \circ y) = f(x) \odot f(y) = f(y) \odot f(x) = f(y \circ x)\,.$$

Weil f injektiv ist, folgt daraus $x \circ y = y \circ x$.
Weil auch ein Isomorphismus $f^{-1} : H \to G$ existiert, gilt auch die umgekehrte Implikation: Wenn $(G, \circ)$ kommutativ ist, dann ist auch $(H, \odot)$ kommutativ. □

Satz 2.2.21. *Seien $(G, \circ)$ und $(H, \odot)$ zwei Gruppen, und sei $f : G \to H$ ein Gruppenhomomorphismus. Seien e_G und e_H die neutralen Elemente in G bzw. H. Sei $a \in G$.*

Dann gilt:

1) $f(e_G) = e_H$.
2) $f(a^{-1}) = (f(a))^{-1}$.

Beweis. 1) Weil f ein Homomorphismus ist, gilt

$$f(e_G) \odot f(e_G) = f(e_G \circ e_G) = f(e_G) = f(e_G) \odot e_H.$$

Nach der Kürzungsregel folgt $f(e_G) = e_H$.
2.) Es ist

$$f(a) \odot f(a^{-1}) = f(a \circ a^{-1}) = f(e_G) = e_H = f(a) \odot (f(a))^{-1}.$$

Nach der Kürzungsregel folgt $f(a^{-1}) = (f(a))^{-1}$. □

2.2.3. *Untergruppen.*

Definition 2.2.22. Sei $(G, \circ)$ eine Gruppe und $U \subset G$ eine nichtleere Teilmenge. U heißt eine Untergruppe von G wenn gilt:

1) $(\forall a, b \in U)\ a \circ b \in U$,
2) $(\forall a \in U)\ a^{-1} \in U$.

Bemerkung 2.2.23. Sei $(G, \circ)$ eine Gruppe und $U \subset G$ eine Untergruppe. Ist $e \in G$ das neutrale Element, so ist auch $e \in U$. Ferner ist $(U, \circ)$ eine Gruppe.

Satz 2.2.24. *Seien $(G, \circ)$ und $(H, \odot)$ zwei Gruppen, und sei $f : G \to H$ ein Gruppenhomomorphismus. Seien $U \subset G$ und $V \subset H$ zwei Untergruppen von G bzw. H. Dann gilt:*

1) *$f(U)$ ist eine Untergruppe von H.*
2) *$f^{-1}(V)$ ist eine Untergruppe von G.*

Beweis. 1.) Seien $c, d \in f(U)$, d.h., es gibt $a, b \in U$ mit $f(a) = c$, $f(b) = d$. Es ist dann $a^{-1} \in U$, und folglich $c^{-1} = f(a^{-1}) \in f(U)$. Es ist $a \circ b \in U$, folglich ist $c \odot d = f(a \circ b) \in f(U)$.
2.) Seien $a, b \in f^{-1}(V)$, d.h. $f(a) \in V$, $f(b) \in V$. Weil V Untergruppe ist, ist auch $f(a^{-1}) = (f(a))^{-1} \in V$, daher ist $a^{-1} \in f^{-1}(V)$. Ferner ist $f(a \circ b) = f(a) \odot f(b) \in V$, also $a \circ b \in f^{-1}(V)$. □

Korollar 2.2.25. *Seien $(G, \circ)$ und $(H, \odot)$ zwei Gruppen, und sei $f : G \to H$ ein Gruppenhomomorphismus. Sei e_H das neutrale Element in H. Dann ist $\{x \in G \mid f(x) = e_H\}$ eine Untergruppe von G.*

Beweis. Es lässt sich leicht zeigen (können Sie es?), dass $\{e_H\}$ eine Untergruppe von H ist. Folglich ist $f^{-1}(\{e_H\})$ eine Untergruppe von G. □

2.3. **Ringe, Restklassen.**

2.3.1. *Ringe.*

Definition 2.3.1. Sei R eine nichtleere Menge mit zwei inneren Verknüpfungen $+$ und $\cdot$, welche folgende Bedingungen erfüllen:

1) $(R, +)$ ist eine kommutative Gruppe.
2) Die Verknüpfung $\cdot$ ist assoziativ.
3) Für alle $a, b, c \in R$ gelten die zwei Distributivgesetze:

$$\begin{aligned} c \cdot (a + b) &= (c \cdot a) + (c \cdot b) \\ (a + b) \cdot c &= (a \cdot c) + (b \cdot c). \end{aligned}$$

Dann heißt das Tripel $(R, +, \cdot)$ ein Ring.
Gilt zusätzlich, dass $\cdot$ kommutativ ist, dann heißt der Ring ein kommutativer Ring.
Besitzt $(R, \cdot)$ ein neutrales Element, so heißt $(R, +, \cdot)$ ein Ring mit Eins.

Schreibweise 2.3.2.

- In einem Ring $(R, +, \cdot)$ heißt $+$ die Ringaddition und $\cdot$ die Ringmultiplikation.
- Das neutrale Element der Gruppe $(R, +)$ bezeichnen wir mit 0 und nennen es auch Nullelement.
- Das neutrale Element von $(R, \cdot)$ (falls es eines gibt) bezeichnen wir mit 1 und nennen es auch Einselement.
- Das inverse Element zu $x \in R$ bezüglich $+$ bezeichnen wir mit $-x$, und verwenden die Notation

$$a - b := a + (-b).$$

Satz 2.3.3. *In einem Ring $(R, +, \cdot)$ mit Nullelement 0 gilt für alle $x \in R$:*

$$0 \cdot x = x \cdot 0 = 0.$$

Beweis. Es gilt

$$0 + 0 \cdot x = 0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x.$$

Nun verwenden wir die Kürzungsregel bezüglich $+$ und erhalten

$$0 = 0 \cdot x.$$

Ebenso geht $0 = x \cdot 0$. □

Beispiel 2.3.4. $(\mathbb{R}, +, \cdot)$, $(\mathbb{Z}, +, \cdot)$ sind Ringe, aber nicht $(\mathbb{N}, +, \cdot)$, weil $(\mathbb{N}, +)$ keine Gruppe ist.

Ein besonders wichtiger Ring ist auch der Polynomring. Wir werden diesen aber erst im zweiten Semester besprechen, denn da werden wir ihn brauchen.

2.3.2. *Restklassen.*

Definition 2.3.5.

(1) Seien $a, b \in \mathbb{Z}$. Wir sagen, a ist ein Teiler von b (Kurzschrift: $a \mid b$), genau dann, wenn es ein $z \in \mathbb{Z}$ gibt, sodass $az = b$.
(2) Sei $m \in \mathbb{N}$ (ohne Null). Dann ist $m\mathbb{Z}$ die Menge aller ganzzahligen Vielfachen von m, d.h.

$$m\mathbb{Z} := \{x \in \mathbb{Z} \,\big|\, m \mid x\}.$$

Bemerkung 2.3.6. Es ist leicht zu zeigen, dass $(m\mathbb{Z}, +, \cdot)$ ein Ring ist. Außerdem gilt[8]: Ist $p \in \mathbb{Z}$ und $x \in m\mathbb{Z}$, dann ist $px \in m\mathbb{Z}$.

[8]Ein Unterring J eines kommutativen Ringes R mit der Eigenschaft $(\forall p \in R)\, pJ \subset J$ heißt auch ein Ideal. Ideale spielen eine fundamentale Rolle in der Theorie der Ringhomomorphismen und der Teilbarkeitslehre.

Beweis. "Leicht zu zeigen ..." Können Sie's? □

Definition 2.3.7. Sei $m \in \mathbb{N}$.

(1) Ist $x \in \mathbb{Z}$, so bezeichnen wir mit $[x]_{\mathrm{mod}\, m}$ oder kurz $[x]$ die Restklasse von x modulo m:

$$[x]_{\mathrm{mod}\, m} := \{y \in \mathbb{Z} \mid \ y - x \in m\mathbb{Z}\}\,.$$

(2) Die Menge aller Restklassen modulo m bezeichnen wir mit $\mathbb{Z}/m\mathbb{Z}$.

Beispiel 2.3.8. Die Menge $\mathbb{Z}/4\mathbb{Z}$ der Restklassen modulo 4 besteht aus folgenden Klassen:

$$\begin{aligned}
[0] &= \{0, -4, 4, -8, 8, -12, 12, -16, 16, \cdots\}\,,\\
[1] &= \{1, -3, 5, -7, 9, -11, 13, -15, 17, \cdots\}\,,\\
[2] &= \{2, -2, 6, -6, 10, -10, 14, -14, 18, \cdots\}\,,\\
[3] &= \{3, -1, 7, -5, 11, -9, 15, -13, 19, \cdots\}\,.
\end{aligned}$$

Beachten Sie, dass damit alle Restklassen aufgezählt sind. Die Restklasse von 4 ist wieder die Restklasse von 0, usw.

Bemerkung 2.3.9. Tatsächlich gibt es in der Algebra viele Arten der Restklassenbildung, nicht nur für ganze Zahlen. Das zunächst seltsam anmutende Divisionszeichen im Symbol $\mathbb{Z}/m\mathbb{Z}$ hat sich für Restklassenbildung eingebürgert. Über dem "Bruchstrich" steht die Struktur, die in Restklassen unterteilt wird. Unter dem "Bruchstrich" steht die Restklasse der 0.

Definition 2.3.10. Sei M eine Menge und $\mathcal{U}$ eine Menge von Teilmengen von M. Wir sagen, $\mathcal{U}$ ist eine Klasseneinteilung oder Partition auf M, wenn folgende Eigenschaften gelten:

(1) $M = \bigcup_{A \in \mathcal{U}} A$, d.h., jedes Element aus M ist in mindestens einer der Mengen aus $\mathcal{U}$ enthalten, und die Mengen aus $\mathcal{U}$ enthalten nur Elemente aus M.
(2) Je zwei verschiedene Mengen aus $\mathcal{U}$ sind disjunkt, d.h., kein Element aus M ist in mehr als einer der Mengen aus $\mathcal{U}$ enthalten.
(3) $\emptyset \notin \mathcal{U}$, das heißt, keine Menge aus $\mathcal{U}$ ist leer.

Bemerkung 2.3.11. Sei $m \in \mathbb{N}$. Die Menge $\mathbb{Z}/m\mathbb{Z}$ der Restklassen modulo m ist eine Partition auf $\mathbb{Z}$.

Beweis. Aussage 1: Zu jedem $x \in \mathbb{Z}$ gehört nach Definition 2.3.7 eine Restklasse $[x]$. Nun ist $x - x = 0 \in m\mathbb{Z}$, also ist $x \in [x]$. Also ist jede ganze Zahl in einer Restklasse enthalten.
Aussage 2: Sei $[x] \cap [y] \neq \emptyset$, also können wir ein $z \in [x] \cap [y]$ wählen. Es sind also $z - x$ und $z - y$ Vielfache von m, d.h., es gibt $a, b \in \mathbb{Z}$ mit $z - x = am$, $z - y = bm$. Sei nun $t \in [x]$, also $t - x = cm$. Dann ist

$$t - y = (t - x) + (x - z) + (z - y) = cm - am + bm = (c - a + b)m$$

ein Vielfaches von m. Also gilt $t \in [y]$. Wir haben gezeigt: $[x] \subset [y]$. Ebenso zeigt man $[y] \subset [x]$ durch Tauschen der Rollen von x und y. Daher ist $[x] = [y]$. Restklassen, die nicht disjunkt sind, sind in Wirklichkeit dieselbe Restklasse.
Aussage 3: Es ist $z \in [z]$, daher ist keine Restklasse leer. □

Lemma 2.3.12. *Sei $m \in \mathbb{N}$, seien $x, y, \tilde{x}, \tilde{y} \in \mathbb{Z}$, dabei sei $\tilde{x} \in [x]$ und $\tilde{y} \in [y]$. Dann gilt:*

$$[\tilde{x} + \tilde{y}] = [x + y]\,, \quad [\tilde{x} - \tilde{y}] = [x - y] \quad \textit{und} \quad [\tilde{x} \cdot \tilde{y}] = [x \cdot y]\,.$$

Beweis. Es gibt $a, b \in \mathbb{Z}$ sodass $\tilde{x} - x = am$ und $\tilde{y} - y = bm$. Dann ist

$$(\tilde{y} + \tilde{x}) - (y + x) = (\tilde{y} - y) + (\tilde{x} - x) = bm + am = (b + a)m\,,$$
$$(\tilde{y}\tilde{x}) - (yx) = (\tilde{y} - y)\tilde{x} + y(\tilde{x} - x) = bm\tilde{x} + yam = (b\tilde{x} + ay)m\,.$$

Also ist $\tilde{x} + \tilde{y} \in [x + y]$, und da Klassen gleich sind, wenn sie nicht disjunkt sind, ist $[\tilde{x} + \tilde{y}] = [x + y]$. Ebenso zeigt man $[\tilde{x} - \tilde{y}] = [x - y]$. Schließlich ist nach der obigen Rechnung $\tilde{x}\tilde{y} \in [xy]$, und folglich $[\tilde{x}\tilde{y}] = [xy]$. □

Definition 2.3.13. Sei $m \in \mathbb{N}$, und $[x], [y] \in \mathbb{Z}/m\mathbb{Z}$. Wir definieren:

$$\begin{aligned} [x] + [y] &:= [x + y]\,, \\ [x] - [y] &:= [x - y]\,, \\ [x][y] &:= [xy]\,. \end{aligned}$$

Erklärung: Man darf definieren, was man will, aber man muss sicher stellen, dass das definierte Objekt eindeutig bestimmt ist. Die Schwierigkeit in dieser Definition könnte daran liegen, dass verschiedene ganze Zahlen dieselbe Restklasse ergeben können. Es sei also $[x] = [\tilde{x}]$ und $[y] = [\tilde{y}]$. Wie soll $[x] + [y] = [\tilde{x}] + [\tilde{y}]$ definiert werden, durch $[x+y]$ oder durch $[\tilde{x}+\tilde{y}]$? Lemma 2.3.12 sagt, dass beide Definitionen dieselbe Restklasse ergeben, also ist $[x + y]$ eindeutig festgelegt, wir sagen auch, wohldefiniert. Dasselbe gilt für $[x - y]$ und $[xy]$. □

Satz 2.3.14. *Sei $m \in \mathbb{N}$. Dann ist $\mathbb{Z}/m\mathbb{Z}$ ein kommutativer Ring mit Eins. Dabei sind $[0]$ das Nullelement und $[1]$ das Einselement, und das Inverse in der additiven Gruppe ist $-[x] = [-x]$.*

Beweis. Wir müssten nun Rechenregel für Rechenregel des Ringes nachprüfen. Sie übertragen sich direkt von den Rechenregeln der ganzen Zahlen. Wir zeigen als Beispiel nur das Distributivgesetz:

$$([x] + [y])[z] = [x + y][z] = [(x + y)z] = [xz + yz] = [xz] + [yz] = [x][z] + [y][z]\,.$$

□

Definition 2.3.15. Sei $(R, +, \cdot)$ ein kommutativer Ring und $x \in R$. Dann heißt x ein Nullteiler, wenn $x \neq 0$ und wenn es ein $y \neq 0$ gibt, so dass $xy = 0$.
Ein Ring, der keine Nullteiler enthält, heißt nullteilerfrei.

Satz 2.3.16. *Sei $m \in \mathbb{Z}$. Dann ist $\mathbb{Z}/m\mathbb{Z}$ genau dann nullteilerfrei, wenn m eine Primzahl ist.*

Beweisskizze: Wir müssen uns hier leider auf Vorwissen über ganze Zahlen berufen, das wir vielleicht noch aus der Schule haben. Einen sauberen Beweis bekommen Sie in einer Vorlesung über elementare Algebra oder über Zahlentheorie vorgelegt. Sei zunächst m eine Primzahl, das ist eine Zahl, die sich nur auf triviale Art in ein Produkt zerlegen lässt: $m = 1.m = m.1$. Wenn für $x, y \in \mathbb{Z}$ das Produkt xy durch m teilbar ist, muss m mindestens eine der beiden Zahlen, x oder y, teilen. Es gilt also, in anderen Worten: Wenn $[x][y] = [0]$, dann ist $[x] = 0$ oder $[y] = 0$.
Nun sei m keine Primzahl, also sei $m = xy$ mit $1 < x < m$, $1 < y < m$. Dann sind $[x] \neq [0]$ und $[y] \neq [0]$, aber $[x][y] = [m] = [0]$. □

Beispiel 2.3.17. Wir zeigen die Verknüpfungstafeln für $\mathbb{Z}/6\mathbb{Z}$. Beachten Sie die Nullteiler [2], [3] und [4]. Nur die Restklassen von Elementen, die mit 6 keinen gemeinsamen Teiler haben, haben inverse Elemente bezüglich der Multiplikation.

+	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[1]	[2]	[3]	[4]	[5]
[1]	[1]	[2]	[3]	[4]	[5]	[0]
[2]	[2]	[3]	[4]	[5]	[0]	[1]
[3]	[3]	[4]	[5]	[0]	[1]	[2]
[4]	[4]	[5]	[0]	[1]	[2]	[3]
[5]	[5]	[0]	[1]	[2]	[3]	[4]

·	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]	[5]
[2]	[0]	[2]	[4]	[0]	[2]	[4]
[3]	[0]	[3]	[0]	[3]	[0]	[3]
[4]	[0]	[4]	[2]	[0]	[4]	[2]
[5]	[0]	[5]	[4]	[3]	[2]	[1]

2.4. **Körper, komplexe Zahlen.**

2.4.1. *Körper.*

Definition 2.4.1. Sei K eine nichtleere Menge mit zwei inneren Verknüpfungen $+$ und $\cdot$, welche folgende Bedingungen erfüllen:

1) $(K, +)$ ist eine kommutative Gruppe.
2) $(K \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe.
3) Es gelten die Distributivgesetze: Für alle $a, b, c \in K$ gilt:

$$\begin{aligned} c \cdot (a + b) &= (c \cdot a) + (c \cdot b) \\ (a + b) \cdot c &= (a \cdot c) + (b \cdot c). \end{aligned}$$

Dann heißt das Tripel $(K, +, \cdot)$ ein Körper.

Bemerkung 2.4.2. Insbesondere ist also jeder Körper ein kommutativer Ring mit Eins. Es gibt aber kommutative Ringe mit Eins, die keine Körper sind (zum Beispiel $(\mathbb{Z}, +, \cdot)$).

Schreibweise 2.4.3. In einem Körper bezeichnen wir das inverse Element zu $x \in K \setminus \{0\}$ mit x^{-1} oder $\frac{1}{x}$. Wir verwenden auch die Bruchschreibweise für $a \in K$, $b \in K \setminus \{0\}$:

$$\frac{a}{b} := ab^{-1}.$$

Bemerkung 2.4.4. In einem Ring $(R, +, \cdot)$ mit Eins, der mindestens 2 Elemente enthält kann das Nullelement kein inverses Element bezüglich der Multiplikation haben.

Beweis. Zunächst kann 0 nicht das Einselement sein, denn es gibt mindestens ein Element $a \neq 0$. Es gilt aber $1 \cdot a = a \neq 0 = 0 \cdot a$. Hätte 0 bezüglich der Multiplikation ein inverses Element x, so wäre $1 = 0 \cdot x$ und zugleich $0 \cdot x = 0$, dann wäre aber $1 = 0$, was wir eben ausgeschlossen haben. □

Bemerkung 2.4.5. Körper sind immer nullteilerfrei.

Beweis. Sei $(K, +, \cdot)$ ein Körper, $x, y \in K$ mit $xy = 0$. Sei $x \neq 0$. Dann ist

$$y = 1y = (x^{-1}x)y = x^{-1}(xy) = x^{-1}0 = 0\,,$$

also ist mindestens einer der beiden Faktoren Null. □

In einem Körper gelten die Rechenregeln bezüglich Addition und Multiplikation, welche wir von $(\mathbb{R}, +, \cdot)$ gewöhnt sind. Wenn wir aber nur wissen, dass $(K, +, \cdot)$ irgendein Körper ist, haben wir die folgenden Begriffe und ihre Gesetzmäßigkeiten nicht zur Verfügung: Ordnungsrelation $x < y$, Absolutbetrag $|x|$, Konvergenz von Folgen.

Beispiel 2.4.6. Bekannte Körper von Zahlen sind $(\mathbb{R}, +, \cdot)$ und $(\mathbb{Q}, +, \cdot)$. Die algebraische Struktur $(\mathbb{Z}, +, \cdot)$ ist zwar ein nullteilerfreier Ring mit Eins, aber kein Körper, weil es im Allgemeinen keine inversen Elemente für die Multiplikation gibt, man kann nicht dividieren.

Satz 2.4.7. *Sei $m \in \mathbb{N}$. Der Restklassenring $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ ist genau dann ein Körper, falls m eine Primzahl ist.*

Beweis. Ist m keine Primzahl, so hat $\mathbb{Z}/m\mathbb{Z}$ Nullteiler und kann daher kein Körper sein. Ist aber m eine Primzahl, so hat $\mathbb{Z}/m\mathbb{Z}$ keine Nullteiler. Sei $[x] \neq [0] \in \mathbb{Z}/m\mathbb{Z}$. Wir zeigen zunächst, dass die Abbildung

$$f_x : \begin{cases} \mathbb{Z}/m\mathbb{Z} & \to \mathbb{Z}/m\mathbb{Z} \\ [a] & \mapsto [a][x] \end{cases}$$

injektiv ist. Sei $f_x([a]) = f_x([b])$, also $[a][x] = [b][x]$. Dann ist $([a]-[b])[x] = [0]$, und weil $\mathbb{Z}/m\mathbb{Z}$ keine Nullteiler hat, ist $[a]-[b] = [0]$, also $[a] = [b]$. Also ist f_x injektiv. Nun ist $\mathbb{Z}/m\mathbb{Z}$ eine endliche Menge mit m Elementen. Weil f_x injektiv ist, bildet es jedes Element aus $\mathbb{Z}/m\mathbb{Z}$ auf ein anderes Element ab, es müssen also mindestens m verschiedene Elemente im Bild von f_x sein. Daher füllt das Bild von f_x ganz $\mathbb{Z}/m\mathbb{Z}$ aus, und f_x ist surjektiv.
Wir wählen nun $[y]$ so dass $f_x([y]) = [1]$, das heißt aber, $[x][y] = 1$ und wir haben das inverse Element zu $[x]$ gefunden. □

Bemerkung 2.4.8.

(1) In Wirklichkeit verwendet der obige Beweis nur, dass $\mathbb{Z}/m\mathbb{Z}$ nur endlich viele Elemente hat und nullteilerfrei ist, wenn m eine Primzahl ist. Wir haben gezeigt: Ein endlicher, kommutativer, nullteilerfreier Ring ist ein Körper.
(2) Die endlichen Körper sehen im ersten Augenblick wie eine Zahlenspielerei aus. Sie haben sehr große Bedeutung in der Kryptographie (Verschlüsselung von Daten) erlangt.

Beispiel 2.4.9. Wir konstruieren den kleinsten Körper von reellen Zahlen, der alle rationalen Zahlen aber auch $\sqrt{2}$ enthält:
Sei $K := \{x + y\sqrt{2} \mid x, y \in \mathbb{Q}\} \subset \mathbb{R}$. Mit der üblichen Addition und Multiplikation ist $(K, +, \cdot)$ ein Körper.

Beweis. Wir benötigen zunächst, dass die Verknüpfungen $+$ und $\cdot$ auf K innere Verknüpfungen sind. Betrachten wir $a + b\sqrt{2} \in K$ und $c + d\sqrt{2} \in K$, dann ist (unter der Ausnützung der Rechenregeln in $\mathbb{R}$)

$$\begin{aligned} (a + b\sqrt{2}) + (c + d\sqrt{2}) &= (a + c) + (b + d)\sqrt{2} \in K, \\ (a + b\sqrt{2}) \cdot (c + d\sqrt{2}) &= (ac + 2bd) + (ad + bc)\sqrt{2} \in K. \end{aligned}$$

Assoziativitäts-, Kommutativitäts- und Distributivgesetze erbt K von den reellen Zahlen. Die Elemente $0 = 0+0\sqrt{2}$ und $1 = 1+0\sqrt{2}$ sind das Null- und Einselement in $\mathbb{R}$ und daher erst recht in K. Es fehlt noch die Existenz der inversen Elemente für Addition und Multiplikation. Sei $a+b\sqrt{2} \neq 0+0\sqrt{2}$. Es ist leicht nachzuprüfen:

$$\begin{aligned} &[a + b\sqrt{2}] + [(-a) + (-b)\sqrt{2}] = 0, \\ &[a + b\sqrt{2}] \cdot [\frac{a}{a^2 - 2b^2} - \frac{b}{a^2 - 2b^2}\sqrt{2}] = 1 + 0\sqrt{2}. \end{aligned}$$

Beachten Sie, dass die Brüche in der zweiten Zeile oben wohldefiniert sind, denn $a^2 - 2b^2$ kann nur null sein, wenn sowohl a als auch b gleich Null sind. Andernfalls wäre $2 = (a/b)^2$, aber bekanntlich ist $\sqrt{2}$ keine rationale Zahl. □

In der Algebra spielen solche "Erweiterungen" des Körpers der rationalen Zahlen eine große Rolle. Man kann sie zum Beispiel dazu heranziehen, zu untersuchen, welche Vielecke man exakt mit Zirkel und Lineal konstruieren kann, oder welche Gleichungen mit Ausdrücken von Summen, Multiplikation, Potenz, Wurzeln exakt lösbar sind.

2.4.2. *Komplexe Zahlen.*

Heuristik 2.4.10. Wenn es eine Zahl i mit der Eigenschaft $i^2 = -1$ gäbe, wie müssten die Multiplikation und Division

$$(a+bi)(c+di), \quad \frac{a+bi}{c+di}$$

gehen, wenn a, b, c, d reelle Zahlen wären?

$$\begin{aligned}(a+bi)(c+di) &= ac+bci+adi+i^2bd = (ac-bd)+(bc+ad)i,\\ \frac{a+bi}{c+di} &= \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{(ac-bdi^2)+(bc-ad)i}{c^2-d^2i^2}\\ &= \frac{(ac+bd)+(bc-ad)i}{c^2+d^2}.\end{aligned}$$

Hat man einmal gefunden, wie die Rechenregeln lauten müssten, zäumt man das Pferd vom Schwanz auf: Man definiert sich die komplexen Zahlen $a+ib$ einfach als Paare (a,b) von reellen Zahlen, und führt auf der Menge $\mathbb{R}\times\mathbb{R}$ die passenden Rechenregeln ein. Dann zeigt man, dass das Ganze eine sinnvolle Definition ist, indem man zeigt, dass man die reellen Zahlen als Teilmenge der komplexen Zahlen auffassen kann, und dass die komplexen Zahlen einen Körper bilden.

Definition 2.4.11. Die komplexen Zahlen $(\mathbb{C}, +, \cdot)$ sind die Menge $\mathbb{C} = \mathbb{R}\times\mathbb{R}$ mit den folgenden inneren Verknüpfungen:

$$\begin{aligned}(a,b)+(c,d) &:= (a+c, b+d),\\ (a,b)\cdot(c,d) &:= (ac-bd, ad+bc).\end{aligned}$$

Statt (a,b) schreiben wir auch $a+ib$, insbesondere schreiben wir $i := (0,1)$. Durch die Abbildung

$$j : \begin{cases}\mathbb{R} & \to \mathbb{C}\\ a & \mapsto (a,0)\end{cases}$$

bilden wir die reellen Zahlen injektiv in die komplexen Zahlen ab.

Satz 2.4.12. *Die komplexen Zahlen $(\mathbb{C}, +, \cdot)$ bilden einen Körper mit dem Nullelement $0+0i$ und dem Einselement $1+0i$. Es gilt insbesondere*

$$\begin{aligned}(a+bi)-(c+di) &= (a-c)+(b-d)i,\\ \frac{a+bi}{c+di} &= \frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}i.\end{aligned}$$

Die Abbildung j ist ein injektiver Körperhomomorphismus im folgenden Sinn: Sind $x, y \in \mathbb{R}$, so gilt

$$\begin{aligned}j(x+y) &= j(x)+j(y),\\ j(xy) &= j(x)\cdot j(y).\end{aligned}$$

Beweis. Der Beweis ist im Detail mühsam und erfolgt durch einfaches Nachrechnen aller erforderlichen Rechenregeln.
Beispielsweise die Assoziativität der Multiplikation. (Beachten Sie, dass $a+bi$ in Wirklichkeit das Paar (a,b) bezeichnet, und wir zwar wissen, dass die Multiplikation der reellen Zahlen assoziativ ist, aber ob die Multiplikation auch assoziativ bleibt, wenn die nicht reelle Zahl i vorkommt, ist ja erst zu beweisen!)

$$\begin{aligned}[(a,b)\cdot(c,d)]\cdot(e,f) &= (ac-bd, ad+bc)\cdot(e,f)\\ &= ((ac-bd)e-(ad+bc)f\,,\,(ac-bd)f+(ad+bc)e)\\ &= (ace-bde-adf-bcf\,,\,acf-bdf+ade+bce),\\ (a,b)\cdot[(c,d)\cdot(e,f)] &= (a,b)\cdot(ce-df, cf+de)\\ &= (a(ce-df)-b(cf+de)\,,\,a(cf+de)+b(ce-df))\\ &= (ace-adf-bcf-bde\,,\,acf+ade+bce-bdf).\end{aligned}$$

In beiden Multiplikationen kommt dasselbe heraus, daher gilt das Assoziativgesetz in $(\mathbb{C}, \cdot)$. Ebenso sind die Kommutativität von Addition und Multiplikation, die Assoziativität der Addition und die Distributivgesetze zu beweisen.
Beispielsweise zeigen wir, dass das neutrale Element und das inverse Element zu $a + ib$ bezüglich der Multiplikation die Elemente $1 + 0i$ und $\frac{a}{a^2+b^2} - \frac{b}{a^2+b^2} i$ sind:

$$\begin{aligned} (1,0) \cdot (x,y) &= (1x - 0y, 1y + 0x) = (x,y), \\ (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}) \cdot (a,b) &= (\frac{a}{a^2+b^2}\, a + \frac{b}{a^2+b^2}\, b \ , \ \frac{a}{a^2+b^2}\, b + \frac{-b}{a^2+b^2}\, a) \\ &= (1,0)\,. \end{aligned}$$

Die anderen Behauptungen über neutrale, inverse Elemente, Subtraktion und Division folgen ebenso.
Die Injektivität von j ist klar, denn

$$j(x) = j(y) \ \Rightarrow \ (x,0) = (y,0) \ \Rightarrow \ x = y.$$

Die Homomorphieeigenschaften prüft man nach, z.B. für die Multiplikation

$$j(x) \cdot j(y) = (x,0) \cdot (y,0) = (x.y - 0.0, x.0 + 0.y) = (xy, 0) = j(xy).$$

□

Schreibweise 2.4.13. Wir identifizieren die Menge der reellen Zahlen mit der Teilmenge $\{(a,0) \in \mathbb{C} \mid, a \in \mathbb{R}\}$, indem wir jedes $x \in \mathbb{R}$ mit $j(x)$ identifizieren. Da j injektiv ist, werden verschiedene reelle Zahlen mit verschiedenen komplexen Zahlen identifiziert, da j ein Körperhomomorphismus ist, bleiben die Rechenregeln der reellen Zahlen auch bei dieser Identifikation erhalten.

Künftig schreiben wir also unbesorgt $a + ib$ statt (a,b) und können auf die formale Definition als Paare und den Einbettungshomomorphismus j vergessen: Sie wurden nur gebraucht, um $\mathbb{C}$ auf eine solide formale Basis zu stellen.

Definition 2.4.14. Sei $z = a + ib$ eine komplexe Zahl: Wir definieren

- den Realteil von z durch $\Re(z) := a$,
- den Imaginärteil von z durch $\Im(z) := b$,
- Betrag $|z| := r$ und Argument $\arg(z) := \phi$ von z als jene nichtnegative reelle Zahl r und jenen Winkel ϕ für die gilt:

  $$z = r(\cos(\phi) + i \sin(\phi)).$$

  (In diesem Zusammenhang geben wir Winkel immer im Bogenmaß an: 2π entspricht einem Winkel von 360°.)
- die konjugiert komplexe Zahl zu z durch $\overline{z} = a - bi$.

Wir identifizieren auch die komplexen Zahlen zur geometrischen Interpretation mit $\mathbb{R}^2$:

$$\begin{cases} \mathbb{C} & \to \mathbb{R}^2, \\ a + ib & \mapsto \begin{pmatrix} a \\ b \end{pmatrix}. \end{cases}$$

Diese Definition wird mit einer grafischen Darstellung der komplexen Zahlen untermauert.

Bemerkung 2.4.15.

- Der Betrag einer komplexen Zahl $a + bi$ errechnet sich durch

  $$|a + bi| = \sqrt{a^2 + b^2}.$$
- Es gilt $|z|^2 = z\overline{z}$. Insbesondere ist $z\overline{z}$ immer reell und nichtnegativ, und Null nur dann, wenn $z = 0$ ist.
- Die Addition von komplexen Zahlen stellt sich in der Zahlenebene als Vektoraddition dar.

- Bei der Multiplikation multiplizieren sich die Beträge, und addieren sich die Argumente.
- Bei der Division dividieren sich die Beträge und subtrahieren sich die Argumente, insbesondere ist $\arg(\frac{1}{z}) = -\arg(z)$.

Beweis. Für den Betrag: Wir verwenden $\cos^2(\phi)+\sin^2(\phi) = 1$. Ist also $a = r\cos(\phi)$, $b = r\sin(\phi)$, dann gilt

$$a^2 + b^2 = r^2\cos^2(\phi) + r^2\sin^2(\phi) = r^2.$$

Offensichtlich ist

$$(a + bi) \cdot \overline{(a + bi)} = a^2 + b^2 = |a + bi|^2.$$

Die geometrische Deutung der Addition ist sehr einfach. Schwieriger ist die Multiplikation. Wir müssen dazu die Additionstheoreme der Winkelfunktionen verwenden:

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

Betrachten wir also das Produkt zweier komplexer Zahlen $u = r\cos(\phi) + r\sin(\phi)i$, $v = s\cos(\psi) + s\sin(\psi)i$. Das Produkt ist

$$\begin{aligned} uv &= [r\cos(\phi)s\cos(\psi) - r\sin(\phi)s\sin(\psi)] + [r\cos(\phi)s\sin(\psi) + r\sin(\phi)s\cos(\psi)]i \\ &= rs[\cos(\phi)\cos(\psi) - \sin(\phi)\sin(\psi)] + rs[\cos(\phi)\sin(\psi) + \sin(\phi)\cos(\psi)]i \\ &= rs\cos(\phi + \psi) + rs\sin(\phi + \psi). \end{aligned}$$

Aus dieser Darstellung von uv folgt $|uv| = rs$ und $\arg(uv) = \phi + \psi$.
Die Folgerung für die Division ist nun offensichtlich. □

Bemerkung 2.4.16. Das Argument einer komplexen Zahl ist nur modulo 2π bestimmt, da die Winkelfunktionen periodisch sind: Für alle ganzen Zahlen $k \in \mathbb{Z}$ gilt $\cos(\phi + 2k\pi) = \cos(\phi)$, $\sin(\phi + 2k\pi) = \sin(\phi)$. Ist also $\arg(z) = \phi$, so könnte man auch z.B. schreiben $\arg(z) = \phi + 28\pi$.

Eine mathematisch saubere Definition des Arguments müsste also den Wertebereich einschränken z.B. $\arg(z) \in (-\pi, \pi]$: Diese Variante ist aber problematisch, wenn man schreiben will: $\arg(uv) = \arg(u) + \arg(v)$, weil die Summe zweier Zahlen in $(-\pi, \pi]$ nicht wieder in $(-\pi, \pi]$ liegen muss. Besser wäre, das Argument als Restklasse modulo 2π zu definieren. Wir erlauben uns statt dessen hier eine Schlamperei, aber wir ertappen uns dabei und schämen uns pflichtschuldigst — siehe später!

Satz 2.4.17. *Sei $u = r(\cos(\phi) + i\sin(\phi))$ (mit $r > 0$ und $\phi \in \mathbb{R}$) eine komplexe Zahl ungleich Null. Sei $p \in \mathbb{N}$ (hier wird $\mathbb{N}$ ohne Null verstanden, also $p \neq 0$). Die Gleichung*

$$z^p = u$$

besitzt genau p Lösungen, und zwar

$$z = r^{1/p}\left[\cos(\frac{\phi + 2k\pi}{p}) + i\sin(\frac{\phi + 2k\pi}{p})\right],$$

mit $k \in \{0, \cdots, p-1\}$.

Beweis. Ansatz: Setze $z = s(\cos(\psi)+i\sin(\psi))$. Nach den Regeln der Multiplikation (p-mal auf z angewendet) muss sich also der Betrag potenzieren, und das Argument multiplizieren:

$$r(\cos(\phi) + i\sin(\phi)) = s^p(\cos(p\psi) + i\sin(p\psi)).$$

Also $r = s^p$ oder $s = r^{1/p}$. Und $\psi = \phi/p$? Aber halt, das Argument ist ja nur eindeutig bis auf modulo 2π (siehe Bemerkung 2.4.16), also muss gelten

$$p\psi = \phi + 2k\phi \text{ für irgendein } k \in \mathbb{Z}.$$

Wir bekommen also die p verschiedenen Lösungen. □

Die folgenden wichtigen Eigenschaften führen wir nun ohne Beweis an. Zum Großteil würde man auch gewisse Grumndkenntnisse der Analysis brauchen, betrachten Sie das als einen Vorgriff, den Sie jetzt noch nicht zur Gänze verstehen können. Sie müssen das natürlich bei der Prüfung noch nicht können.

Bemerkung 2.4.18. Der Betrag erüllt die Dreiecksungleichung: Für alle $u, v \in \mathbb{C}$ gilt

$$|u + v| \leq |u| + |v|.$$

Wir werden im zweiten Semester sehen, wie man die Dreiecksungleichung im $\mathbb{R}^n$ beweist, angewandt auf die Ebene $\mathbb{R}^2$ gibt das die Dreiecksungleichung für $\mathbb{C}$. Die Dreiecksungleichung ist die Grundlage dafür, dass man $|u - v|$ als Maß für den Abstand zweier komplexer Zahlen u, v verwenden kann. Kann man erst einen Abstand erklären, kann man auch anfangen, von Grenzwerten zu sprechen. Der Absolutbetrag ist also die Grundlage für die Grenzwertrechnung in $\mathbb{C}$.

Bemerkung 2.4.19. Jedes nicht konstante Polynom über $\mathbb{R}$ hat mindestens eine komplexe Nullstelle.

Das Polynom $x^2 + 1$ hätte zum Beispiel keine reelle Nullstelle. Braucht man also zur Behandlung eines Problems die Nullstellen eines Polynoms, so muss man im Reellen Fallunterscheidungen treffen zwischen Situationen, in denen das Polynom Nullstellen hat, und dem Fall, dass es keine Nullstelle gibt. Im Komplexen gibt es immer Nullstellen, und daher keine umständlichen Fallunterscheidungen. Wir gehen auf dieses Thema im zweiten Semester noch einmal gründlich ein.

Bemerkung 2.4.20. Die Eulersche Zahl

$$e := \lim_{n\to\infty} (1 + \frac{1}{n})^n \approx 2,7183$$

läßt sich auch zu komplexen Potenzen erheben. Ist ϕ eine reelle Zahl, so ist

$$e^{i\phi} = \cos(\phi) + i \sin(\phi).$$

In der Theorie der linearen Systeme gibt es exponentielles Wachstum. Eine exponentiell wachsende oder fallende Funktion f lässt sich durch die Formel

$$f(t) = Ce^{\lambda t}$$

beschreiben. Dabei sind $C, \lambda \in \mathbb{R}$. Das Vorzeichen von λ entscheidet, ob die Funktion exponentiell steigt oder fällt.
Es gibt aber auch Systeme, welche schwingen:

$$f(t) = C\cos(\omega t) \quad \text{mit } C, \omega \in \mathbb{R}.$$

Nach der Formel oben lässt sich f als der Realteil einer Exponentialfunktion schreiben

$$f(t) = \Re(Ce^{i\omega t}).$$

Durch die komplexen Zahlen lassen sich Systeme mit exponentiellem Wachstumsverhalten und Systeme die schwingen mit derselben Theorie anfassen, der komplexen Exponentialfunktion.

Bemerkung 2.4.21. Die differenzierbaren Funktionen $f : \mathbb{C} \to \mathbb{C}$ haben eine sehr reiche Theorie mit überraschenden Ergebnissen, zum Beispiel

- lassen sie sich immer durch Potenzreihen ausdrücken,
- können sich Nullstellen nicht häufen,
- wird der maximale Betrag, den die Funktion auf einem Kreis (oder auch allgemeineren Gebieten) annimmt, immer am Rand angenommen,
- bestimmen die Werte, die sie am Rand des Kreises annimmt, bereits eindeutig die Werte, die sie im Inneren des Kreises annimmt.

Die Theorie dieser Funktionen heißt Funktionentheorie oder komplexe Analysis, und hat ungeheuer weittragende Anwendungen in Differentialgleichungen, Statistik und Wahrscheinlichkeitstheorie, Regeltechnik, Signalverarbeitung, u.a.m. Was zuerst wie ein Spleen weltfremder Mathematiker aussieht, die unbedingt die Wurzel aus -1 ziehen wollen, hat sich als eines der fruchtbarsten und erstaunlichsten Konzepte der Mathematik bewährt!

3. Vektorräume

3.1. **Definitionen und Beispiele, Unterräume.**

3.1.1. *Definition von Vektorräumen, einfache Rechenregeln.*

Definition 3.1.1. Sei $(\mathbb{K}, +, \cdot)$ ein Körper, und V eine nichtleere Menge. Gegeben seien zwei Verknüpfungen $\oplus$ und $\odot$:

$$\oplus \begin{cases} V \times V & \to V \\ (a,b) & \mapsto a \oplus b \end{cases}, \qquad \odot \begin{cases} \mathbb{K} \times V & \to V \\ (\alpha, a) & \mapsto \alpha \odot a \end{cases}$$

$(V, \oplus, \odot)$ heißt ein Vektorraum über $\mathbb{K}$, wenn folgende Gesetze gelten:

1) $(V, \oplus)$ ist eine abelsche Gruppe.
2) $$(\forall \alpha, \beta \in \mathbb{K}, v \in V)\ (\alpha\beta) \odot v = \alpha \odot (\beta \odot v).$$
3) Es gelten die Distributivgesetze (für alle $\alpha, \beta \in \mathbb{K}$, $v, w \in V$):
$$\begin{aligned} \alpha \odot (v \oplus w) &= \alpha \odot v \oplus \alpha \odot w, \\ (\alpha + \beta) \odot v &= \alpha \odot v \oplus \beta \odot v. \end{aligned}$$
4) Für das Einselement 1 des Körpers gilt
$$(\forall v \in V)\ 1 \odot v = v.$$

Schreibweise 3.1.2. Die Elemente von V heißen Vektoren, die Elemente von $\mathbb{K}$ heißen Skalare. Das neutrale Element der Gruppe $(V, \oplus)$ heißt Nullvektor, und wir bezeichnen es mit 0 — aus dem Zusammenhang muss erkannt werden, ob mit 0 der Nullvektor oder das Nullelement des Körpers gemeint sind. Das inverse Element zu $v \in V$ bezüglich $\oplus$ bezeichnen wir mit $-v$, und $v - w$ bedeutet wie üblich $v \oplus (-w)$. In der Praxis unterscheidet man auch nicht die Körper- und Vektorraumaddition und das Produkt in $\mathbb{K}$ und das Produkt von Skalaren mit Vektoren in der Schreibung, wir schreiben also statt $\oplus$ und $\odot$ einfach $+$ und $\cdot$. Aus dem Zusammenhang muss gesehen werden, welche Verknüpfung gemeint ist.
In dieser Vorlesung werden wir zumeist Vektoren mit römischen Kleinbuchstaben und Skalare mit griechischen Kleinbuchstaben bezeichnen.

Beispiel 3.1.3.

(1) Sei $\mathbb{K}$ ein Körper. Ein n-dimensionaler Spaltenvektor über $\mathbb{K}$ ist ein n-Tupel von Skalaren, als Spalte angeordnet. Mit $\mathbb{K}^n$ bezeichnen wir den Raum der n-dimensionalen Spaltenvektoren über $\mathbb{K}$:
$$\mathbb{K}^n := \left\{ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \;\middle|\; xi_i \in \mathbb{K} \text{ für } i = 1, \cdots n \right\}.$$

(2) Eine $m \times n$-Matrix über einem Körper $\mathbb{K}$ ist eine Tabelle, in der Elemente von $\mathbb{K}$ in m Zeilen und n-Spalten angeordnet sind. $\mathbb{K}^{m \times n}$, Raum der $m \times n$-Matrizen über $\mathbb{K}$.

(3) Sei $\Omega \neq \emptyset$ eine Menge und $\mathbb{K}$ ein Körper. Mit $\Omega^{\mathbb{K}}$ oder $\mathcal{F}(\Omega, \mathbb{K})$ bezeichnen wir die Menge aller Funktionen von Ω nach $\mathbb{K}$.
Für $f, g \in \mathcal{F}(\Omega, \mathbb{K})$, $\alpha, \beta \in \mathbb{K}$ definieren wir die Funktion $\alpha f + \beta g \in \mathcal{F}(\Omega, \mathbb{K})$ durch
$$[\alpha f + \beta g](x) := \alpha f(x) + \beta g(x).$$
Mit diesen Operationen ist $\mathcal{F}(\Omega, \mathbb{K})$ ein Vektorraum über $\mathbb{K}$. Der Nullvektor ist die konstante Funktion $f(x) = 0$.

(4) Zumeist verwendet man Räume von Funktionen, die noch spezielle Eigenschaften besitzen. Zum Beispiel den Raum der beschränkten Funktionen von $[0,1]$ nach $\mathbb{R}$, oder den Raum der Funktionen $f : \mathbb{R} \to \mathbb{R}$, für die $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$ ist, oder Räume von stetigen oder gar differenzierbaren Funktionen.
(5) Ist $\mathbb{K}_1$ ein Körper, der einen Körper $\mathbb{K}_2$ als Unterkörper enthält, so ist $\mathbb{K}_1$ mit den Körperverknüpfungen $+, \cdot$, als Vektorraumaddition und Multiplikation mit Skalaren aufgefasst, ein Vektorraum über $\mathbb{K}_2$. Beispielsweise ist $\mathbb{C}$ ein zweidimensionaler Vektorraum über $\mathbb{R}$.

Satz 3.1.4. *Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$. Ausnahmsweise verwenden wir zwei verschiedene Symbole für Null im Körper (0) und Nullvektor (Θ).*

1) $(\forall v \in V)\ 0v = \Theta$.
2) $(\forall \alpha \in \mathbb{K})\ \alpha\Theta = \Theta$.
3) $(\forall \alpha \in \mathbb{K},\ v \in V)\ \alpha v = \Theta \Rightarrow \alpha = 0 \vee v = \Theta$.
4) $(\forall v \in V)\ (-1)v = -v$.

Beweis. 1.)Aus $0 + 0 = 0$ folgern wir mittels Distributivität:

$$0v + 0v = (0 + 0)v = 0v = 0v + \Theta.$$

Nun liefert die Kürzungsregel $0v = \Theta$.
2.)Wir beginnen mit $\Theta + \Theta = \Theta$ und verwenden Distributivität:

$$\alpha\Theta + \alpha\Theta = \alpha(\Theta + \Theta) = \alpha\Theta = \alpha\Theta + \Theta.$$

Kürzungsregel.
3.) Sei $\alpha v = \Theta$ aber $\alpha \neq 0$. Zeige: $v = \Theta$. Da $\alpha \neq 0$, gibt es das inverse Element α^{-1}, und es ist

$$\Theta = \alpha^{-1}\Theta = \alpha^{-1}(\alpha v) = (\alpha^{-1}\alpha)v = 1v = v.$$

4.) Es gilt wegen Distributivität:

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = \Theta.$$

Also ist $(-1)v$ invers zu v bezüglich der Vektorraumaddition. □

Zur Schärfung der Beobachtung markieren wir in diesem Beweis nachträglich die Vektorraumadditionen und Produkte von Skalaren und Vektoren zum Unterschied von Summen und Produkten in $\mathbb{K}$.

3.1.2. *Unterräume.*

Definition 3.1.5. Sei $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$ und W eine Teilmenge von V. Die Menge W heißt ein Unterraum (oder Teilraum) von V, wenn folgende Gesetze gelten:

1) $W \neq \emptyset$.
2) $(\forall v, w \in W)\ v + w \in W$.
3) $(\forall \alpha \in \mathbb{K},\ v \in W)\ \alpha v \in W$.

Nicht verwechseln mit affinen Unterräumen, von denen wir in der formlosen Einführung geredet haben. Jeder Unterraum ist ein affiner Unterraum, aber nicht jeder affine Unterraum ist ein Unterraum. Klingt verkehrt, ist aber so.

In $\mathbb{R}^n$ hatten wir früher eine andere Definition von Unterraum als lineare Hülle. Wir werden später darauf zurückkommen, dass sich das mit der neuen Definition verträgt.

Satz 3.1.6. *Eine nichtleere Teilmenge W eines Vektorraumes $(V, +, \cdot)$ über $\mathbb{K}$ ist genau dann ein Unterraum wenn gilt:*

$$(\forall \alpha, \beta \in \mathbb{K},\ v, w \in W)\ \alpha v + \beta w \in W. \tag{3.1.1}$$

Beweis. Beweis einer Äquivalenz, zwei Richtungen. - Könnte auch als Übung gegeben werden!
Es gelte (3.1.1). Wir zeigen dass W ein Unterraum ist. Nach Voraussetzung ist W nichtleer. Seien $v, w \in W$, dann gilt

$$v + w = 1v + 1w \in W$$

wegen (3.1.1). Ist $\alpha \in \mathbb{K}$ und $v \in W$, so gilt

$$\alpha v = \alpha v + 0v \in W.$$

Sei nun W ein Unterraum von V, wir zeigen (3.1.1): Seien $\alpha, \beta \in \mathbb{K}$, $v, w \in W$. Zunächst sind wegen Definition 3.1.5(3) $\alpha v \in W$, $\beta w \in W$. Wegen Definition 3.1.5(2) folgt nun $\alpha v + \beta w \in W$. □

Satz 3.1.7. *Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$, sei W ein Unterraum von V. Dann gilt:*

1) $(\forall v \in W)\ -v \in W$.
2) *Der Nullvektor* $0 \in W$.
3) $(W, +, \cdot)$ *ist ein Vektorraum über* $\mathbb{K}$.

Beweis. 1) folgt aus $-v = (-1)v$.
2) Nach Voraussetzung ist $W \neq \emptyset$, wähle ein beliebiges $v \in W$. Es ist dann der Nullvektor $0 = 0v \in W$.
3)Nullvektor und inverses Element sind in W wegen Punkten (1) und (2). Nach Voraussetzung ist $+$ eine innere Verknüpfung in W, und die Multiplikation $\cdot$ zwischen Skalar und Vektor bildet $\mathbb{K} \times W$ nach W ab. Alle Rechengesetze gelten für alle Vektoren in V, daher erst recht für alle Vektoren in der Teilmenge W. □

Satz 3.1.8. *Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$. Sei I eine Indexmenge (endlich oder unendlich), und für jedes $i \in I$ sei U_i ein Unterraum von V. Dann ist auch der Durchschnitt*

$$\bigcap_{i \in I} U_i$$

ein Unterraum.

Beweis. Der Durchschnitt ist nicht leer, denn weil jedes U_i ein Unterraum ist, ist $(\forall i \in I)\ 0 \in U_i$. Das heißt aber gerade $0 \in \bigcap_{i \in I} U_i$.
Seien $v, w \in \bigcap_{i \in I} U_i$ und seien $\alpha, \beta \in \mathbb{K}$. Dann sind für alle $i \in I$ die Vektoren v, w in U_i und folglich (Satz 3.1.6) auch $\alpha v + \beta w \in U_i$. Da das für alle i gilt, ist $\alpha v + \beta w \in \bigcap_{i \in I} U_i$. Nach Satz 3.1.6 ist der Durchschnitt also ein Unterraum. □

Bemerkung 3.1.9. Sei $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$, seien U und W Unterräume von V. Wenn $U \cup W$ auch ein Unterraum von V ist, dann ist $U \subset W$ oder $W \subset U$.

Beweis. Beispiel eines Beweises durch Widerspruch
Sei $U \cup W$ ein Unterraum und sei weder $U \subset W$ noch $W \subset U$. Dann gibt es also $u \in U \setminus W$ und $w \in W \setminus U$. Es sind $u, w \in U \cup W$, und da $U \cup W$ ein Unterraum ist, ist auch $u + w \in U \cup W$. OBdA sei $u + w \in U$. Da auch $u \in U$ und U ein Unterraum ist, ist $w = u + w - u \in U$, im Widerspruch zu $w \in W \setminus U$. □

3.2. **Lineare Hülle und Erzeugendensysteme.**

3.2.1. *Lineare Hülle.*

Definition 3.2.1. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$. Seien $v_1, \cdots, v_n \in V$ und $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$. Dann heißt der Vektor

$$\lambda_1 v_1 + \cdots + \lambda_n v_n$$

eine Linearkombination von $v_1, \cdots, v_n$.

Definition 3.2.2. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$. Sei $M \subset V$ eine nichtleere Teilmenge (endlich oder unendlich). Die lineare Hülle von M ist die Menge

$$\mathcal{L}(M) := \{\sum_{i=1}^{n} \lambda_i v_i \mid n \in \mathbb{N},\, \lambda_i \in \mathbb{K},\, v_i \in M\}.$$

Für die leere Menge definieren wir

$$\mathcal{L}(\emptyset) := \{0\}.$$

Andere Namen für die lineare Hülle sind Span von M oder von M aufgespannter Unterraum.

Satz 3.2.3. *Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$. Sei $M \subset V$ eine nichtleere Teilmenge.*

1) *Die lineare Hülle $\mathcal{L}(M)$ ist ein Unterraum von V*
2) *$M \subset \mathcal{L}(M)$.*
3) *$\mathcal{L}(M)$ ist der kleinste Unterraum, welcher ganz M enthält, im folgenden Sinne:*
   Ist U ein Unterraum von V, welcher M enthält, so ist $\mathcal{L}(M) \subset U$.

Beweis. 1.) Seien $v, w \in \mathcal{L}(M)$ und $\alpha, \beta \in \mathbb{K}$. Wir zeigen: $\alpha v + \beta w \in \mathcal{L}(M)$. Nach Satz 3.1.6 folgt dann, dass $\mathcal{L}(M)$ ein Unterraum ist. Da $v, w \in \mathcal{L}(M)$, können sie als Linearkombinationen von Vektoren aus M geschrieben werden:

$$\begin{aligned} v &= \lambda_1 v_1 + \cdots + \lambda_n v_n, \\ w &= \mu_1 w_1 + \cdots + \mu_m w_m, \end{aligned}$$

mit $\lambda_i, \mu_i \in \mathbb{K}$ und $v_i, w_i \in M$. Es ist dann

$$\alpha v + \beta w = \alpha\lambda_1 v_1 + \cdots + \alpha\lambda_n v_n + \beta\mu_1 w_1 + \cdots + \beta\mu_m w_m$$

wiederum eine Linearkombination von Vektoren aus M, also $\alpha v + \beta w \in \mathcal{L}(M)$.
2.) Sei $v \in M$. Dann lässt sich v trivialerweise als Linearkombination schreiben: $v = 1v$. Also ist $v \in \mathcal{L}(M)$.
3.) Sei U ein Unterraum, der M enthält. Sei $v \in \mathcal{L}(M)$. Wir zeigen: $v \in U$. Tatsächlich lässt sich ja v als Linearkombination $v = \sum_{i=1}^{n} \lambda_i v_i$ schreiben, mit $v_i \in M$. Weil $M \subset U$, folgt dass alle v_i auch in U liegen. Linearkombinationen führen aber aus einem Unterraum nicht hinaus, daher ist auch $v \in U$. □

Im obigen Beweis fehlt etwas: Nämlich der Fall dass M die leere Menge ist. Können Sie das selbst?

Die folgenden Sätzchen können evtl ohne Beweis gebracht werden, evtl. als Übungsbeispiele dienen.

Bemerkung 3.2.4. Jeder Unterraum eines Vektorraums ist seine eigene lineare Hülle.

Beweis. Sei U Unterraum. Offensichtlich ist U ein Unterraum, der U enthält, und zwangsläufig gibt es keinen kleineren, in dem ganz U enthalten ist. Also ist (nach Satz 3.2.3 (3)) $U = \mathcal{L}(U)$. □

Bemerkung 3.2.5. Sei $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$, seien M_1 und M_2 nichtleere Teilmengen von V mit $M_1 \subset M_2$. Dann ist $\mathcal{L}(M_1) \subset \mathcal{L}(M_2)$.

Beweis. $\mathcal{L}(M_2)$ ist ein Unterraum, der M_2, und damit a fortiori (mathematisch für "erst recht") M_1 enthält. Nach Satz 3.2.3 (3) ist $\mathcal{L}(M_1) \subset \mathcal{L}(M_2)$. □

3.2.2. *Lineare Abhängigkeit.*

Wir motivieren zunächst den Begriff der linearen Unabhängigkeit in $\mathbb{R}^2$ und $\mathbb{R}^3$ durch Zeichnungen.

Definition 3.2.6. Sei $(V,+,\cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$, seien $v_1,\cdots,v_n$ Vektoren aus V. Das n-Tupel $(v_1,\cdots,v_n)$ heißt linear unabhängig, wenn es kein n-Tupel $(\lambda_1,\cdots,\lambda_n)\in\mathbb{K}^n$ außer $(0,\cdots,0)$ gibt, für welches gilt:

$$\lambda_1 v_1+\cdots+\lambda_n v_n=0.$$

Andernfalls heißt das Vektoren-n-Tupel $(v_1,\cdots,v_n)$ linear abhängig.
Die leere Menge betrachten wir immer als linear unabhängig.

Wir versuchen einige äquivalente Formulierungen:
$(v_1,\cdots,v_n)$ ist linear unabhängig, wenn gilt:

$$(\forall\lambda_1,\cdots,\lambda_n\in\mathbb{K})\ \big[\sum_{i=1}^n\lambda_i v_i=0\ \Rightarrow\ (\lambda_1,\cdots,\lambda_n)=(0,\cdots,0)\big].$$

$(v_1,\cdots,v_n)$ ist linear abhängig, wenn gilt:

$$(\exists\lambda_1,\cdots,\lambda_n\in\mathbb{K})\ \big[[(\exists i\in\{1,\cdots,n\})\ \lambda_i\neq 0]\quad\wedge\quad[\sum_{i=1}^n\lambda_i v_i=0]\big].$$

Bemerkung 3.2.7. Vektoren-n-Tupel, welche den Nullvektor enthalten, und n-Tupel, in denen ein Vektor mehrmals vorkommt, sind immer linear abhängig.

Beweis. Im n-Tupel $(v_1,\cdots,v_n)$ sei $v_k=0$ für einen Index k. Wir setzen

$$\lambda_i=\begin{cases}0 & \text{falls } i\neq k,\\ 1 & \text{falls } i=k.\end{cases}$$

Dann ist

$$\sum_{i=1}^n\lambda_i v_i=1\cdot v_k=0,$$

aber nicht alle λ_i sind Null. Also ist $(v_1,\cdots,v_n)$ linear abhängig.
Sei nun $(v_1,\cdots,v_n)$ ein n-Tupel, in dem für zwei Indices $p\neq q$ gilt: $v_p=v_q$. Wir haben

$$0v_1+\cdots+1v_p+\cdots+(-1)v_q+\cdots+0v_n=0,$$

aber nicht alle Koeffizienten dieser Linearkombination sind Null. Also ist das n-Tupel linear abhängig. □

Satz 3.2.8. *Sei $(V,+,\cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$. Seien $v_1,\cdots,v_n\in V$. Dann sind folgende Aussagen äquivalent:*

1) *Das n-Tupel $(v_1,\cdots,v_n)$ ist linear unabhängig.*
2) *Für jedes $x\in\mathcal{L}(\{v_1,\cdots,v_n\})$ gibt es nur ein eindeutiges n-Tupel $(\lambda_1,\cdots,\lambda_n)\in\mathbb{K}^n$ sodass $x=\sum_{i=1}^n\lambda_i v_i$.*
3) *Für jedes $i\in\{1,\cdots,n\}$ gilt:*
$$v_i\notin\mathcal{L}(\{v_1,\cdots,v_{i-1},v_{i+1},\cdots,v_n\}).$$
4) *Für jedes $i\in\{1,\cdots,n\}$ gilt:*
$$\mathcal{L}(\{v_1,\cdots,v_{i-1},v_{i+1},\cdots,v_n\})\neq\mathcal{L}(\{v_1,\cdots,v_{i-1},v_i,v_{i+1},\cdots,v_n\}).$$

Beweis. (1) ⇒ (2): Sei $(v_1,\cdots,v_n)$ linear unabhängig. Sei

$$x=\sum_{i=1}^n\lambda_i v_i=\sum_{i=1}^n\mu_i v_i.$$

Dann ist

$$\sum_{i=1}^n(\lambda_i-\mu_i)v_i=x-x=0.$$

Wegen der linearen Unabhängigkeit folgt für alle i dass $\lambda_i - \mu_i = 0$, also $\lambda_i = \mu_i$.
(2) $\Rightarrow$ (3): Angenommen, es gilt (2). Sei $i \in \{1, \cdots, n\}$ beliebig, und nehmen wir an, $v_i \in \mathcal{L}(\{v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n\})$. Dann lässt sich v_i als Linearkombination schreiben:

$$v_i = \lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \cdots + \lambda_n v_n.$$

Andererseits ist natürlich

$$v_i = 0v_1 + \cdots + 1v_i + \cdots + 0v_n.$$

Das sind zwei offensichtlich verschiedene Arten, v_i als Linearkombination der $v_1, \cdots, v_n$ anzusetzen, und ist ein Widerspruch zu (2).
(3) $\Rightarrow$ (4): Wenn (3) gilt, ist für alle i der Vektor $v_i \notin \mathcal{L}(\{v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n\})$. Andererseits ist offensichtlich $v_i \in \mathcal{L}(\{v_1, \cdots, v_{i-1}, v_i, v_{i+1}, \cdots, v_n\})$. Daher sind diese beiden linearen Hüllen verschieden.
(4) $\Rightarrow$ (1): Angenommen, $(v_1, \cdots, v_n)$ sind linear abhängig. Dann gibt es $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$ sodass

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0.$$

Dabei ist mindestens ein $\lambda_i \neq 0$. Wir erhalten

$$v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \cdots - \frac{\lambda_{i-1}}{\lambda_i} v_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} v_{i+1} - \cdots - \frac{\lambda_n}{\lambda_i} v_n.$$

Offensichtlich ist

$$\mathcal{L}(\{v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n\}) \subset \mathcal{L}(\{v_1, \cdots, v_{i-1}, v_i, v_{i+1}, \cdots, v_n\}).$$

Ist $x \in \mathcal{L}(\{v_1, \cdots, v_{i-1}, v_i, v_{i+1}, \cdots, v_n\})$, so können wir x in folgender Form schreiben:

$$\begin{aligned} x &= \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1} + \mu_i v_i + \mu_{i+1} v_{i+1} + \cdots + \mu_n v_n \\ &= \mu_1 v_1 + \cdots + \mu_{i-1} v_{i-1} \\ &\quad + \mu_i (-\frac{\lambda_1}{\lambda_i} v_1 - \cdots - \frac{\lambda_{i-1}}{\lambda_i} v_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} v_{i+1} - \cdots - \frac{\lambda_n}{\lambda_i} v_n) \\ &\quad + \mu_{i+1} v_{i+1} + \cdots + \mu_n v_n \\ &= (\mu_1 - \frac{\mu_i \lambda_1}{\lambda_i}) v_1 + \cdots + (\mu_{i-1} - \frac{\mu_i \lambda_{i-1}}{\lambda_i}) v_{i-1} + (\mu_{i+1} - \frac{\mu_i \lambda_{i+1}}{\lambda_i}) v_{i+1} + \cdots \\ &\quad + (\mu_n - \frac{\mu_i \lambda_n}{\lambda_i}) v_n. \end{aligned}$$

Also ist $x \in \mathcal{L}(\{v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n\})$. Daher ist

$$\mathcal{L}(\{v_1, \cdots, v_{i-1}, v_i, v_{i+1}, \cdots, v_n\}) \subset \mathcal{L}(\{v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n\}),$$

und letztlich sind die beiden linearen Hüllen gleich, im Widerspruch zu (4). □

3.2.3. *Erzeugendensysteme.*

Definition 3.2.9. Sei $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$, und $S \subset V$. Wenn $V = \mathcal{L}(S)$, dann heißt S ein Erzeugendensystem von V.
Besitzt V eine endliche Menge als Erzeugendensystem, so heißt V ein endlich erzeugter Vektorraum.

Bemerkung 3.2.10. Sei $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$. Sei $S \subset T \subset V$. Wenn S ein Erzeugendensystem von V ist, dann ist auch T ein Erzeugendensystem von V.

Beweis. Easy to see! (auch für Sie?) □

Definition 3.2.11. Sei $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$. Seien $v_1, \cdots, v_n \in V$. Das System $(v_1, \cdots, v_n)$ heißt eine Basis von V, wenn folgende zwei Bedingungen erfüllt sind:

1) $(v_1, \cdots, v_n)$ sind linear unabhängig.
2) $\{v_1, \cdots, v_n\}$ ist ein Erzeugendensystem von V.

Satz 3.2.12. *Sei $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$. Seien $v_1, \cdots, v_n \in V$. Folgende Aussagen sind äquivalent:*

1) *$(v_1, \cdots, v_n)$ ist eine Basis von V.*
2) *$(v_1, \cdots, v_n)$ ist ein maximales linear unabhängiges System in folgendem Sinn: $(v_1, \cdots, v_n)$ ist linear unabhängig, und für jeden Vektor $w \in V$ ist das System $(v_1, \cdots, v_n, w)$ linear abhängig.*
3) *Für jeden Vektor $x \in V$ existieren eindeutig bestimmte $\kappa_1, \cdots, \kappa_n \in \mathbb{K}$ so, dass*
$$x = \kappa_1 v_1 + \cdots + \kappa_n v_n.$$
4) *$(v_1, \cdots, v_n)$ ist ein minimales Erzeugendensystem in folgendem Sinne: Das System $(v_1, \cdots, v_n)$ erzeugt den Vektorraum V, aber für keinen Index i erzeugt das System $(v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n)$ ganz V.*

Beweis. (1) $\Rightarrow$ (2): Nach Definition einer Basis ist $(v_1, \cdots, v_n)$ linear unabhängig. Weil $\mathcal{L}(\{v_1, \cdots, v_n\}) = V$, gibt es für jedes $w \in V$ Skalare $\lambda_1, \cdots, \lambda_n$ sodass
$$w = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$
Es ist dann also
$$\lambda_1 v_1 + \cdots + \lambda_n v_n - w = 0,$$
und weil der letzte Koeffizient dieser Linearkombination $-1 \neq 0$ ist, ist das System $(v_1, \cdots, v_n, w)$ linear abhängig.
(2) $\Rightarrow$ (3): Es gelte (2). Sei $x \in V$. Dann ist das System $(v_1, \cdots, v_n, x)$ linear abhängig, also gibt es $\lambda_1, \cdots, \lambda_n, \mu \in \mathbb{K}$ sodass
$$\lambda_1 v_1 + \cdots + \lambda_n v_n + \mu x = 0,$$
und mindestens einer der Skalare $\lambda_1, \cdots, \lambda_n, \mu$ ungleich null ist. Wäre $\mu = 0$, so wäre
$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$
und mindestens eines der λ_i ungleich null. Das widerspricht aber der linearen Unabhängigkeit von $(v_1, \cdots, v_n)$. Also ist $\mu \neq 0$. Wir erhalten
$$x = -\frac{\lambda_1}{\mu} v_1 - \cdots - \frac{\lambda_n}{\mu} v_n.$$
Also lässt sich x als Linearkombination der v_i schreiben. Da $(v_1, \cdots, v_n)$ linear unabhängig sind, ist diese Linearkombination nach Satz 3.2.8 eindeutig bestimmt.
(3) $\Rightarrow$ (4): Es gelte (3). Da sich jedes $x \in V$ als Linearkombination der v_i schreiben lässt, ist $(v_1, \cdots, v_n)$ ein Erzeugendensystem. Angenommen, $(v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n)$ wäre auch ein Erzeugendensystem. Dann lässt sich v_i schreiben:
$$v_i = \lambda_1 v_1 + \cdots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \cdots + \lambda_n v_n.$$
Andererseits ist aber auch $v_i = 1.v_i$, im Widerspruch zur Eindeutigkeit der Darstellung, wie sie in (3) gefordert ist.
(4) $\Rightarrow$ (1): Es gelte (4). Es ist also $(v_1, \cdots, v_n)$ ein Erzeugendensystem. Andererseits ist für jedes i
$$\mathcal{L}(v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n) \neq V = \mathcal{L}(v_1, \cdots, v_n).$$
Nach Satz 3.2.8 folgt die lineare Unabhängigkeit. □

Satz 3.2.13 (Basisauswahlsatz)**.** *Aus jedem endlichen Erzeugendensystem eines Vektorraums lässt sich eine Basis auswählen. Insbesondere besitzt jeder endlich erzeugte Vektorraum eine Basis.*

Beweis. Sei $(v_1, \cdots, v_n)$ ein Erzeugendensystem von V. Wenn $(v_1, \cdots, v_n)$ keine Basis ist, dann ist das System linear abhängig, und es lässt sich ein v_i weglassen, sodass $(v_1, \cdots, v_{i-1}, v_{i+1}, \cdots, v_n)$ auch ein Erzeugendensystem ist. Wenn nun dieses System linear unabhängig ist, haben wir eine Basis, sonst können wir wieder einen Vektor weglassen, solange bis das Verfahren bei einer Basis abbricht. Da es nur endlich viele Vektoren zum Weglassen gibt, muss das Verfahren einmal enden. □

Im folgenden definieren wir lineare Unabhängigkeit für (endliche oder unendliche) Mengen von Vektoren. Der feine Unterschied liegt darin, dass in n-Tupeln von Vektoren derselbe Vektor mehrfach vorkommen kann.

Definition 3.2.14. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$, sei $M \subset V$ eine (endliche oder unendliche) Teilmenge von V. Die Menge M heißt linear abhängig, wenn es ein endliches System von paarweise verschiedenen[9] Vektoren $v_1, \cdots, v_n \in M$ und ein System von Skalaren $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$ gibt, sodass für mindestens einen Index i gilt $\lambda_i \neq 0$, und

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

Anders ausgedrückt, eine (möglicherweise unendliche) Menge M ist linear unabhängig, wenn jedes endliche System von paarweise verschiedenen Vektoren aus M linear unabhängig ist.

Definition 3.2.15. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$, sei $M \subset V$ eine (endliche oder unendliche) Teilmenge von V. Die Menge M heißt eine Basis von V, wenn sie ein linear unabhängiges Erzeugendensystem ist.

Mit Aufwendung hochkarätiger Mittel der Mengenlehre (Auswahlaxiom, Zornsches Lemma) kann man dann beweisen, dass jeder Vektorraum (auch ein nicht endlich erzeugter) eine Basis besitzt. Allerdings haben Basen in nicht endlich erzeugten Vektorräumen eine weit geringere Bedeutung. Statt jeden Vektor von V als Linearkombination von jeweils endlich vielen Vektoren aus der Basis darzustellen, benützt man Vektorsysteme, mit Hilfe derer sich jeder Vektor von V als unendliche Reihe schreiben lässt.

3.3. **Erste Schritte in der Matrizenrechnung.**

3.3.1. *Multiplikation einer Matrix mit einem Vektor.*

Definition 3.3.1. Sei $\mathbb{K}$ ein Körper, $m, n \in \mathbb{N}$. Sei

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \in \mathbb{K}^{m \times n}, \qquad x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{K}^n.$$

Wir definieren das Produkt $Ax \in \mathbb{K}^m$ durch

$$Ax := \begin{pmatrix} \alpha_{1,1}\xi_1 + \cdots + \alpha_{1,n}\xi_n \\ \vdots \\ \alpha_{m,1}\xi_1 + \cdots + \alpha_{m,n}\xi_n \end{pmatrix}.$$

Es sei ausdrücklich darauf hingewiesen, dass die Reihenfolge der Schreibung Ax wichtig ist. Das Produkt xA würde, wie wir später sehen werden, etwas anderes bedeuten!

Bemerkung 3.3.2. Damit man eine Matrix A mit einem Spaltenvektor x multiplizieren kann, muss A soviele Spalten haben wie x Komponenten. Das Produkt Ax ist dann wieder ein Spaltenvektor, und hat soviele Komponenten, wie A Zeilen hat.

[9] das heißt: $v_i \neq v_j$ falls $i \neq j$.

Beispiel 3.3.3.

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -2 & 0 \\ 2 & 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 4 \\ -5 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 2 + 2 \cdot 4 + 3 \cdot (-5) + 4 \cdot 1 \\ 0 \cdot 2 + 1 \cdot 4 + (-2) \cdot (-5) + 0 \cdot 1 \\ 2 \cdot 2 + 3 \cdot 4 + 0 \cdot (-5) + 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 14 \\ 17 \end{pmatrix}.$$

Beispiel 3.3.4. Das Produkt

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & -2 & 0 \\ 2 & 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ -6 \end{pmatrix}$$

ist nicht definiert, weil die Spaltenzahl der Matrix (=4) und die Zeilenzahl des Vektors (=3) nicht gleich sind.

Bemerkung 3.3.5. Ein Grund (nicht der einzige) für die zunächst seltsam anmutende Definition von Ax ist, dass man damit lineare Gleichungssysteme anschreiben kann:
Ist $\mathbb{K}$ ein Körper, $m, n \in \mathbb{N}$, sind

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \in \mathbb{K}^{m \times n}, \qquad x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{K}^n, \qquad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \in \mathbb{K}^m.$$

Dann ist die Gleichung $Ax = b$ nichts anderes als das Gleichungssystem

$$\begin{array}{ccccccc} \alpha_{1,1}\xi_1 & + & \cdots & + & \alpha_{1,n}\xi_n & = & \beta_1 \\ \vdots & & & & & & \\ \alpha_{m,1}\xi_1 & + & \cdots & + & \alpha_{m,n}\xi_n & = & \beta_m \end{array}$$

Satz 3.3.6. *Sei $\mathbb{K}$ ein Körper, $m, n \in \mathbb{N}$, $A \in \mathbb{K}^{m \times n}$. Die Abbildung*

$$\begin{cases} \mathbb{K}^n & \to \mathbb{K}^m \\ x & \mapsto Ax \end{cases}$$

erfüllt die folgenden Bedingungen: Sind $\lambda \in \mathbb{K}$, $x, y \in \mathbb{K}^n$, so gilt

1) $A(\lambda x) = \lambda(Ax)$.
2) $A(x + y) = Ax + Ay$.

Beweis. Wir rechnen Punkt (2) nach, Punkt (1) geht ebenso durch direktes Nachrechnen. Seien also

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \in \mathbb{K}^{m \times n}, \qquad x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{K}^n, \qquad y = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} \in \mathbb{K}^n.$$

Es ist

$$\begin{aligned} A(x+y) &= \begin{pmatrix} \sum_{j=1}^n \alpha_{1,j}(\xi_j + \eta_j) \\ \vdots \\ \sum_{j=1}^n \alpha_{m,j}(\xi_j + \eta_j) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n \alpha_{1,j}\xi_j + \sum_{j=1}^n \alpha_{1,j}\eta_j \\ \vdots \\ \sum_{j=1}^n \alpha_{m,j}\xi_j + \sum_{j=1}^n \alpha_{m,j}\eta_j \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^n \alpha_{1,j}\xi_j \\ \vdots \\ \sum_{j=1}^n \alpha_{m,j}\xi_j \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^n \alpha_{1,j}\eta_j \\ \vdots \\ \sum_{j=1}^n \alpha_{m,j}\eta_j \end{pmatrix} = Ax + Ay. \end{aligned}$$

□

Später werden wir eine Abbildung, die Punkte (1) und (2) aus Satz 3.3.6 erfüllt, eine lineare Abbildung nennen. Lineare Abbildungen und die Tatsache, dass die Multiplikation mit Matrizen linear ist, werden eine ganz zentrale Rolle in dieser Vorlesung spielen.

Definition 3.3.7. Sei $\mathbb{K}$ ein Körper, $m, n \in \mathbb{N}$. Sei

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \in \mathbb{K}^{m\times n}.$$

Die Spalten von A sind die Vektoren

$$\begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix}, \cdots, \begin{pmatrix} \alpha_{1,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix} \in \mathbb{K}^m.$$

Die Zeilen von A sind die Zeilenvektoren (d.s. $1 \times n$-Matrizen)

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \end{pmatrix}, \cdots, \begin{pmatrix} \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \in \mathbb{K}^{1\times n}.$$

Satz 3.3.8. *Sei $\mathbb{K}$ ein Körper, $m, n \in \mathbb{N}$. Seien*

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \in \mathbb{K}^{m\times n}, \qquad x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{K}^n.$$

Dann ist das Produkt Ax eine Linearkombination der Spalten von A:

$$Ax = \xi_1 \begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix} + \cdots + \xi_n \begin{pmatrix} \alpha_{1,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix}.$$

Beweis. Direkt nachrechnen! □

3.3.2. *Kern, Spaltenraum, Zeilenraum.*

Definition 3.3.9. Sei $\mathbb{K}$ ein Körper und A eine $m\times n$-Matrix über $\mathbb{K}$. Wir definieren die folgenden Mengen:

(1) Den Kern von A:
$$\ker(A) := \{x \in \mathbb{K}^n \mid Ax = 0\}$$
(eine Teilmenge des $\mathbb{K}^n$).
(2) Den Spaltenraum von A (geschrieben $\mathrm{rg}(A)$) als die lineare Hülle der Spalten von A (eine Teilmenge von $\mathbb{K}^m$).
(3) Den Zeilenraum von A als die lineare Hülle der Zeilen von A (eine Teilmenge von $\mathbb{K}^{1\times n}$)

Bemerkung 3.3.10.

(1) Der Kern einer Matrix $A \in \mathbb{K}^{m\times n}$ ist die Menge aller Lösungen des homogenen Gleichungssystems $Ax = 0$, vgl. Bemerkung 3.3.5. Der Kern heißt im Englischen auch oft Nullspace, darf aber nicht mit dem Nullraum $\{0\}$ verwechselt werden.
(2) Der Spaltenraum einer Matrix ist
$$\mathrm{rg}(A) = \{Ax \mid x \in \mathbb{K}^n\},$$
vgl. Satz 3.3.8. Anders ausgedrückt, der Spaltenraum von A ist
$$\mathrm{rg}(A) = \{b \in \mathbb{K}^m \mid Ax = b \text{ besitzt eine Lösung}\}.$$

Die Schreibweise rg kommt vom englischen Wort "range" (Wertebereich, Bildraum), nicht zu verwechseln mit "rank", dem Rang, den wir später definieren werden.

(3) Für den Zeilenraum führen wir hier keine Schreibweise ein, wir werden ihn relativ selten verwenden.

Satz 3.3.11. *Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{m\times n}$.*

(1) *Der Kern von A ist ein Unterraum des $\mathbb{K}^n$.*
(2) *Der Spaltenraum von A ist ein Unterraum des $\mathbb{K}^m$.*
(3) *Der Zeilenraum von A ist ein Unterraum des $\mathbb{K}^{1\times n}$.*

Beweis. Da Spaltenraum und Zeilenraum als lineare Hüllen definiert sind, sind sie Vektorräume. Um zu zeigen, dass $\ker(A)$ ein Unterraum ist, nehmen wir $x, y \in \ker(A)$ und $\lambda, \mu \in \mathbb{K}$. Es ist dann $Ax = Ay = 0$, und unter Benützung von Satz 3.3.6 erhalten wir:

$$A(\lambda x + \mu y) = A(\lambda x) + A(\mu y) = \lambda A(x) + \mu A(y) = \lambda 0 + \mu 0 = 0.$$

Also ist auch $\lambda x + \mu y \in \ker(A)$. □

3.3.3. *Elementare Zeilenumformungen.*

Definition 3.3.12. Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{m\times n}$ (kann auch ein Spaltenvektor in $\mathbb{K}^{m\times 1}$ sein). Die folgenden Umformungen in der Matrix A heißen elementare Zeilenumformungen:

(1) Vertauschen zweier Zeilen.
(2) Multiplikation einer Zeile mit einem Skalar $\lambda \in \mathbb{K} \setminus \{0\}$.
(3) Addition eines Vielfachen einer Zeile zu einer anderen Zeile.

Lemma 3.3.13. *Jede elementare Zeilenumformung lässt sich durch eine elementare Zeilenumformung des gleichen Typs rückgängig machen:*

(1) *Vertauschen der Zeilen i und k durch ein nochmaliges Vertauschen der Zeilen i und k.*
(2) *Multiplikation der Zeile i mit λ durch anschließende Multiplikation der Zeile i mit λ^{-1}.*
(3) *Addition des λ-fachen der Zeile i zur Zeile k durch anschließende Addition des $(-\lambda)$-fachen der Zeile i zur Zeile k.*

Beweis. Offensichtlich. (D.h. wissen Sie, warum?) □

Satz 3.3.14. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{m\times n}$, $b \in \mathbb{K}^m$. Mit $z_1, \cdots, z_m$ bezeichnen wir die Zeilen von A, mit $s_1, \cdots, s_n$ die Spalten von A.*
Eine elementare Zeilenumformung forme A in die Matrix $\tilde{A}$ um, und durch dieselbe Zeilenumformung werde b in den Vektor $\tilde{b}$ umgeformt. Mit $\tilde{z}_1, \cdots, \tilde{z}_m$ bezeichnen wir die Zeilen von $\tilde{A}$, mit $\tilde{s}_1, \cdots, \tilde{s}_n$ die Spalten von $\tilde{A}$.
Dann gilt:

(1) *Ein Vektor $x \in \mathbb{K}^n$ löst genau dann das Gleichungssystem $Ax = b$, wenn er das Gleichungssystem $\tilde{A}x = \tilde{b}$ löst.*
(2) *A und $\tilde{A}$ haben denselben Kern.*
(3) *Sind $j_1, \cdots, j_r \in \{1, \cdots, n\}$, so ist das r-Tupel $(s_{j_1}, \cdots, s_{j_r})$ genau dann linear abhängig, wenn das r-Tupel $(\tilde{s}_{j_1}, \cdots, \tilde{s}_{j_r})$ linear abhängig ist.*
(4) *Der Vektor b liegt genau dann im Spaltenraum von A, wenn $\tilde{b}$ im Spaltenraum von $\tilde{A}$ liegt.*
(5) *Sind $j_1, \cdots, j_r \in \{1, \cdots, n\}$, so spannt $\{s_{j_1}, \cdots, s_{j_r}\}$ genau dann den Spaltenraum von A auf, wenn $\{\tilde{s}_{j_1}, \cdots, \tilde{s}_{j_r}\}$ den Spaltenraum von $\tilde{A}$ aufspannt.*

(6) *Die Matrizen A und $\tilde{A}$ haben denselben Zeilenraum.*

Beweis. Es seien $\alpha_{i,j}$ die Koeffizienten von A, β_i die Koeffizienten von b, und ξ_j die Koeffizienten von x. Wir werden in diesem Beweis immer wieder verwenden, dass die folgenden drei Aussagen äquivalent sind (Bemerkung 3.3.5, Satz 3.3.8):

$$Ax = b, \tag{3.3.1}$$

$$\begin{array}{ccccccc} \alpha_{1,1}\xi_1 & + & \cdots & + & \alpha_{1,n}\xi_n & = & \beta_1 \\ \vdots & & & & \vdots & & \vdots \\ \alpha_{m,1}\xi_1 & + & \cdots & + & \alpha_{m,n}\xi_n & = & \beta_m \end{array} \tag{3.3.2}$$

$$\xi_1 s_1 + \cdots + \xi_n s_n = b. \tag{3.3.3}$$

Punkt (1)*:* Wir nehmen an, dass x eine Lösung von $Ax = b$ ist, und zeigen, dass $\tilde{A}x = \tilde{b}$.
Vertauschung der Reihenfolge der einzelnen Gleichungen ändert natürlich nichts an der Lösungsmenge, in diesem Fall ist also auch $\tilde{A}x = \tilde{b}$.
Wird die Zeile Nummer i mit einem Faktor $\lambda \neq 0$ multipliziert, so löst x natürlich alle unveränderten Gleichungen des Systems (3.3.2), und zusätzlich die Zeile

$$\lambda\alpha_{i,1}\xi_1 + \cdots + \lambda\alpha_{i,n}\xi_n = \lambda\beta_i.$$

Damit löst x wiederum das gesamte Gleichungssystem $\tilde{A}x = \tilde{b}$.
Wird das λ-fache der i-ten Zeile zur k-ten Zeile addiert, so gehen wir davon aus, dass gilt:

$$\begin{array}{ccccccc} \alpha_{i,1}\xi_1 & + & \cdots & + & \alpha_{i,n}\xi_n & = & \beta_i \\ \alpha_{k,1}\xi_1 & + & \cdots & + & \alpha_{k,n}\xi_n & = & \beta_k. \end{array}$$

Folglich gilt auch

$$(\alpha_{k,1} + \lambda\alpha_{i,1})\xi_1 + \cdots + (\alpha_{k,n} + \lambda\alpha_{i,n})\xi_n = \beta_k + \lambda\beta_i$$

Dies ist aber die einzige neue Zeile im Gleichungssystem $\tilde{A}x = \tilde{b}$.
In jedem Fall ist also jede Lösung von $Ax = b$ erst recht eine Lösung von $\tilde{A}x = \tilde{b}$. Weil nach Lemma 3.3.13 auch A aus $\tilde{A}$ und b aus $\tilde{b}$ durch eine elementare Zeilenumformung hervorgehen, gilt nun auch: Wenn x eine Lösung von $\tilde{A}x = \tilde{b}$ ist, so ist auch $Ax = b$.

Punkt (2)*:* Das folgt aus Punkt (1), denn der Kern von A ist nichts anderes als die Lösungsmenge von $Ax = 0$.

Punkt (3)*:* Sei das r-Tupel $(s_{j_1}, \cdots, s_{j_r})$ linear abhängig. Es gibt also $\xi_{j_1}, \cdots, \xi_{j_r} \in \mathbb{K}$, nicht alle gleich null, sodass

$$\xi_{j_1} s_{j_1} + \cdots + \xi_{j_r} s_{j_r} = 0.$$

Setzen wir

$$\xi_j = 0 \text{ für } \; j \notin \{j_1, \cdots, j_r\}$$

(für die anderen j ist ξ_j ja bereits gegeben), und

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix},$$

so erhalten wir

$$\xi_1 s_1 + \cdots + \xi_n s_n = 0.$$

Das heißt aber zugleich dass x eine Lösung des homogenen Gleichungssystems $Ax = 0$ ist. Nach Punkt (1) löst x auch das Gleichungssystem $\tilde{A}x = 0$, also gilt

$$\xi_1 \tilde{s}_1 + \cdots + \xi_n \tilde{s}_n = 0.$$

Lassen wir aus dieser Summe die Nullterme weg, die für $j \notin \{j_1, \cdots, j_r\}$ auftreten, erhalten wir

$$\xi_{j_1}\tilde{s}_{j_1} + \cdots + \xi_{j_r}\tilde{s}_{j_r} = 0.$$

Weil nach Voraussetzung nicht alle ξ_{j_k} gleich null sind, ist das r-Tupel $(\tilde{s}_{j_1}, \cdots, \tilde{s}_{j_r})$ linear abhängig. Wieder entstehen auch A und b durch eine Zeilentransformation aus $\tilde{A}$, $\tilde{b}$, sodass ebenso gilt: Wenn $(\tilde{s}_{j_1}, \cdots, \tilde{s}_{j_r})$ linear abhängig sind, so sind auch $(s_{j_1}, \cdots, s_{j_r})$ linear abhängig.

Punkt (4)*:* Punkt (1) und Bemerkung 3.3.10.

Punkt (5)*:* Wir nehmen an, dass der Spaltenraum von $\tilde{A}$ durch $\{\tilde{s}_{j_1}, \cdots, \tilde{s}_{j_r}\}$ erzeugt wird, und beweisen, dass der Spaltenraum von A durch $\{s_{j_1}, \cdots, s_{j_r}\}$ aufgespannt wird. Sei also $b \in \mathrm{rg}(A)$. Wir wenden auf b dieselbe Zeilentransformation an wie auf A und erhalten $\tilde{b}$. Nach Punkt (4) ist $\tilde{b} \in \mathrm{rg}(\tilde{A})$ und kann deshalb als Linearkombination von $\tilde{s}_{j_1}, \cdots, \tilde{s}_{j_r}$ geschrieben werden.

$$\tilde{b} = \xi_{j_1}\tilde{s}_{j_1} + \cdots + \xi_{j_r}\tilde{s}_{j_r}.$$

Wir setzen wieder $\xi_j = 0$ für $j \notin \{j_1, \cdots, j_r\}$ und $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$. Dann ist also $\tilde{b} = \tilde{A}x$ und folglich $b = Ax$. Wenn wir das als Linearkombination der Spalten umschreiben und die Nullterme bei $j \notin \{j_1, \cdots, j_r\}$ weglassen, erhalten wir b als Linearkombination der Spalten $s_{j_1}, \cdots, s_{j_r}$:

$$b = \xi_{j_1}s_{j_1} + \cdots + \xi_{j_r}s_{j_r}.$$

Punkt (6)*:* Es ist klar, dass Umordnung der Zeilen an der linearen Hülle der Zeilen nichts ändert. Bei den beiden anderen elementaren Umformungen geht die neue Zeile aus den alten Zeilen durch Linearkombination hervor. Es sind also die Zeilen $\tilde{z}_1, \cdots, \tilde{z}_m$ Linearkombinationen von $z_1, \cdots, z_m$ und folglich ist $\{\tilde{z}_1, \cdots, \tilde{z}_m\} \subset \mathcal{L}(\{z_1, \cdots, z_m\})$. Da aber nach Satz 3.2.3 die lineare Hülle $\mathcal{L}(\{\tilde{z}_1, \cdots, \tilde{z}_m\})$ in jedem Unterraum als Teilmenge enthalten ist, welcher $\{\tilde{z}_1, \cdots, \tilde{z}_m\}$ enthält, gilt insbesondere

$$\mathcal{L}(\{\tilde{z}_1, \cdots, \tilde{z}_m\}) \subset \mathcal{L}(\{z_1, \cdots, z_m\}).$$

Wieder kann man die Rollen von $\tilde{A}$ und A vertauschen, und erhält ebenso

$$\mathcal{L}(\{z_1, \cdots, z_m\}) \subset \mathcal{L}(\{\tilde{z}_1, \cdots, \tilde{z}_m\}).$$

□

3.3.4. *Verallgemeinerte Zeilenstufenform.*

In der Lehrbuchliteratur wird bei der Erörterung des Gausschen Eliminationsverfahrens üblicherweise als Ergebnis der Eliminationsphase eine Matrix in Zeilenstufenform (siehe unten) präsentiert. Damit man eine Zeilenstufenform erhält, muss man bei der Pivotsuche stets jenes Element wählen, das unter allen möglichen Pivotelementen am weitesten links in der Matrix steht. Tatsächlich könnte aber in vielen Fällen die Wahl eines anderen Pivotelementes bequemer oder aus numerischen Gründen günstiger sein (z.B. könnte ein kleines Pivotelement, das in einer Zeile steht, in der sonst sehr große Elemente stehen, sich ungünstig auf Rundungsfehler auswirken). Um freie Wahl der Pivotspalte zuzulassen, werden wir als Endform der Elimination eine Form wählen, in der kein weiterer Pivotschritt möglich ist. Wir werden diese Form, die sich durch Umordnung der Zeilen und Spalten in eine Zeilenstufenform umwandeln lässt, als eine "verallgemeinerte Zeilenstufenform" benennen. Mir ist leider aus der Literatur kein Name bekannt. Die große Beliebtheit der Zeilenstufenform unter Lehrbuchautoren kommt sicher daher, dass es leichter ist, die Zeilenstufenform zu beschreiben[10]. Damit die schwer lesbare Definition 3.3.15 besser verständlich wird, habe ich unter jede Zeile eine Interpretation im Sinne von Pivotelementen geschrieben.

[10]Faulheit war schon immer eine starke Antriebsfeder für mathematische Kreativität.

Definition 3.3.15. Sei $\mathbb{K}$ ein Körper und $A = (\alpha_{i,j})_{i=1,\cdots,m;\, j=1,\cdots,n} \in \mathbb{K}^{m\times n}$. Wir sagen, dass A in verallgemeinerter Zeilenstufenform vorliegt, wenn gilt:
Es gibt r Indexpaare $(i_1, j_1), \cdots, (i_r, j_r) \in \{1, \cdots, m\} \times \{1, \cdots, n\}$ sodass die folgenden Bedingungen erfüllt sind:

(1) Für alle $k = 1, \cdots, r$ ist $\alpha_{i_k,j_k} \neq 0$.
Den Koeffizienten α_{i_k,j_k} bezeichnen wir als das k-te Pivotelement, und fordern hier, dass die Pivotelemente nicht Null sind.

(2) Die Indices $i_1, \cdots, i_m$ sind paarweise verschieden, und die Indices $j_1, \cdots, j_n$ sind paarweise verschieden.
Keine Zeile und keine Spalte enthält zwei Pivotelemente.

(3) Für alle $k, l \in \{1, \cdots, r\}$ mit $k > l$ gilt: $\alpha_{i_k,j_l} = 0$.
Enthält eine Spalte das l-te Pivotelement, so sind die Einträge in dieser Spalte gleich Null ab der $(l+1)$-ten Pivotzeile.

(4) Für $i \notin \{i_1, \cdots, i_r\}$ und jedes $j = 1, \cdots, n$ ist $\alpha_{i,j} = 0$.
Alle Zeilen, die kein Pivotelement enthalten, sind Nullzeilen.

Wir bezeichnen die Spalten mit Indices $j_1, \cdots, j_r$ als basische[11] Spalten, die anderen (falls es welche gibt) sind nichtbasische Spalten.
Gilt zusätzlich: $i_1 < i_2 < \cdots < i_r$, $j_1 < j_2 < \cdots < j_r$, so sagen wir, die Matrix A liegt in Zeilenstufenform vor.

Beispiel 3.3.16. Die folgenden Schemata sind Beispiele von Matrizen in verallgemeinerter Zeilenstufenform und in Zeilenstufenform. Dabei sind die Pivotelemente in ihrer Reihenfolge mit $(1), (2), \cdots$ markiert, $*$ bedeutet ein Element, das nicht unbedingt 0 sein muss. Die linke Matrix kann man durch Vertauschung der Zeilen und Spalten in die rechte Matrix überführen.

$$\begin{pmatrix} * & 0 & 0 & 0 & 0 & (4) & * \\ * & (3) & 0 & 0 & 0 & * & * \\ * & * & * & (1) & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & (2) & 0 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} (1) & * & * & * & * & * & * \\ 0 & (2) & * & * & * & * & * \\ 0 & 0 & 0 & (3) & * & * & * \\ 0 & 0 & 0 & 0 & (4) & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Bemerkung 3.3.17. Wir wissen bereits aus Algorithmus 1.1.19, dass sich jede Matrix durch Zeilentransformationen auf verallgemeinerte Zeilenstufenform bringen lässt (Eliminationsphase), und von dort aus weiter auf eine verallgemeinerte Zeilenstufenform, in der alle basischen Spalten Einheitsvektoren sind (Rücksubstitution). Ist das System $Ax = b$ überhaupt lösbar, und geben wir für jeden Spaltenindex j der nichtbasischen Spalten einen Wert ξ_j vor, so lassen sich eindeutige ξ_j für die basischen Spaltenindices finden, sodass der Vektor $x = (\xi_1, \cdots, \xi_n)^T$ das System $Ax = b$ löst.

Satz 3.3.18. *Sei $\mathbb{K}$ ein Körper, sei $A = (\alpha_{i,j})_{i=1,\cdots,m\, j=1,\cdots,n} \in \mathbb{K}^{m\times n}$ in verallgemeinerter Zeilenstufenform mit $i_1, \cdots, i_r$, $j_1, \cdots, j_r$ wie in Definition 3.3.15. Die Zeilen von A bezeichnen wir mit $z_1, \cdots, z_m$, die Spalten mit $s_1, \cdots, s_n$. Wir bezeichnen mit $B = \{j_1, \cdots, j_r\}$ die Menge der basischen Spaltenindices.*

(1) *Sei $b = (\beta_i)_{i=1,\cdots,m} \in \mathbb{K}^m$ so, dass $\beta_i = 0$ falls $i \notin \{i_1, \cdots, i_r\}$. Wählt man für jeden nichtbasischen Spaltenindex j je einen Skalar $\xi_j \in \mathbb{K}$, so lassen sich für die basischen Spaltenindices in eindeutiger Weise Skalare ξ_{j_k} bestimmen, sodass der Vektor $x = (\xi_j)_{j=1,\cdots,n}$ das System $Ax = b$ löst.*

[11] Die Bezeichnung basische Spalten ist in der linearen Optimierung üblich. Der Simplexalgorithmus der linearen Optimierung bedient sich ebenfalls der elementaren Zeilentransformationen. Wir werden sehen, dass die basischen Spalten eine Basis des Spaltenraumes bilden.

(2) *Für jeden nichtbasischen Spaltenindex k sei $x_k \in \mathbb{K}^n$ jeweils die eindeutige Lösung $Ax_k = 0$, für die die k-te Komponente eins und die j-ten Komponenten für $j \neq k$, $j \notin B$ gleich Null sind. Dann ist das Vektorsystem $(x_k)_{k\notin B}$ eine Basis des Kerns von A.*
(3) *Die Spalten $(s_{j_1}, \cdots, s_{j_r})$ sind eine Basis des Spaltenraumes von A.*
(4) *Die Zeilen $(z_{i_1}, \cdots, z_{i_r})$ sind eine Basis des Zeilenraumes von A.*
(5) *Insbesondere gibt es eine Basis des Zeilenraumes und eine Basis des Spaltenraumes, jeweils mit Mächtigkeit r, und eine Basis des Kerns mit Mächtigkeit $n - r$.*

Beweis. Punkt (1)*:* Durch Rücksubstitution, siehe Algorithmus 1.1.19.

Punkt (2)*:* Wir beobachten zuerst, dass für jeden der so berechneten Vektoren x_k und jeden nichtbasischen Spaltenindex j die Komponente Nummer j entweder null oder eins ist, wobei eins nur im Falle $k = j$ gilt. Bilden wir also eine Linearkombination $x = \sum_{j\notin B} \lambda_j x_j$, so ist für jeden nichtbasischen Index k die k-te Komponente von x gleich λ_k. Daraus folgt unmittelbar, dass keine nichttriviale [12] Linearkombination der x_j den Nullvektor ergibt, also ist $(x_j)_{j\notin B}$ linear unabhängig.
Sei nun $y = (\eta_j)_{j=1,\cdots,n} \in \ker(A)$. Wir definieren $x = \sum_{j\notin B} \eta_j x_j$ und erhalten wieder eine Lösung $x = (\xi_i)_{i=1,\cdots,n}$ von $Ax = 0$ mit $\xi_j = \eta_j$ für alle $j \notin B$. Nach Punkt (1) ist diese Lösung aber eindeutig bestimmt, also muss $x = y$ sein. Daher ist $y = x$ und folglich $y \in \mathcal{L}(\{x_1, \cdots, x_{n-r}\})$. Also spannt $(x_1, \cdots, x_{n-r})$ den Kern von A auf.

Punkt (3) Sei $b = (\beta_i)_{i=1,\cdots,m} \in \mathrm{rg}(A)$. Ist $i \notin \{i_1, \cdots, i_r\}$, dann ist die i-te Zeile von A eine Nullzeile, und weil $b = Ax$ mindestens eine Lösung hat, muss auch $\beta_i = 0$ sein. Wegen Punkt (1) lässt sich b eindeutig als Linearkombination $\sum_{j=1}^n \xi_j s_j$ der Spalten von A so schreiben, dass die Koeffizienten der nichtbasischen Spalten null sind: $\xi_k = 0$ für $k \notin B$. Entfernt man aus der Linearkombination die nichtbasischen Spalten (welche ohnehin nur mit Koeffizient 0 vorkommen), so hat man b in eindeutiger Weise als Linearkombination der basischen Spalten geschrieben. Damit ist (vgl. Satz 3.2.12) $(s_{j_1}, \cdots, s_{j_r})$ eine Basis von $\mathrm{rg}(A)$.

Punkt (4)*:* Da es sonst nur Nullzeilen gibt, spannen die Zeilen $z_{i_1}, \cdots, z_{i_r}$ jedenfalls den ganzen Zeilenraum von A auf. Zu zeigen bleibt, dass diese Zeilen linear unabhängig sind. Seien $\lambda_{i_1}, \cdots, \lambda_{i_r} \in \mathbb{K}$ so, dass gilt

$$\lambda_{i_1} z_{i_1} + \cdots + \lambda_{i_r} z_{i_r} = 0.$$

Ausgeschrieben heißt das für alle Spaltenindices $j = 1, \cdots, n$:

$$\lambda_{i_1} \alpha_{i_1,j} + \cdots + \lambda_{i_r} \alpha_{i_r,j} = 0.$$

Wir führen einen Induktionsbeweis, dass für alle $s \in \{1, \cdots, r\}$ gilt:

$$\lambda_{i_1} = \cdots = \lambda_{i_s} = 0.$$

Der Fall $s = r$ liefert dann, dass alle $\lambda_{i_k} = 0$ sind, und damit die lineare Unabhängigkeit der z_{i_k}.
Induktionsverankerung: $s = 1$. Wir betrachten die j_1-te Spalte der Matrix A. Es gilt $\alpha_{i_1,j_1} \neq 0$ und $\alpha_{i_k,j_1} = 0$ für alle $k > 1$. Daher ist

$$0 = \sum_{k=1}^{r} \lambda_{i_k} \alpha_{i_k,j_1} = \lambda_{i_1} \alpha_{i_1,j_1},$$

und weil $\alpha_{i_1,j_1} \neq 0$, folgt $\lambda_{i_1} = 0$. Damit ist die Induktion verankert.
Induktionsschritt: Wir nehmen an, $\lambda_{i_1} = \cdots = \lambda_{i_s} = 0$ für ein $s < r$ und müssen nun zeigen, dass $\lambda_{i_{s+1}} = 0$. Dazu betrachten wir die j_{s+1}-te Spalte von A. Hier sind

[12] d.h. keine Linearkombination, bei der einer oder mehrere Koeffizienten ungleich Null sind.

für $k > s+1$ die Koeffizienten $\alpha_{i_k,j_{s+1}} = 0$, und $\alpha_{i_{s+1},j_{s+1}} \neq 0$. Wir verwenden auch dass $\lambda_{i_k} = 0$ für $k \leq s$. Damit ist

$$0 = \sum_{k=1}^{r} \lambda_{i_k} \alpha_{i_k,j_{s+1}} = \lambda_{i_{s+1}} \alpha_{i_{s+1},j_{s+1}}.$$

Also ist $\lambda_{i_{s+1}} = 0$, und der Induktionsschritt ist fertig.

Punkt (5)*:* ist eine direkte Folge der oberen Punkte. □

Korollar 3.3.19. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{m\times n}$, und durch elementare Zeilentransformationen werde A in eine Matrix $\tilde{A}$ übergeführt. $\tilde{A}$ sei in verallgemeinerter Zeilenstufenform, mit $i_1, \cdots, i_r$, $j_1, \cdots, j_r$ wie in Definition 3.3.15. Sei B die Menge der Spaltenindices, die basische Spalten beschreiben. Dann gilt:*

1. *Die Spalten mit Indices $j_1, \cdots, j_r$ der Matrix A bilden eine Basis des Spaltenraumes von A.*
2. *Die Zeilen der Matrix $\tilde{A}$, welche nicht Nullzeilen sind, bilden eine Basis des Zeilenraumes von A.*
3. *Für $k \notin B$ sei $x_k = (\xi_{k,1}, \cdots, \xi_{k,n})^T$ die eindeutige Lösung des Gleichungssystems $Ax_k = 0$ mit $\xi_{k,j} = 0$ für alle $j \notin B$, $j \neq k$ und $\xi_{k,k} = 1$. Dann ist $(x_k)_{k\notin B}$ eine Basis von $\ker(A)$.*
4. *Es gibt Basen des Zeilenraumes und des Spaltenraumes von A jeweils mit Mächtigkeit r, und eine Basis des Kerns von A mit Mächtigkeit $n - r$.*

Beweis. Satz 3.3.14 und Satz 3.3.18. □

Bemerkung 3.3.20. Aus der Lage der Pivotelemente sieht man, wo die Zeilen und Spalten liegen, die man zur Bildung einer Basis von Zeilen- und Spaltenraum heranziehen kann. Für die Basis des Spaltenraumes nimmt man dann Spalten aus der ursprünglichen Matrix A, für die Basis des Zeilenraumes Zeilen der transformierten Matrix $\tilde{A}$.

Noch haben wir nicht gezeigt, dass jedesmal, wenn eine Matrix in verallgemeinerte Zeilenstufenform übergeführt wird, dieselbe Anzahl r von basischen Spalten herauskommt. Zunächst könnte man sich ja vorstellen, dass die Anzahl der Pivotelemente, die gebraucht werden, vielleicht von der geschickten Wahl der Pivotelemente abhängt. Wir werden aber bald wissen, dass alle Basen eines endlich erzeugten Vektorraums gleich viel Vektoren enthalten, und damit ist die Zahl r vom Rechengang unabhängig. Dann werden wir r als den Rang der Matrix bezeichnen.

3.3.5. *Ein Beispiel.*

Beispiel 3.3.21. Wir untersuchen die Matrix:

$$A = \begin{pmatrix} 2 & 6 & 8 & -5 & 17 \\ 2 & 8 & 6 & -10 & 8 \\ 2 & 6 & 7 & -8 & 11 \\ 0 & 0 & 1 & 3 & 6 \end{pmatrix}$$

Pivotschritte auf verallgemeinerte Zeilenstufenform:

$$A = \begin{pmatrix} 2 & 6 & 8 & -5 & 17 \\ 2 & 8 & 6 & -10 & 8 \\ 2 & 6 & 7 & -8 & 11 \\ 0 & 0 & 1* & 3 & 6 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 6 & 0 & -29 & -31 \\ 2 & 8 & 0 & -28 & -28 \\ 2* & 6 & 0 & -29 & -31 \\ 0 & 0 & 1* & 3 & 6 \end{pmatrix}$$

$$\tilde{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1* & 3 \\ 2* & 6 & 0 & -29 & -31 \\ 0 & 0 & 1* & 3 & 6 \end{pmatrix}$$

1) Die Spalten von A sind linear abhängig, denn es gibt nichtbasische Spalten. Daher hat das Gleichungssystem $Ax = 0$ nichttriviale Lösungen.[13]
2) Eine Basis des Spaltenraumes: aus der alten Matrix A die basischen Spalten, also die Spalten, aus denen Pivotspalten wurden:

$$\begin{pmatrix} 2 \\ 2 \\ 2 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 8 \\ 6 \\ 7 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -5 \\ -10 \\ -8 \\ 3 \end{pmatrix}$$

3) Eine Basis des Zeilenraumes: aus der neuen Matrix B jene Zeilen, die Pivotzeilen waren (also nicht Nullzeilen sind).

$$\begin{pmatrix} 0 & 2 & 0 & 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 2 & 6 & 0 & -29 & -31 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 3 & 6 \end{pmatrix}.$$

Rücksubstitutionsschritte:

$$\tilde{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1* & 3 \\ 2* & 6 & 0 & -29 & -31 \\ 0 & 0 & 1* & 3 & 6 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1* & 3 \\ 2* & 64 & 0 & 0 & 56 \\ 0 & -6 & 1* & 0 & -3 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1* & 3 \\ 1* & 32 & 0 & 0 & 28 \\ 0 & -6 & 1* & 0 & -3 \end{pmatrix}$$

4) Eine Basis des Kerns. Je eine nichtbasische Variable wird mit 1 angesetzt, die andere mit 0, und das Gleichungssystem $Cx = 0$ (äquivalent zu $Ax = 0$) durch Rücksubstitution aufgelöst.

$$\begin{pmatrix} -32 \\ 1 \\ 6 \\ -2 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -28 \\ 0 \\ 3 \\ -3 \\ 1 \end{pmatrix}$$

5) Eine andere Basis des Zeilenraumes: Die Zeilen in der neuen Matrix, die nicht Nullzeilen sind. Falls man mit der Basis weiterrechnen muss, ist diese Basis des Zeilenraumes bequemer als die in Punkt (3) gegebene, weil sie mehr Nullen enthält.

$$\begin{pmatrix} 0 & 2 & 0 & 1 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 32 & 0 & 0 & 28 \end{pmatrix}, \quad \begin{pmatrix} 0 & -6 & 1 & 0 & -3 \end{pmatrix}.$$

[13] In diesem Fall könnten wir die lineare Abhängigkeit der Spalten viel leichter feststellen: Wenn im $\mathbb{R}^m$ mehr als m Vektoren gegeben sind, dann müssen sie linear abhängig sein. Aber noch haben wir die theoretischen Grundlagen dafür nicht entwickelt.

3.4. **Eindeutigkeit der Dimension.**

3.4.1. *Austauschsatz von Steinitz.*

Lemma 3.4.1 (Austauschlemma von Steinitz)**.** *Sei $(V,+,\cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$, sei $(v_1,\cdots,v_n)$ eine Basis von V, und sei $w\in V$ mit*

$$w=\lambda_1 v_1+\cdots+\lambda_n v_n.$$

Für einen Index $k\in\{1,\cdots,n\}$ sei $\lambda_k\neq 0$. Dann ist

$$(v_1,\cdots,v_{k-1},w,v_{k+1},\cdots,v_n)$$

eine Basis von V.

Beweis. Wir zeigen zunächst, dass das n-Tupel $(v_1,\cdots,w_k,\cdots,v_n)$ linear unabhängig ist. Sei

$$\mu_1 v_1+\cdots\mu_{k-1}v_{k-1}+\mu_k w+\mu_{k+1}v_{k+1}+\cdots+\mu_n v_n=0.$$

Wir müssen zeigen, dass alle $\mu_i=0$ sind. Wir schreiben die Summe um:

$$0=\sum_{i\neq k}\mu_i v_i+\mu_k w=\sum_{i\neq k}\mu_i v_i+\mu_k\sum_{i=1}^{n}\lambda_i v_i=\sum_{i\neq k}(\mu_i+\lambda_i\mu_k)v_i+\lambda_k\mu_k v_k.$$

Da die v_i eine Basis bilden, sind sie linear unabhängig, und es folgt

$$\lambda_k\mu_k=0 \text{ und } (\forall i\neq k)\ \mu_i+\lambda_i\mu_k=0.$$

Da $\lambda_k\neq 0$, folgt nun $\mu_k=0$ und daraus $\mu_i=\mu_i+\lambda_i\mu_k=0$ für alle anderen i.
Wir zeigen nun, dass $\{v_1,\cdots,w,\cdots,v_n\}$ ganz V aufspannt. Dazu genügt es zu zeigen, dass $v_k\in\mathcal{L}(\{v_1,\cdots,w,\cdots,v_n\})$, denn dann ist die Menge $\{v_1,\cdots,v_k,\cdots,v_n\}$ in $\mathcal{L}(\{v_1,\cdots,w,\cdots,v_n\})$ enthalten, und damit ist

$$V=\mathcal{L}(\{v_1,\cdots,v_k,\cdots,v_n\})\subset\mathcal{L}(\{v_1,\cdots,w,\cdots,v_n\}).$$

Es ist aber

$$v_k=\lambda_k^{-1}\left(w-\sum_{i\neq k}\lambda_i v_i\right).$$

□

Satz 3.4.2 (Austauschsatz von Steinitz)**.** *Sei $(V,+,\cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$, sei $(v_1,\cdots,v_n)$ eine Basis von V, und $(w_1,\cdots,w_r)$ ein linear unabhängiges r-Tupel von Vektoren aus V. Dann gilt:*

1) *$r\leq n$.*
2) *Ist $r=n$, so ist $(w_1,\cdots,w_n)$ eine Basis von V.*
3) *Ist $r<n$, so gibt es Indices $i_1,\cdots,i_{n-r}$ so, dass $(w_1,\cdots,w_r,v_{i_1},\cdots,v_{i_{n-r}})$ eine Basis von V ist.*

Beweis. Vollständige Induktion nach r.
Verankerung: Für $r=1$ folgt wegen der linearen Unabhängigkeit von (w_1) dass $w_1\neq 0$. Da V also nicht der Nullraum sein kann, enthält auch die Basis $(v_1,\cdots,v_n)$ mindestens einen Vektor. Damit ist $n\geq 1$. Mit Hilfe der Basis schreibt sich $w_1=\sum_{i=1}^{n}\lambda_i v_i$ mit mindestens einem $\lambda_k\neq 0$. Nach dem Austauschlemma von Steinitz (Lemma 3.4.1) ist $(v_1,\cdots,v_{k-1},w_1,v_{k+1},\cdots,v_n)$ eine Basis von V. Im Falle, dass $n=1$ ist, bleiben beim Austausch keine v_i übrig, und es ist also (w_1) eine Basis von V.
Induktionsschritt: Wir nehmen an, der Satz gilt für r und zeigen, dass er auch für linear unabhängige Systeme $(w_1,\cdots,w_{r+1})$ gilt. Sei also $(w_1,\cdots,w_{r+1})$ ein linear

unabhängiges $(r+1)$-Tupel von Vektoren. Wir können die Induktionsvoraussetzung auf $(w_1, \cdots, w_r)$ anwenden. Daraus folgt $r \leq n$. Wäre aber $r = n$, so wäre $(w_1, \cdots, w_r)$ bereits eine Basis von V, und $(w_1, \cdots, w_{r+1})$ könnte nicht mehr linear unabhängig sein. Also ist $r < n$, das heißt, $r + 1 \leq n$. Wiederum nach der Induktionsvoraussetzung hat V eine Basis $(w_1, \cdots, w_r, v_{i_1}, \cdots, v_{i_{n-r}})$. Der Vektor w_{r+1} schreibt sich

$$w_{r+1} = \sum_{i=1}^{r} \mu_i w_i + \sum_{j=1}^{n-r} \lambda_j v_{i_j}.$$

Es muss mindestens ein $\lambda_k \neq 0$ sein, sonst wäre $w_{r+1} \in \mathcal{L}(\{w_1, \cdots, w_r\})$ im Widerspruch zur linearen Unabhängigkeit von $(w_1, \cdots, w_{r+1})$. Nach dem Austauschlemma kann nun v_{i_k} gegen w_{r+1} getauscht werden, und es ist

$$(w_1, \cdots, w_{r+1}, v_{i_1}, \cdots, v_{i_{k-1}}, v_{i_{k+1}}, \cdots, v_{i_{n-r}})$$

eine Basis. Ist $n = r + 1$, dann bleiben nach dem Austausch keine v_{i_j} in der Basis übrig, und es ist $(w_1, \cdots, w_{r+1})$ eine Basis. Damit ist der Induktionsschritt fertig. □

3.4.2. *Dimension.*

Satz 3.4.3. *Sei $(V, +, \cdot)$ ein endlich erzeugter Vektorraum über einem Körper $\mathbb{K}$ mit einer Basis $(v_1, \cdots, v_n)$. Dann besteht auch jede andere Basis von V aus genau n Elementen.*

Beweis. Sei $(w_1, \cdots, w_r)$ eine andere Basis von V. Nach Satz 3.4.2 ist $r \leq n$, weil $(w_1, \cdots, w_r)$ linear unabhängig ist. Umgekehrt ist aber $(v_1, \cdots, v_n)$ linear unabhängig und $(w_1, \cdots, w_r)$ eine Basis, und wieder wegen Satz 3.4.2 ist $n \leq r$. Also ist $r = n$. □

Definition 3.4.4. Sei $(V, +, \cdot)$ ein endlich erzeugter Vektorraum über einem Körper $\mathbb{K}$ mit einer Basis $(v_1, \cdots, v_n)$.
Dann heißt n die Dimension von V ($\dim(V) := n$).

Bemerkung 3.4.5. Eine Definition muss man nicht beweisen, aber man muss sicher stellen, dass sie eindeutig ihr Objekt festlegt. Deshalb mussten wir erst Satz 3.4.3 haben, um sicher zu gehen, dass die Dimension wirklich nur von V und nicht von der speziellen Wahl der Basis abhängt! Auch wissen wir bereits (Satz 3.2.13) dass jeder endlich erzeugte Vektorraum mindestens eine Basis besitzt.

Bemerkung 3.4.6. Die Dimension des Nullraumes $\{0\}$ ist Null. Die Dimension des $\mathbb{K}^n$ ist n.

Satz 3.4.7. *Ist $(V, +, \cdot)$ ein endlich erzeugter Vektorraum über einem Körper $\mathbb{K}$, und W ein beliebiger Unterraum von V.*
Dann ist auch W endlich erzeugt, und es gilt $\dim(W) \leq \dim(V)$. *Wenn* $\dim(W) = \dim(V)$, *dann ist $W = V$.*

Beweis. Sei $n = \dim(V)$, also besitzt V eine Basis von n Vektoren. Nach Satz 3.4.2 kann kein linear unabhängiges Vektorsystem in V (und somit auch keines in W) mehr als n Vektoren enthalten. Damit muss aber W ein maximal linear unabhängiges System, also eine Basis, mit höchstens n Vektoren haben. Hat diese Basis von W genau n Vektoren, so ist sie nach Satz 3.4.2 auch eine Basis von V, und somit ist $W = V$. □

3.4.3. *Basisergänzungssatz.*

Satz 3.4.8 (Basisergänzungssatz)**.** *Sei $(V, +, \cdot)$ ein endlich erzeugter Vektorraum über einem Körper $\mathbb{K}$, sei $(w_1, \cdots, w_m)$ ein linear unabhängiges m-Tupel von Vektoren aus V. Wenn $(w_1, \cdots, w_m)$ nicht selbst Basis ist, dann gibt es Vektoren $w_{m+1}, \cdots, w_n \in V$ so, dass $(w_1, \cdots, w_n)$ eine Basis von V ist.*

Beweis. Beweis 1: Da V endlich erzeugt ist, besitzt V eine Basis $(v_1, \cdots, v_n)$. Nach dem Austauschsatz von Steinitz (Satz 3.4.2) lassen sich aus dieser Basis m Vektoren gegen die Vektoren $w_1, \cdots, w_m$ austauschen.
Beweis 2: Wir bringen diesen Beweis, weil er gleichzeitig eine bequeme Methode zeigt, ein lineares System zu einer Basis aufzufetten. Da V endlich erzeugt ist, besitzt V eine Basis aus endlich vielen Vektoren (Satz 3.2.13). Wir setzen $n := \dim(V)$. Jedes linear unabhängige System hat höchstens n Vektoren (Satz 3.4.2). Erzeugt das System $(w_1, \cdots, w_m)$ bereits den ganzen Raum V, so ist $m = n$ und $(w_1, \cdots, w_m)$ ist eine Basis. Andernfalls gibt es mindestens einen Vektor $w_{m+1} \notin \mathcal{L}(\{w_1, \cdots, w_m\})$. Wir zeigen, dass das System $(w_1, \cdots, w_{m+1})$ linear unabhängig ist. Wäre es nicht linear unabhängig, gäbe es $\lambda_1, \cdots, \lambda_m, \lambda_{m+1} \in \mathbb{K}$, nicht alle gleich null, so dass $\sum_{i=1}^{m+1} \lambda_i w_i = 0$. Wäre $\lambda_{m+1} = 0$, könnte man den Nullvektor durch eine nichttriviale Linearkombination von $w_1, \cdots, w_m$ schreiben, und $(w_1, \cdots, w_m)$ wäre linear abhängig. Also ist $\lambda_{m+1} \neq 0$ und $w_{m+1} = -\frac{1}{\lambda_{m+1}} \sum_{i=1}^{m} \lambda_i w_i$. Dann wäre aber $w_{m+1} \in \mathcal{L}(\{w_1, \cdots, w_m\})$, im Widerspruch zur Wahl von w_{m+1}. Nun haben wir also das linear unabhängige System um einen Vektor erweitert. Wir wiederholen diesen Schritt so oft, bis wir eine Basis erreicht haben. Da ein linear unabhängiges System höchstens n Vektoren enthält, muss das Verfahren abbrechen, wenn wir n Vektoren gesammelt haben. □

3.4.4. *Rang einer Matrix.*

Definition 3.4.9. Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{m \times n}$. Der Rang[14] von A (geschrieben $\rho(A)$ oder $\mathrm{rk}(A)$) ist die Dimension des Spaltenraumes von A.

Satz 3.4.10. *Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{m \times n}$, sei $\rho(A)$ der Rang von A. Dann gilt:*

1) *Der Rang wird durch elementare Zeilenumformungen nicht verändert.*
2) *Wird A durch elementare Zeilentransformationen in eine Matrix $\tilde{A}$ in verallgemeinerter Zeilenstufenform umgewandelt, so ist $\rho(A)$ die Zahl der Zeilen in $\tilde{A}$, welche nicht Nullzeilen sind, zugleich die Zahl der basischen Spalten.*
3) *$\rho(A)$ ist zugleich die Dimension des Zeilenraumes von A (man sagt: "Zeilenrang ist Spaltenrang").*
4) *$n - \rho(A)$ ist die Dimension des Kernes von A.*
5) *Es gilt $\rho(A) \leq \min(m, n)$.*

Beweis. 1.) Wir nehmen an, dass A durch eine elementare Zeilenumformung in $\tilde{A}$ umgeformt wird. Seien $s_1, \cdots, s_n$ die Spalten von A, $\tilde{s}_1, \cdots, \tilde{s}_n$ die Spalten von $\tilde{A}$, und sei $r = \rho(A)$. Es gibt also eine Basis $(s_{j_1}, \cdots, s_{j_r})$ des Spaltenraumes von A aus r Vektoren. Nach Satz 3.3.14, Punkte (3) und (5) ist $(\tilde{s}_{j_1}, \cdots, \tilde{s}_{j_r})$ eine Basis des Spaltenraumes von $\tilde{A}$, und somit hat auch $\mathrm{rg}(\tilde{A})$ die Dimension r.
2.) Sei r die Anzahl der basischen Spalten in $\tilde{A}$. Nach Definition der verallgemeinerten Zeilenstufenform ist r gleichzeitig die Anzahl der Zeilen von $\tilde{A}$, die nicht

[14] englisch: rank, nicht zu verwechseln mit range!

Null sind. Nach Satz 3.3.18 bilden die basischen Spalten von $\tilde{A}$ eine Basis des Spaltenraumes von $\tilde{A}$, und die Nicht-Nullzeilen von $\tilde{A}$ eine Basis des Zeilenraumes. Also ist sowohl die Dimension des Spaltenraumes als auch des Zeilenraumes von $\tilde{A}$ gleich r. Da nach Punkt (1) der Rang durch die elementaren Zeilenumformungen nicht verändert wird, ist auch $\rho(A) = r$.
3.) Wir knüpfen gleich an Punkt (2) an. Nach Satz 3.3.14 (6) haben $\tilde{A}$ und A denselben Zeilenraum. Daher ist die Dimension des Zeilenraumes von A ebenfalls gleich r.
4.) Wieder greifen wir auf die verallgemeinterte Zeilenstufenform $\tilde{A}$ zurück. (Dann ist, wie wir jetzt wissen, $r = \rho(A)$.) Wieder nach Satz 3.3.14 (2) ist $\ker(\tilde{A}) = \ker(A)$ und besitzt eine Basis von $n - r$ Vektoren. Also ist $\dim(\ker(A)) = n - r$.
5.) Da der Zeilenraum von A von m Zeilen erzeugt wird, aus denen man nach Satz 3.2.13 eine Basis auswählen kann, kann seine Dimension nicht größer als m sein. Da der Spaltenraum von A von n Spalten erzeugt wird, kann seine Dimension nicht größer als n sein. □

Bemerkung 3.4.11. Man bestimmt den Rang einer Matrix A, indem man sie durch Pivotschritte auf verallgemeinerte Zeilenstufenform bringt, und zählt, wieviele Zeilen nicht Null geworden sind (zugleich die Anzahl der Pivotelemente).

Satz 3.4.12. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{m\times n}$, $b \in \mathbb{K}^m$. Dann ist das lineare Gleichungssystem $Ax = b$ genau dann lösbar, wenn die Systemmatrix A und die erweiterte Matrix (A, b) denselben Rang haben.*

Beweis. Da der Spaltenraum von (A, b) alle Spalten von A enthält, enthält er auch den ganzen Spaltenraum von A als Unterraum. Nach Satz 3.4.7 sind die beiden Räume gleich, wenn ihre Dimensionen gleich sind. Sind nun die Dimensionen gleich, so ist $b \in \mathrm{rg}((A, b)) = \mathrm{rg}(A)$ und damit ist $Ax = b$ lösbar (siehe Bemerkung 3.3.10). Sind die Dimensionen verschieden, so kann b nicht im Spaltenraum von A liegen, und das Gleichungssystem ist unlösbar. □

Dieser Satz darf in keiner Einführung über lineare Algebra fehlen. In der Praxis konnte ich ihn allerdings nie brauchen. Die Rangbestimmung geht durch Pivotschritte, und damit kann man auch gleich ausprobieren, ob das Gleichungssystem lösbar ist, und gegebenenfalls die Lösung ausrechnen.

3.5. **Summenräume und Dimensionsformel.**

3.5.1. *Summenraum.*

Wir wissen, dass der Durchschnitt zweier Unterräume wieder ein Unterraum ist. Dagegen ist die Vereinigung nur dann ein Unterraum, wenn einer der Unterräume im anderen enthalten ist. Wenn wir zwei oder mehrere Unterräume zu einem Unterraum verbinden wollen, bietet sich der Begriff des Summenraumes an. — An dieser Stelle einige Zeichungen in $\mathbb{R}^2$ und $\mathbb{R}^3$ zur Vorbereitung des Begriffes.

Definition 3.5.1. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$. Für $i = 1, \cdots, n$ sei W_i jeweils ein Unterraum von V. Wir definieren den Summenraum

$$W_1 + \cdots + W_n := \{w_1 + \cdots + w_n \mid (\forall i = 1 \cdots n)\ w_i \in W_i\}.$$

Satz 3.5.2. *Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$. Seien $W_1, \cdots, W_n$ Unterräume von V. Dann gilt:*

1. *Der Summenraum $W_1 + \cdots + W_n$ ist ein Unterraum von V.*
2. *Für jedes $i = 1, \cdots, n$ sei M_i jeweils ein Erzeugendensystem von W_i. Dann ist die Vereinigung $\bigcup_{i=1}^n M_i$ ein Erzeugendensystem von $W_1 + \cdots + W_n$.*

(3) *Der Summenraum $W_1 + \cdots + W_n$ ist der kleinste Unterraum, der alle W_i enthält.*

Beweis. Punkt (1)*:* Seien $x, y \in W_1 + \cdots + W_n$, $\alpha, \beta \in \mathbb{K}$. Wir zeigen: $\alpha x + \beta y \in W_1 + \cdots + W_n$ (das reicht nach Satz 3.1.6). Nach Definition des Summenraumes lassen sich x, y in folgender Weise schreiben:

$$x = \sum_{i=1}^{n} w_i, \quad y = \sum_{i=1}^{n} \tilde{w}_i, \quad \text{mit } w_i, \tilde{w}_i \in W_i.$$

Dann ist

$$\alpha x + \beta y = \sum_{i=1}^{n} (\alpha w_i + \beta \tilde{w}_i).$$

Da die W_i Unterräume von V sind, ist $\alpha w_i + \beta \tilde{w}_i \in W_i$, und somit ist $\alpha x + \beta y \in W_1 + \cdots + W_n$.
Punkt (2)*:* Sei nun U ein Unterraum von V, welcher alle M_i enthält. Wir müssen zeigen, dass $W_1 + \cdots + W_n$ ein Unterraum von U ist. Zunächst enthält U jedes M_i, und damit auch $\mathcal{L}(M_i) = W_i$. Ist $x \in W_1 + \cdots + W_n$, so lässt sich x schreiben als $x = \sum_{i=1}^{n} w_i$ mit $w_i \in W_i$. Da alle W_i in U enthalten sind, sind alle $w_i \in U$, und weil U ein Unterraum ist, ist damit auch $\sum_{i=1}^{n} w_i \in U$. Somit ist $W_1 + \cdots + W_n \subset U$.
Punkt (3)*:* Setze in Punkt (2) einfach $M_i := W_i$. □

Achtung: Wenn $(w_{i,1}, \cdots, w_{i,r_i})$ jeweils eine Basis von W_i ist, folgt daraus noch nicht, dass $(w_{1,1}, \cdots, w_{n,r_n})$ eine Basis des Summenraumes ist. Zwar haben wir noch ein Erzeugendensystem des Summenraumes, aber dieses könnte linear abhängig sein. — Weiteres im Abschnitt über direkte Summen weiter unten.

3.5.2. *Dimensionsformel.*

Satz 3.5.3 (Dimensionsformel)**.** *Sei $(V, +, \cdot)$ ein endlich erzeugter Vektorraum über einem Körper $\mathbb{K}$. Seien A, B zwei Unterräume von V. Dann gilt*

$$\dim(A + B) + \dim(A \cap B) \;=\; \dim(A) + \dim(B).$$

Beweis. Da V endlich erzeugt ist, ist jeder Unterraum von V endlich erzeugt und besitzt eine Basis aus endlich vielen Vektoren. Wir beginnen mit einer Basis $(x_1, \cdots, x_r)$ von $A \cap B$. Weil $A \cap B$ sowohl in A als auch in B enthalten ist, ist $(x_1, \cdots, x_r)$ ein linear unabhängiges Vektorsystem sowohl in A als auch B, und läßt sich nach dem Basisergänzungssatz 3.4.8 zu je einer Basis von A und B erweitern. Wir haben also die folgenden Basen und Dimensionen:

$$\begin{array}{lll} (x_1, \cdots, x_r) & \text{Basis von } A \cap B; & \dim(A \cap B) = r \\ (x_1, \cdots, x_r, a_1, \cdots, a_s) & \text{Basis von } A \quad ; & \dim(A) = r + s \\ (x_1, \cdots, x_r, b_1, \cdots, b_t) & \text{Basis von } B \quad ; & \dim(B) = r + t \end{array}$$

Wir zeigen nun, dass $(x_1, \cdots, x_r, a_1, \cdots, a_s, b_1, \cdots, b_t)$ eine Basis von $A + B$ ist. Nach Satz 3.5.2 ist dieses System jedenfalls ein Erzeugendensystem von $A + B$, da es Vereinigung einer Basis von A mit einer Basis von B ist. Wir zeigen die lineare Unabhängigkeit. Es sei also

$$\sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j a_j + \sum_{k=1}^{t} \nu_k b_k = 0.$$

Wir schreiben das in der Form

$$\sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j a_j = -\sum_{k=1}^{t} \nu_k b_k.$$

Da alle $x_i \in A$, $a_j \in A$, ist $\sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^s \mu_j a_j \in A$. Andererseits ist $-\sum_{k=1}^t \nu_k b_k \in B$. Also ist $-\sum_{k=1}^t \nu_k b_k \in A \cap B$. Weil aber $(x_1, \cdots, x_r)$ eine Basis von $A \cap B$ ist, muss es jedenfalls eine Darstellung der Form geben (mit geeigneten Skalaren ξ_i):

$$-\sum_{k=1}^t \nu_k b_k = \sum_{i=1}^r \xi_i x_i.$$

Das System $(x_1, \cdots, x_r, b_1, \cdots, b_t)$ ist aber eine Basis von B. Wir können also denselben Vektor auf zwei Arten als Linearkombination dieser Basis schreiben. Wegen der linearen Unabhängigkeit der Basis muss es sich in Wirklichkeit um dieselbe Linearkombination handeln. Die erste Darstellung enthält keinen der Basisvektoren x_i, die zweite enthält keinen der Basisvektoren b_k. Folglich sind alle Koeffizienten ν_k und ξ_i gleich Null.

$$\sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^s \mu_j a_j = -\sum_{k=1}^t \nu_k b_k = 0,$$

und wegen der linearen Unabhängigkeit von $(x_1, \cdots, x_r, a_1, \cdots, a_s)$ sind auch alle λ_i und alle μ_j gleich Null. Also ist das System $(x_1, \cdots, x_r, a_1, \cdots, a_s, b_1, \cdots, b_t)$ linear unabhängig, und somit ist es eine Basis von $A + B$. Insbesondere ist $\dim(A + B) = r + s + t$.
Nun können wir die Dimensionsformel zusammensetzen:

$$\dim(A + B) + \dim(A \cap B) = (r + s + t) + r = (r + s) + (r + t) = \dim(A) + \dim(B).$$

□

3.5.3. *Direkte Summen innerhalb eines Vektorraums.*

Definition 3.5.4. Sei $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$. Seien $W_1, \cdots, W_n$ Unterräume von V. Der Summenraum $W_1 + \cdots + W_n$ heißt direkte Summe von $W_1, \cdots, W_n$ wenn gilt:
Für jedes $x \in W_1 + \cdots + W_n$ gibt es eindeutig bestimmte $w_i \in W_i$ sodass $x = \sum_{i=1}^n w_i$.

Schreibweise 3.5.5. Manche Autoren schreiben für die Summe von Unterräumen innerhalb eines Vektorraumes, wenn sie direkt ist, $W_1 \oplus \cdots \oplus W_n$. In diesem Skriptum werde ich diese Schreibweise vermeiden, um nicht Konfusion mit der äußeren direkten Summe zu schaffen.

Bemerkung 3.5.6. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$. Sei $W = W_1 + \cdots + W_n$ die direkte Summe von Unterräumen $W_i \subset V$. Dann ist für jedes $k = 1, \cdots, n$ die folgende Abbildung wohldefiniert:

$$\pi_k : \begin{cases} W & \to W_k, \\ \sum_{i=1}^n w_i & \mapsto w_k. \end{cases}$$

Diese Abbildung heißt die Projektion des Summenraumes auf den Teilraum W_k.

Beweis. Nach Definition des Summenraumes läßt sich jedes $w \in W$ als die Summe $\sum_{i=1}^n w_i$ schreiben. Weil die Summe direkt ist, ist diese Zerlegung von w eindeutig, sodass die obige Definition von π_k eine eindeutige Abbildung festlegt. □

Satz 3.5.7. *Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper $\mathbb{K}$. Seien $W_1, \cdots, W_n$ Unterräume von V. Dann gilt:*

1) *Ist die Summe $W_1 + \cdots + W_n$ direkt, dann gilt: Jedes System von Vektoren $(w_1, \cdots, w_n)$ mit $w_i \in W_i \setminus \{0\}$ ist linear unabhängig. Falls $W_i \neq \{0\}$ für alle i, so gilt auch die Umkehrung.*
2) *Ist $n = 2$, so gilt: Die Summe $W_1 + W_2$ ist genau dann direkt, wenn gilt $W_1 \cap W_2 = \{0\}$.*
3) *Ist V endlich erzeugt, so gilt: Die Summe $W_1 + \cdots + W_n$ ist genau dann direkt, wenn gilt: Sind $(w_{i,1}, \cdots, w_{i,r_i})$ jeweils Basen von W_i, so ist das System $(w_{1,1}, \cdots, w_{n,r_n})$ eine Basis von $W_1 + \cdots + W_n$.*
4) *Ist V endlich erzeugt, so ist die Summe der W_i genau dann direkt, wenn gilt:*

$$\dim(W_1 + \cdots + W_n) = \sum_{i=1}^{n} \dim(W_i).$$

Beweis. Punkt (1): Sei die Summe direkt, seien $w_i \in W_i$, alle ungleich Null. Wir zeigen, dass $(w_1, \cdots, w_n)$ linear unabhängig ist. Sei also $\sum_{i=1}^n \lambda_i w_i = 0$. Da die Summe direkt ist, läßt sich der Nullvektor nur in einer eindeutigen Form $0 = \sum_{i=1}^n \tilde{w}_i$ mit $\tilde{w}_i \in W_i$ zerlegen, und natürlich geschieht das durch $\tilde{w}_i = 0$. Also muss für jedes i gelten: $\lambda_i w_i = 0$. Weil aber $w_i \neq 0$, folgt $\lambda_i = 0$ für alle i.
Umgekehrt sei nun jedes n-Tupel von Vektoren $w_i \in W_i \setminus \{0\}$ linear unabhängig. Wir zeigen, dass die Summe direkt ist, also das jedes $w \in W_1 + \cdots + W_n$ sich nur auf eine Weise in Vektoren aus W_i zerlegen lässt. Sei also $w = \sum_{i=1}^n u_i = \sum_{i=1}^n v_i$ mit $u_i, v_i \in W_i$. Dann ist $0 = \sum_{i=1}^n (u_i - v_i)$. Wir setzen:

$$\begin{cases} w_i := u_i - v_i & \text{falls } u_i - v_i \neq 0, \\ w_i \in W_i \setminus \{0\} & \text{beliebig, falls } u_i - v_i = 0. \end{cases}$$

$$\lambda_i := \begin{cases} 1 & \text{falls } u_i - v_i \neq 0, \\ 0 & \text{falls } u_i - v_i = 0. \end{cases}$$

Dann ist $\sum_{i=1}^n \lambda_i w_i = 0$, und da alle $w_i \neq 0$, gilt wegen der linearen Unabhängigkeit $\lambda_i = 0$ für alle i. Nach Definition von λ_i muss also für alle i gelten: $u_i = v_i$.

Punkt (2): Sei $W_1 + W_2$ direkt, und $u \in W_1 \cap W_2$. Dann läßt sich u als Summe $u = w_1 + w_2$ mit $w_i \in W_i$ auf die folgenden Weisen zerlegen: $u = u + 0 = 0 + u$. Da die Zerlegung aber nach Voraussetzung eindeutig ist, folgt $u = 0$.
Umgekehrt sei $W_1 \cap W_2 = \{0\}$ und $u = w_1 + w_2 = v_1 + v_2$ mit $v_i, w_i \in W_i$. Dann ist $w_1 - v_1 = v_2 - w_2$. Nun ist $w_1 - v_1 \in W_1$, derselbe Vektor $v_2 - w_2 \in W_2$. Daher ist $w_1 - v_1 \in W_1 \cap W_2$, also $w_1 - v_1 = 0$, und ebenso $v_2 - w_2 = 0$. Also ist die Zerlegung eindeutig: $v_1 = w_1$, $v_2 = w_2$.
Punkt (3): Sei die Summe direkt. Wir zeigen: $(w_{1,1}, \cdots, w_{n,r_n})$ ist linear unabhängig. Da dieses Vektorsystem den Summenraum erzeugt, ist es dann eine Basis. Sei also

$$\sum_{i=1}^{n} \sum_{j=1}^{r_i} \lambda_{i,j} w_{i,j} = 0.$$

Da die Summe direkt ist, läßt sich der Nullvektor nur auf eine Weise zerlegen, und es muss also gelten:

$$\sum_{j=1}^{r_i} \lambda_{i,j} w_{i,j} = 0 \text{ für alle } i.$$

Da aber $(w_{i,1}, \cdots, w_{i,r_i})$ eine Basis von W_i, und daher linear unabhängig ist, folgt $\lambda_{i,j} = 0$ für alle i, j.
Sei umgekehrt $(w_{1,1}, \cdots, w_{n,r_n})$ eine Basis des Summenraumes. Wir zeigen, dass die Summe direkt ist. Sei $x = \sum v_i = \sum u_i$ mit $u_i, v_i \in W_i$. Es lassen sich u_i, v_i

mit Hilfe der Basen schreiben:

$$u_i = \sum_{j=1}^{r_i} \lambda_{i,j} w_{i,j}, \quad v_i = \sum_{j=1}^{r_i} \mu_{i,j} w_{i,j}.$$

Dann ist

$$x = \sum_{i=1}^{n} \sum_{j=1}^{r_i} \lambda_{i,j} w_{i,j} = \sum_{i=1}^{n} \sum_{j=1}^{r_i} \mu_{i,j} w_{i,j}.$$

Wegen der linearen Unabhängigkeit der $w_{i,j}$ folgt $\lambda_{i,j} = \mu_{i,j}$ für alle i, j, und folglich $u_i = v_i$ für alle i.
Punkt (4): Direkte Folge aus Punkt (3). □

Satz 3.5.8. *Sei $(V, +, \cdot)$ ein endlich erzeugter Vektorraum über einem Körper $\mathbb{K}$ und U ein Unterraum von V. Dann gibt es mindestens einen Unterraum W von V sodass $V = U + W$, wobei die Summe direkt ist.*

Beweis. Sei $u_1, \cdots, u_m$ eine Basis von U. Nach dem Basisergänzungssatz 3.4.8 läßt sich diese Basis zu einer Basis $(u_1, \cdots, u_m, v_1, \cdots, v_{n-m})$ von V erweitern. Setze nun $W := \mathcal{L}(\{v_1, \cdots, v_{n-m}\})$. Nach Satz 3.5.7 (3) ist die Summe $U + W$ direkt, und offensichtlich[15] ist $U + W = V$. □

Dieser Satz gilt auch in unendlich dimensionalen Vektorräumen, man muss nur die Theorie von unendlichen Basen beherrschen. Allerdings interessiert man sich in unendlich dimensionalen Vektorräumen zumeist für abgeschlossene Unterräume. (Was eine abgeschlossene Menge ist, lernen Sie in Analysis.)

3.6. Äußere direkte Summe.

Manchmal will man aus zwei Vektorräumen, die nicht von vorneherein Teilräume eines gemeinsamen Oberraumes sind, eine Summe konstruieren.

Definition 3.6.1. Seien $(V_1, +_1, \cdot_1)$ und $(V_2, +_2, \cdot_2)$ zwei Vektorräume über demselben Körper $\mathbb{K}$. Die (äußere) direkte Summe von V_1 und V_2 wird definiert als Produktmenge:

$$V_1 \oplus V_2 := \{(v_1, v_2) \mid v_i \in V_i\},$$

mit folgenden Rechenoperationen für $(x_1, x_2), (y_1, y_2) \in V_1 \oplus V_2$, $\alpha \in \mathbb{K}$.

$$\begin{aligned} (x_1, x_2) + (y_1, y_2) &:= (x_1 +_1 y_1, x_2 +_2 y_2), \\ \alpha(x_1, x_2) &:= (\alpha x_1, \alpha x_2). \end{aligned}$$

Durch die injektiven Abbildungen

$$j_1 : \begin{cases} V_1 & \to V_1 \oplus V_2, \\ x_1 & \mapsto (x_1, 0), \end{cases} \quad j_2 : \begin{cases} V_2 & \to V_1 \oplus V_2, \\ x_2 & \mapsto (0, x_2), \end{cases}$$

betten wir V_1 und V_2 in den Summenraum ein. (d.h. wir identifizieren $x_1 \in V_1$ mit dem Vektor $j_1(x_1) \in V_1 \oplus V_2$, ebenso wird $x_2 \in V_2$ mit $j_2(x_2) \in V_1 \oplus V_2$ identifiziert.) Die Abbildungen

$$\pi_1 : \begin{cases} V_1 \oplus V_2 & \to V_1, \\ (x_1, x_2) & \mapsto x_1, \end{cases} \quad \pi_2 : \begin{cases} V_1 \oplus V_2 & \to V_2, \\ (x_1, x_2) & \mapsto x_2, \end{cases}$$

heißen die kanonischen Projektionen von $V_1 \oplus V_2$ auf V_1 bzw. V_2.

Bemerkung 3.6.2. Diese Definition läßt sich analog ohne Schwierigkeiten auf die Summe von mehr als 2 Vektorräumen verallgemeinern.

[15] Oder ist es falsch und ich bin nur zu faul, einen richtigen Beweis zu überlegen? Überprüfen Sie selbst.

Schreibweise 3.6.3. Manchmal wird die äußere direkte Summe als direktes Produkt bezeichnet (ein Unterschied zwischen direkter Summe und direktem Produkt tritt erst auf, wenn man unendlich viele Vektorräume verbindet). Auch die Schreibweise $V_1 \otimes V_2$ kommt vor. Sie ist nicht ganz unproblematisch, weil es ein (völlig anderes) sogenanntes Tensorprodukt gibt, und auch dieses wird von manchen Autoren als $V_1 \otimes V_2$ bezeichnet. Facit: Schauen Sie genau, und erfassen Sie aus dem Zusammenhang, was der Autor meint!

Satz 3.6.4. *Seien $(V_1, +_1, \cdot_1)$ und $(V_2, +_2, \cdot_2)$ zwei Vektorräume über demselben Körper $\mathbb{K}$, und sei $V_1 \oplus V_2$ deren äußere direkte Summe.*

1) *$V_1 \oplus V_2$ ist ein Vektorraum über $\mathbb{K}$.*
2) *Identifiziert man V_i mit $j_i(V_i)$, so sind V_1, V_2 Unterräume von $V_1 \oplus V_2$, und $V_1 \oplus V_2$ ist auch ihre direkte Summe im Sinne einer direkten Summe innerhalb eines Vektorraums.*

Beweis. Punkt (1): Der Beweis erfolgt durch Nachrechnen aller Rechenregeln.
Punkt (2): Jedes Paar $(x_1, x_2) \in V_1 \oplus V_2$ läßt sich in eindeutiger Weise in zwei Vektoren aus $j_1(V_1)$ und $j_2(V_2)$ zerlegen: $(x_1, x_2) = (x_1, 0) + (0, x_2)$. Weil es eine solche Zerlegung gibt, ist $V_1 \oplus V_2 = j_1(V_1) + j_2(V_2)$, und weil die Zerlegung eindeutig ist, ist diese Summe direkt. □

Man kann auch äußere direkte Summen von unendlich vielen Vektorräumen definieren. Hierbei unterscheidet man die direkte Summe und das direkte Produkt:

$$\bigoplus_{i=1}^{\infty} V_i := \{(v_i)_{i=1,\cdots,\infty} \mid v_i \in V_i \text{ mit } v_i \neq 0 \text{ für höchstens endlich viele } i\},$$

$$\bigotimes_{i=1}^{\infty} V_i := \{(v_i)_{i=1,\cdots,\infty} \mid v_i \in V_i\}.$$

Wir werden diese, durchaus sehr nützliche, Konstruktion in dieser Vorlesung aber nicht weiter brauchen.

4. Lineare Abbildungen

4.1. **Definitionen und Beispiele linearer Abbildungen.**

4.1.1. *Lineare Abbildung.*

Definition 4.1.1. Seien $(V, +_1, \cdot_1)$ und $(W, +_2, \cdot_2)$ Vektorräume über dem selben Körper $\mathbb{K}$. Eine Abbildung $f : V \to W$ heißt linear, wenn die folgenden Eigenschaften gelten:

$$(\forall x, y \in V)\ f(x +_1 y) = f(x) +_2 f(y),$$
$$(\forall x \in V,\ \alpha \in \mathbb{K})\ f(\alpha \cdot_1 x) = \alpha \cdot_2 f(x).$$

Eine lineare Abbildung heißt auch (Vektorraum)-Homomorphismus. Die Menge der linearen Abbildungen $f : V \to W$ bezeichnen wir mit $\mathrm{Hom}(V, W)$.
Eine bijektive lineare Abbildung $f : V \to W$ heißt (Vektorraum)-Isomorphismus. Existiert ein Isomorphismus $f : V \to W$, so heißt V zu W isomorph.

Schreibweise 4.1.2. Für die Menge der linearen Abbildungen von V nach W wird auch oft die Schreibweise $\mathcal{L}(V, W)$ verwendet, $\mathcal{L}(V)$ ist dann die Menge der linearen Abbildungen von V nach V. Autoren, die diese Schreibweise verwenden, müssen natürlich ein anderes Zeichen für die lineare Hülle einer Menge M verwenden, zum Beispiel $\mathrm{span}(M)$.
Wir werden künftig die Zeichen für $+$ und $\cdot$ in verschiedenen Vektorräumen nicht

mehr unterscheiden. Aus dem Zusammenhang wird ja klar, in welchem Raum gerade gerechnet wird.

Beispiel 4.1.3. Folgende Abbildungen sind Beispiele von linearen Abbildungen:

1) $$\begin{cases} \mathbb{K}^n & \to \mathbb{K}^m \\ x & \mapsto Ax \end{cases} \quad \text{mit einer Matrix } A \in \mathbb{K}^{m\times n}.$$
2) Drehungen, Streckungen, Projektionen in $\mathbb{R}^2$ oder $\mathbb{R}^3$: durch Zeichnungen illustrieren.
3) Sei Ω eine Menge und $\mathcal{F} := \{f \mid f : \Omega \to \mathbb{R} \text{ ist Funktion}\}$ (mit der in Funktionenräumen üblichen Addition und Multiplikation mit Skalaren). Sei $\omega \in \Omega$. Dann ist folgende Abbildung linear:
$$\begin{cases} \mathcal{F} & \to \mathbb{R}, \\ f & \mapsto f(\omega). \end{cases}$$
4) Sei $\mathcal{C}([0,1];\mathbb{R})$ der Raum der stetigen Funktionen $f : [0,1] \to \mathbb{R}$ und sei $\mathcal{C}^1([0,1];\mathbb{R})$ der Raum der stetig differenzierbaren Funktionen $f : [0,1] \to \mathbb{R}$, dann ist die Ableitung eine lineare Abbildung:
$$\begin{cases} \mathcal{C}^1([0,1];\mathbb{R}) & \to \mathcal{C}([0,1];\mathbb{R}), \\ f & \mapsto f'. \end{cases}$$
(Auch die höheren Ableitungen und partielle Ableitungen sind linear.)
5) Ist $g \in \mathcal{C}([0,1];\mathbb{R})$, dann ist die folgende Abbildung linear:
$$\begin{cases} \mathcal{C}([0,1];\mathbb{R}) & \to \mathbb{R}, \\ f & \mapsto \int_0^1 g(x)f(x)\,dx. \end{cases}$$
6) Sei $\ell^1 := \{(\alpha_i)_{i=1,\cdots,\infty} \mid \alpha_i \in \mathbb{R},\ \sum_{i=1}^{\infty} |\alpha_i| < \infty\}$, mit den Rechenvorschriften
$$(\alpha_i)_{i=1,\cdots,\infty} + (\beta_i)_{i=1,\cdots,\infty} := (\alpha_i + \beta_i)_{i=1,\cdots,\infty},$$
$$\lambda \cdot (\alpha_i)_{i=1,\cdots,\infty} := (\lambda\alpha_i)_{i=1,\cdots,\infty}.$$
Dann ist die folgende Abbildung linear:
$$\begin{cases} \ell^1 & \to \mathbb{R}, \\ (\alpha_i)_{i=1,\cdots,\infty} & \mapsto \sum_{i=1}^{\infty} \alpha_i. \end{cases}$$
7) Sei V ein Vektorraum über $\mathbb{K}$ und V_1, V_2 Unterräume von V so, dass die Summe $V_1 + V_2$ direkt ist. Dann sind die Projektionen $\pi_i : V_1 + V_2 \to V_i$ linear.
8) Sei $\mathbb{K}$ ein Körper und $T : \mathbb{K}^{m\times n} \to \mathbb{K}^{m\times n}$ eine elementare Zeilenumformung. Dann ist T eine lineare Abbildung.

Satz 4.1.4. *Seien $(V,+,\cdot)$ und $(W,+,\cdot)$ Vektorräume über dem selben Körper $\mathbb{K}$. Sei $f : V \to W$ eine Abbildung. Dann ist f genau dann linear wenn gilt:*

$$(\forall x, y \in V)\ (\forall \alpha, \beta \in \mathbb{K})\ f(\alpha x + \beta y) = \alpha f(x) + \beta f(y).$$

Beweis. Leicht zu sehen. Vergleichen Sie Satz 3.1.6. □

4.1.2. *Hintereinanderausführung linearer Abbildungen.*

Satz 4.1.5. *Seien $(U,+,\cdot)$, $(V,+,\cdot)$ und $(W,+,\cdot)$ Vektorräume über dem selben Körper $\mathbb{K}$.*

1) *Ist $f \in \mathrm{Hom}(U,V)$ und $g \in \mathrm{Hom}(V,W)$, dann ist die Hintereinanderausführung $g \circ f \in \mathrm{Hom}(U,W)$.*
2) *Ist $f \in \mathrm{Hom}(U,V)$ ein Isomorphismus, so ist auch die Umkehrabbildung $f^{-1} : V \to U$ ein Isomorphismus.*
3) *Die Identität $\mathrm{id}_U : U \to U$ ist ein Isomorphismus.*

Beweis. Punkt (1): Seien $x, y \in U$ und $\alpha, \beta \in \mathbb{K}$. Wenn f und g linear sind, dann gilt

$$\begin{aligned}[g \circ f](\alpha x + \beta y) &= g\left(f\left(\alpha x + \beta y\right)\right) = g\left(\alpha f\left(x\right) + \beta f\left(y\right)\right) \\ &= \alpha g\left(f\left(x\right)\right) + \beta g\left(f\left(y\right)\right) = \alpha[g \circ f](x) + \beta[g \circ f](y).\end{aligned}$$

Punkt (2): Seien $\tilde{x}, \tilde{y} \in V$ und $x = f^{-1}(\tilde{x})$, $y = f^{-1}(\tilde{y})$. Seien $\alpha, \beta \in \mathbb{K}$. Wir müssen zeigen: $f^{-1}(\alpha\tilde{x} + \beta\tilde{y}) = \alpha f^{-1}(\tilde{x}) + \beta f^{-1}(\tilde{y})$. Das bedeutet dasselbe wie $f(\alpha x + \beta y) = \alpha\tilde{x} + \beta\tilde{y}$. Das folgt aber unmittelbar aus der Linearität von f, weil $f(x) = \tilde{x}$ und $f(y) = \tilde{y}$.
Punkt (3): Als Übung der geneigten Leserschaft anempfohlen. □

Korollar 4.1.6. *Seien $(U,+,\cdot)$, $(V,+,\cdot)$ und $(W,+,\cdot)$ Vektorräume über dem selben Körper $\mathbb{K}$. Dann gilt:*

1) *Ist U isomorph zu V und V isomorph zu W, dann ist U isomorph zu W.*
2) *Ist U isomorph zu V, so ist V isomorph zu U.*
3) *U ist isomorph zu sich selbst.*

Beweis. Satz 4.1.5, Punkte (12), (2), (3). □

4.1.3. *Lineare Abbildungen und Unterräume.*

Bemerkung 4.1.7. Seien $(V,+,\cdot)$ und $(W,+,\cdot)$ Vektorräume über dem selben Körper $\mathbb{K}$. Sei $f \in \mathrm{Hom}(V,W)$. Dann gilt: $f(0) = 0$.

Beweis. Der Übersicht halber unterscheiden wir hier die Nullvektoren $0_V \in V$, $0_W \in W$, und den Skalar $0 \in \mathbb{K}$. Es ist

$$f(0_V) = f(0 \cdot 0_V) = 0 \cdot f(0_V) = 0_W.$$

□

Zur Erinnerung: Sind V und W Mengen, und ist $f : V \to W$ eine — nicht unbedingt bijektive! — Funktion, so hat das Symbol f^{-1} trotzdem eine Bedeutung, obwohl es keine Umkehrfunktion gibt. Ist nämlich $B \subset W$, so ist $f^{-1}(B) := \{x \in V \mid f(x) \in B\}$ das Urbild der Menge B unter f. Ist $A \subset V$, so ist $f(A) := \{f(x) \mid x \in A\}$ das Bild der Menge A unter f.

Satz 4.1.8. *Seien $(V,+,\cdot)$ und $(W,+,\cdot)$ Vektorräume über dem selben Körper $\mathbb{K}$. Sei $f \in \mathrm{Hom}(V,W)$. Dann gilt:*

1) *Ist A ein Unterraum von V, so ist $f(A)$ ein Unterraum von W.*
2) *Ist B ein Unterraum von W, so ist $f^{-1}(B)$ ein Unterraum von V.*

Beweis. Punkt (1): Seien $\tilde{x}, \tilde{y} \in f(A)$, $\alpha, \beta \in \mathbb{K}$. Wir zeigen: $\alpha\tilde{x} + \beta\tilde{y} \in f(A)$. Nach Definition von $f(A)$ gibt es $x, y \in A$ sodass $\tilde{x} = f(x)$, $\tilde{y} = f(y)$. Da A ein Unterraum von V ist, ist $\alpha x + \beta y \in A$. Wegen der Linearität von f ist

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = \alpha\tilde{x} + \beta\tilde{y}.$$

Daher ist $\alpha\tilde{x} + \beta\tilde{y} \in f(A)$.
Punkt (2): Seien $x, y \in f^{-1}(B)$, $\alpha, \beta \in \mathbb{K}$. Wir zeigen: $\alpha x + \beta y \in f^{-1}(B)$. Wir wissen nach Definition von $f^{-1}(B)$, dass $f(x), f(y) \in B$. Weil B ein Unterraum von W ist, ist auch $\alpha f(x) + \beta f(y) \in B$. Also ist wegen der Linearität von f

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) \in B,$$

das heißt, $\alpha x + \beta y \in f^{-1}(B)$. □

Definition 4.1.9. Seien $(V, +, \cdot)$ und $(W, +, \cdot)$ Vektorräume über dem selben Körper $\mathbb{K}$. Sei $f \in \mathrm{Hom}(V, W)$. Wir definieren die folgenden Unterräume:

1) Der Kern von f ist $\ker(f) := f^{-1}(\{0\})$.
2) Der Bildraum oder Range von f ist $\mathrm{rg}(f) := f(V)$.

Es ist leicht zu sehen, dass für eine Matrix $A \in \mathbb{K}^{m\times n}$ und die lineare Abbildung $f(x) = Ax$ die obigen Definitionen von Kern und Bildraum mit den Definitionen von Kern und Spaltenraum einer Matrix übereinstimmen. Daher führt die neuerliche Verwendung der Schreibweisen ker und rg zu keinen Missverständnissen!

Satz 4.1.10. *Seien $(V, +, \cdot)$ und $(W, +, \cdot)$ Vektorräume über dem selben Körper $\mathbb{K}$. Sei $f \in \mathrm{Hom}(V, W)$.*

1) *Kern und Bildraum von f sind Unterräume von V bzw. W.*
2) *f ist genau dann surjektiv, wenn $\mathrm{rg}(f) = W$.*
3) *f ist genau dann injektiv, wenn $\ker(f) = \{0\}$.*

Beweis. Punkt (1): Folgt aus Satz 4.1.8, weil V ein Unterraum von V und $\{0\}$ ein Unterraum von W ist.
Punkt (2): Das ist genau die Definition der Surjektivität: $f(V) = W$.
Punkt (3): Sei f injektiv. Jedenfalls ist (für jede lineare Abbildung) $f(0) = 0$. Ist dann $x \in V$ so, dass $f(x) = 0$, so ist wegen der Injektivität $x = 0$. Also ist $\ker(f) = \{0\}$.
Sei umgekehrt $\ker(f) = \{0\}$. Seien $x, y \in V$ so, dass $f(x) = f(y)$. Dann ist wegen der Linearität von f

$$f(x - y) = f(x) - f(y) = 0.$$

Also ist $x - y \in \ker(f)$, und folglich ist $x - y = 0$. Damit ist aber $x = y$. Also ist f injektiv. □

Satz 4.1.11. *Sei U ein endlichdimensionaler Vektorraum und V ein Vektorraum über demselben Körper $\mathbb{K}$. Sei $f : U \to V$ eine lineare Abbildung.*
Dann ist

$$\dim(U) = \dim(\mathrm{rg}(f)) + \dim(\ker(f)).$$

Beweis. Nach Satz 3.5.8 gibt es einen Unterraum W von U so, dass $U = \ker(f) + W$, wobei diese Summe direkt ist. Wir betrachten die Einschränkung von f auf W:

$$f\,|_W\colon \begin{cases} W & \to \mathrm{rg}(f), \\ x & \to f(x). \end{cases}$$

Als Einschränkung einer linearen Abbildung ist auch $f\mid_W$ linear. Für jedes $u \in U$ gibt es eine (eindeutige) Zerlegung $u = x + y$ mit $x \in W$ und $y \in \ker(f)$. Dann ist

$$f(u) = f(x) + f(y) = f\mid_W(x) + 0 = f\mid_W(x).$$

Also ist der Bildraum von f zugleich der Bildraum von $f\mid_W$. Insbesondere ist $f\mid_W\colon W \to \mathrm{rg}(f)$ surjektiv. Aber $f\mid_W$ ist auch injektiv: Es sei $x \in \ker(f\mid_W)$, also ist $x \in W$ und $f(x) = f\mid_W(x) = 0$. Damit ist $x \in W \cap \ker(f)$. Weil die Summe $W + \ker(f)$ aber direkt ist, ist der Durchschnitt $W \cap \ker(f)$ der Nullraum und $x = 0$. Der Kern $\ker(f\mid_W)$ ist also der Nullraum, und $f\mid_W$ ist injektiv. Damit haben wir einen Isomorphismus $f\mid_W\colon W \to \mathrm{rg}(f)$, insbesondere haben W und $\mathrm{rg}(f)$ dieselbe Dimension. Nun ist wegen der direkten Summe

$$\dim(U) = \dim(W) + \dim(\ker(f)) = \dim(\mathrm{rg}(f)) + \dim(\ker(f)).$$

□

Wir kennen ja schon den Satz für eine $m \times n$-Matrix A:

$$n = \rho(A) + \dim(\ker(A)).$$

Der obige Satz sagt dasselbe, auf einem abstrakteren Niveau ausgedrückt.

4.1.4. *Lineare Abbildungen und Basen.*

Satz 4.1.12. *Seien U, V zwei Vektorräume über demselben Körper $\mathbb{K}$, sei $(u_1, \cdots, u_n)$ eine Basis von U, und seien $v_1, \cdots, v_n \in V$.*

1. *Es gibt genau eine lineare Abbildung $f : U \to V$, sodass für alle $i = 1, \cdots, n$ gilt $f(u_i) = v_i$.*
2. *Diese Abbildung f ist genau dann injektiv, wenn $(v_1, \cdots, v_n)$ ein linear unabhängiges System ist.*
3. *Die Abbildung f ist genau dann surjektiv, wenn $\{v_1, \cdots, v_n\}$ ein Erzeugendensystem von V ist.*
4. *Die Abbildung f ist genau dann ein Isomorphismus, wenn $(v_1, \cdots, v_n)$ eine Basis von V ist.*

Beweis. Punkt (1)*:* Weil $(u_1, \cdots, u_n)$ eine Basis von U ist, besitzt jedes $x \in U$ eine eindeutige Zerlegung $x = \sum_{i=1}^n \lambda_i u_i$ mit $\lambda_i \in \mathbb{K}$. Wir definieren dann $f(x) := \sum_{i=1}^n \lambda_i v_i$. Offensichtlich ist $f(u_i) = f(1u_i) = 1v_i = v_i$. Um die Linearität von f zu zeigen, betrachten wir $x = \sum_{i=1}^n \lambda_i u_i$, $y = \sum_{i=1}^n \mu_i u_i$, und $\alpha, \beta \in \mathbb{K}$. Es ist dann offensichtlich

$$\alpha x + \beta y = \sum_{i=1}^n (\alpha\lambda_i + \beta\mu_i) u_i\,,$$

und daher ist

$$f(\alpha x + \beta y) = \sum_{i=1}^n (\alpha\lambda_i + \beta\mu_i) v_i = \alpha \sum_{i=1}^n \lambda_i v_i + \beta \sum_{i=1}^n \mu_i v_i = \alpha f(x) + \beta f(y)\,.$$

Sei $g : U \to V$ linear mit $g(u_i) = v_i$. Dann gilt für jedes $x = \sum_{i=1}^n \lambda_i u_i$

$$g(x) = \sum_{i=1}^n \lambda_i g(u_i) = \sum_{i=1}^n \lambda_i v_i = f(x)\,.$$

Daraus folgt die Eindeutigkeit.
Punkt (2)*:* Für $x = \sum_{i=1}^n \lambda_i u_i$ gilt

$$x \in \ker(f) \Leftrightarrow \sum_{i=1}^n \lambda_i v_i = 0\,.$$

Ist das System $(v_1, \cdots, v_n)$ linear abhängig, so gibt es $\lambda_1, \cdots, \lambda_n$, nicht alle gleich null, so dass $\sum_{i=1}^n \lambda_i v_i = 0$. Dann ist aber wegen der linearen Unabhängigkeit der Basis $(u_1, \cdots, u_n)$ der Vektor $x := \sum_{i=1}^n \lambda_i u_i$ nicht der Nullvektor. Andererseits ist $f(x) = 0$. Daher hat f einen nichttrivialen Kern, und ist nicht injektiv.
Sei andererseits $(v_1, \cdots, v_n)$ linear unabhängig und $x = \sum_{i=1}^n \lambda_i u_i \in \ker(f)$. Dann ist also $0 = f(x) = \sum_{i=1}^n \lambda_i v_i$, und wegen der linearen Unabhängigkeit des Systems $(v_1, \cdots, v_n)$ sind alle λ_i gleich Null. Daher ist $x = 0$ der einzige Vektor in $\ker(f)$, und f ist injektiv.
Punkt (3)*:* Sei $\{v_1, \cdots, v_n\}$ ein Erzeugendensystem von V. Dann lässt sich jedes $y \in V$ in der Form

$$y = \sum_{i=1}^n \lambda_i v_i = f(\sum_{i=1}^n \lambda_i u_i) \in f(U)$$

schreiben. Folglich ist f surjektiv.
Sei andererseits $\{v_1, \cdots, v_n\}$ kein Erzeugendensystem von V. Dann gibt es ein $y \in V$ das keine Linearkombination der v_i ist. Weil aber $f(U)$ per definitionem nur aus Linearkombinationen der v_i besteht, kann f nicht surjektiv sein.
Punkt (4)*:* folgt unmittelbar aus Punkt (2) und Punkt (3). □

4.1.5. *Der Vektorraum* $\mathrm{Hom}(V, W)$.

Definition 4.1.13. Seien $(V, +, \cdot)$ und $(W, +, \cdot)$ Vektorräume über dem selben Körper $\mathbb{K}$. Auf $\mathrm{Hom}(V, W)$ führen wir folgende Rechenoperationen ein: $+ : \mathrm{Hom}(V, W) \times \mathrm{Hom}(V, W) \to \mathrm{Hom}(V, W)$, $\cdot : \mathbb{K} \times \mathrm{Hom}(V, W) \to \mathrm{Hom}(V, W)$:
Für $f, g \in \mathrm{Hom}(V, W)$, $\alpha \in \mathbb{K}$ setzen wir (für alle $x \in V$)

$$\begin{aligned} [f + g](x) &:= f(x) + g(x), \\ [\alpha f](x) &:= \alpha f(x). \end{aligned}$$

Satz 4.1.14. *Seien* $(V, +, \cdot)$ *und* $(W, +, \cdot)$ *Vektorräume über dem selben Körper* $\mathbb{K}$*. Mit den Rechenoperationen aus Definiton 4.1.13 ist* $\mathrm{Hom}(V, W)$ *ein Vektorraum über* $\mathbb{K}$*.*

Beweis. Wir müssen zuerst zeigen, dass die Rechenoperationen nicht aus $\mathrm{Hom}(V, W)$ hinausführen. Seien also $f, g \in \mathrm{Hom}(V, W)$, $\alpha \in \mathbb{K}$. Wir zeigen: $f + g$ ist linear. Für alle $x, y \in V$, $\lambda, \mu \in \mathbb{K}$ gilt

$$\begin{aligned} [f + g](\lambda x + \mu y) &= f(\lambda x + \mu y) + g(\lambda x + \mu y) = (\lambda f(x) + \mu f(y)) + (\lambda g(x) + \mu g(y)) \\ &= \lambda(f(x) + g(x)) + \mu(f(y) + g(y)) = \lambda[f + g](x) + \mu[f + g](y). \end{aligned}$$

Also ist auch $f + g$ linear. Die Linearität von αf zeigt man ebenso durch direktes Nachrechnen.
Nun sind alle Rechengesetze eines Vektorraumes nachzuprüfen. Diese Rechnerei führt zu keinen Überraschungen. Der Nullvektor im Raum $\mathrm{Hom}(V, W)$ ist die konstante Abbildung $f(x) = 0$, und die Abbildung $-f$ ist $[-f](x) := -f(x)$. □

4.2. **Lineare Abbildungen in $\mathbb{K}^n$, Matrizen.**

4.2.1. *Lineare Abbildungen $\mathbb{K}^n \to \mathbb{K}^m$.*

Wiederholung: Produkt einer Matrix mit einem Vektor.

Definition 4.2.1. Sei $\mathbb{K}$ ein Körper. Die Einheitsvektoren in $\mathbb{K}^n$ sind die Vektoren

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix},\ \cdots,\ e_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Satz 4.2.2. *Sei $\mathbb{K}$ ein Körper, $m, n \in \mathbb{N}$.*

1) *Ist $A \in \mathbb{K}^{m\times n}$, so ist die Abbildung*

$$f : \begin{cases} \mathbb{K}^n & \to \mathbb{K}^m \\ x & \mapsto Ax \end{cases}$$

linear. Dabei sind die Spalten $s_1, \cdots, s_n$ von A die Bilder der Einheitsvektoren:

$$s_j = Ae_j.$$

2) *Ist $f : \mathbb{K}^n \to \mathbb{K}^m$ eine lineare Abbildung, so gibt es genau eine Matrix $A \in \mathbb{K}^{m\times n}$ so, dass*

$$(\forall x \in \mathbb{K}^n)\ f(x) = Ax.$$

Beweis. Punkt (1): Dass die Abbildung $x \mapsto Ax$ linear ist, wurde bereits in Satz 3.3.6 bewiesen. Sei nun $A = (\alpha_{i,j})_{i=1,\cdots,m\,,\,j=1,\cdots,n}$, sei e_j der j-te Einheitsvektor in $\mathbb{K}^n$. Dann ist

$$Ae_j = \begin{pmatrix} \alpha_{1,1}\cdot 0 + \cdots + \alpha_{1,j-1}\cdot 0 + \alpha_{1,j}\cdot 1 + \alpha_{1,j+1}\cdot 0 + \cdots + \alpha_{1,n}\cdot 0 \\ \vdots \\ \alpha_{m,1}\cdot 0 + \cdots + \alpha_{m,j-1}\cdot 0 + \alpha_{m,j}\cdot 1 + \alpha_{m,j+1}\cdot 0 + \cdots + \alpha_{m,n}\cdot 0 \end{pmatrix} = \begin{pmatrix} \alpha_{1,j} \\ \vdots \\ \alpha_{m,j} \end{pmatrix}$$

und das ist gerade die j-te Spalte von A.
Punkt (2): Wegen Punkt (1) ist die einzige Matrix, für die gelten kann dass $f(x) = Ax$, die Matrix, deren Spalten aus den Bildern der Einheitsvektoren gebildet werden. Sei also $s_j = f(e_j)$, und sei A die Matrix mit den Spalten $s_1, \cdots, s_n$. Wir müssen zeigen, dass für alle $x \in \mathbb{K}^n$ gilt: $f(x) = Ax$. Es sei also

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{K}^n.$$

Nach Satz 3.3.8 ist

$$Ax = \sum_{j=1}^{n} \xi_j s_j = \sum_{j=1}^{n} \xi_j f(e_j) = f(\sum_{j=1}^{n} \xi_j e_j) = f(x).$$

□

Die folgende Definition, mit der wir bereits schweigend manchmal gearbeitet haben, schreiben wir der Vollständigkeit halber endlich auch im Skriptum nieder.

Definition 4.2.3. Sei $\mathbb{K}$ ein Körper. Seien $A = (\alpha_{i,j})_{i=1,\cdots,m,\, j=1,\cdots,n}$ und $B = (\beta_{i,j})$ zwei Matrizen in $\mathbb{K}^{m\times n}$. Sei $\lambda \in \mathbb{K}$. Wir definieren

$$A + B := \begin{pmatrix} \alpha_{1,1}+\beta_{1,1} & \alpha_{1,2}+\beta_{1,2} & \cdots & \alpha_{1,n}+\beta_{1,n} \\ \vdots & & \vdots & \\ \alpha_{m,1}+\beta_{m,1} & \alpha_{m,2}+\beta_{m,2} & \cdots & \alpha_{m,n}+\beta_{m,n} \end{pmatrix},$$

$$\lambda A := \begin{pmatrix} \lambda\alpha_{1,1} & \lambda\alpha_{1,2} & \cdots & \lambda\alpha_{1,n} \\ \vdots & & \vdots & \\ \lambda\alpha_{m,1} & \lambda\alpha_{m,2} & \cdots & \lambda\alpha_{m,n} \end{pmatrix}.$$

Satz 4.2.4. *Sei $\mathbb{K}$ ein Körper, und $M : \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m) \to \mathbb{K}^{m\times n}$ die Abbildung, die jeder linearen Abbildung $f : \mathbb{K}^n \to \mathbb{K}^m$ ihre Matrix im Sinne von Satz 4.2.2 zuordnet.*
Dann ist M ein Vektorraumisomorphismus[16].

Beweis. Da zu jeder linearen Abbildung genau eine Matrix, und zu jeder Matrix genau eine lineare Abbildung gehört, ist M bijektiv. Seien nun $f, g \in \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ und A, B ihre Matrizen, sei $\lambda \in \mathbb{K}$. Nun ist nur mehr die Linearität nachzuweisen:

1) Die Matrix von $f + g$ ist $A + B$.
2) Die Matrix von λf ist λA.

Wir zeigen Punkt (1), der andere Punkt geht ebenso. Sei C die Matrix von $f + g$. Seien $a_1, \cdots, a_n$ die Spalten von A, $b_1, \cdots, b_n$ die Spalten von B, und $c_1, \cdots, c_n$ die Spalten von C. Wegen Satz 4.2.2 ist also $a_j = f(e_j)$, $b_j = g(e_j)$ und $c_j = [f+g](e_j)$. Dann ist also

$$c_j = [f + g](e_j) = f(e_j) + g(e_j) = a_j + b_j.$$

Daraus folgt aber $C = A + B$. □

Definition 4.2.5. Sei $\mathbb{K}$ ein Körper, $m, n \in \mathbb{N}$. Die $m \times n$-Nullmatrix ist die $m \times n$-Matrix

$$0 := \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Verwechslungen der Nullmatrix mit dem Nullvektor und dem Skalar 0 passieren nicht, wenn man auf den Zusammenhang achtet. Auch müsste man streng genommen am Symbol für die Nullmatrix kenntlich machen, um welche Nullmatrix es sich handelt, also die Anzahl der Zeilen und Spalten, z.B in der Art wie $0_{m\times n}$. Das ist aber absolut unüblich und auch nicht nötig, der Zusammenhang macht alles klar.

Bemerkung 4.2.6. Die $m\times n$-Nullmatrix ist der Nullvektor im Vektorraum $\mathbb{K}^{m\times n}$.

Beweis. Klar? □

[16]Sind wir nicht gebildet?! So gelehrt können wir ausdrücken, dass

1) zu jeder linearen Abbildung eine Matrix gehört und umgekehrt,
2) wir die Matrix von $f + g$ einfach erhalten, indem wir die Matrizen von f und von g zusammenzählen,
3) wir die Matrix von λf erhalten, indem wir die Matrix von f mit λ multiplizieren.

4.2.2. *Matrixmultiplikation.*

Definition 4.2.7. Sei $\mathbb{K}$ ein Körper, $k, m, n \in \mathbb{N}$. Sei $A = (\alpha_{i,j}) \in \mathbb{K}^{k\times m}$ und $B = (\beta_{i,j}) \in \mathbb{K}^{m\times n}$. Wir definieren die Produktmatrix $A \cdot B \in \mathbb{K}^{k\times n}$ durch

$$\begin{aligned} A \cdot B &:= (c_{i,j})_{i=1,\cdots,k\, j=1,\cdots,n} \text{ mit} \\ c_{i,j} &:= \sum_{s=1}^{m} \alpha_{i,s}\beta_{s,j}. \end{aligned}$$

Bemerkung 4.2.8.

1) Um $A \cdot B$ zu bilden, muss also die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B sein. Das Element $c_{i,j}$ der Matrix $A \cdot B$ wird dann aus der i-ten Zeile von A und der j-ten Spalte von B berechnet.
2) Ist $A \in \mathbb{K}^{m\times n}$ und $x \in \mathbb{K}^n$, so ist das Produkt Ax nichts anderes als das Matrizenprodukt, wenn man x als eine Matrix mit n Zeilen und einer Spalte auffasst.

Beispiel 4.2.9.

$$\begin{pmatrix} 2 & 1 & 3 \\ -1 & -2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 4 & 5 \\ 3 & 1 \\ 2 & 2 \end{pmatrix} = \begin{pmatrix} 17 & 17 \\ -10 & -7 \end{pmatrix},$$

$$\begin{pmatrix} 4 & 5 \\ 3 & 1 \\ 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & 3 \\ -1 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 3 & -6 & 12 \\ 5 & 1 & 9 \\ 2 & -2 & 6 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 1 & 3 \\ -1 & -2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 4 & 3 & 2 \\ 5 & 1 & 2 \end{pmatrix} \text{ gibt es nicht!}$$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Definition 4.2.10. Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Die $n \times n$-Einheitsmatrix E_n ist die Matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Schreibweise 4.2.11. Den Index n lässt man oft weg und schreibt einfach E für die Einheitsmatrix, wenn ohnehin klar ist, wieviele Zeilen und Spalten sie hat. Andere gebräuchliche Symbole für die Einheitsmatrix sind I oder 1. Ist $\lambda \in \mathbb{K}$ so schreibt man oft einfach λ anstatt λE — aus dem Zusammenhang muss dann geschlossen werden, ob man den Skalar meint oder eine Matrix, die auf der Diagonalen überall λ hat, und sonst Nullen.

Definition 4.2.12. Das Kronecker-Delta $\delta : \mathbb{Z} \times \mathbb{Z} \to \{0, 1\}$ wird definiert durch

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

Bemerkung 4.2.13. Das Kronecker-Delta ist eine sehr praktische Schreibweise! Zum Beispiel kann man damit bequem die Einheitsmatrix definieren

$$E_n = (\delta_{i,j})_{i=1,\cdots,n,\ j=1,\cdots,n}$$

Auch die Einheitsvektoren in $\mathbb{K}^n$ kann man schön hinschreiben:

$$e_k = (\delta_{i,k})_{i=1,\cdots,n}.$$

Satz 4.2.14. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{m\times n}$, E_m und E_n die $m \times m$ bzw. $n \times n$ Einheitsmatrizen über $\mathbb{K}$. Dann ist*

$$E_m A = A E_n = A.$$

Insbesondere gilt für alle $x \in \mathbb{K}^n$: $E_n x = x$. Das heißt, die $n \times n$-Einheitsmatrix ist die Matrix der Identitätsabbildung auf $\mathbb{K}^n$.

Beweis. Wir zeigen $E_m A = A$. Sei also $A = (\alpha_{i,j})_{i=1,\cdots,m,\ j=1,\cdots,n}$. Dann ist der Koeffizient an der Stelle i, j der Matrix $E_m A$ nach Definition des Matrixproduktes

$$\sum_{k=1}^{m} \delta_{i,k}\alpha_{k,j}.$$

Nun ist $\delta_{i,k}$ immer Null, ausser wenn $k = i$ ist, und daher bleibt von der ganzen Summe nur übrig

$$\sum_{k=1}^{m} \delta_{i,k}\alpha_{k,j} = \delta_{i,i}\alpha_{i,j} = \alpha_{i,j}.$$

Die Gleichung $AE_n = A$ geht ebenso. □

Manchmal ist der Beweis für einen sehr einfachen Sachverhalt schwierig formal aufzuschreiben. Wenn Ihnen nicht einleuchtet, was da oben steht, rechnen Sie einmal

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

Es ist grundsätzlich oft hilfreich für das Verständnis, wenn man abstrakte Beweise an Hand einfacher Beispiele nachvollzieht.

Definition 4.2.15. Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix (d.h., eine Matrix, die soviele Zeilen wie Spalten hat). Wir definieren rekursiv die Potenzen von A:

$$\begin{aligned} A^0 &:= E_n \quad (\text{die } n \times n\text{-Einheitsmatrix}), \\ A^{k+1} &:= A \cdot A^k \quad \text{für } k = 0, 1, 2, \cdots . \end{aligned}$$

Wir haben formal einfach das definiert, was wir uns unter einer Potenz vorstellen:

$$A^k := A \cdot A \cdots A \quad (\text{das Produkt von } k \text{ gleichen Faktoren}).$$

Ist die rekursive Definition keine Zirkeldefinition? Nein, denn A^0 ist ja definiert, und sobald A^0 bekannt ist, können wir sagen, dass $A^1 = A \cdot A^0 = A$. Nun kennen wir A^1 und können $A^2 = AA^1 = AA$ definieren, und auf diese Weise erreichen wir jede Potenz. Im Prinzip funktioniert eine rekursive Definition wie ein Induktionsbeweis, nur dass definiert statt bewiesen wird.

Bemerkung 4.2.16. An Beispiel 4.2.9 kann man sehen:

1) Das Matrixprodukt existiert nur, wenn die erste Matrix soviele Spalten hat wie die zweite Matrix Zeilen.

2) Das Matrixprodukt ist nicht kommutativ: Es gibt Matrizen A, B sodass $AB \neq BA$.
3) Das Matrixprodukt hat Nullteiler, das heißt, es gibt Matrizen A, B, beide nicht Null, sodass das Produkt AB trotzdem die Nullmatrix ergibt.
4) Idempotente Matrizen sind Matrizen $A \in \mathbb{K}^{n\times n}$ für die $AA = A$ ist. Das ist klar für die $n \times n$-Nullmatrix und die Einheitsmatrix. Aber darüber hinaus gibt es auch andere idempotente Matrizen.
5) Nilpotente Matrizen sind $n \times n$-Matrizen, für die es eine Potenz A^k gibt, welche die Nullmatrix ergibt. Das ist klar für die Nullmatrix, aber darüber hinaus gibt es auch andere nilpotente Matrizen.

Es folgt eine Litanei von Rechenregeln für das Matrixprodukt. Alle diese Regeln kann man durch direktes Nachrechnen beweisen: Beweis durch Förderung der Papierindustrie. Versuchen Sie selbst, die eine oder andere Regel zu beweisen.

Satz 4.2.17. *Sei $\mathbb{K}$ ein Körper, $k, m, n, p \in \mathbb{N}$, sei*[17] *$A \in \mathbb{K}^{k\times m}$, $B, C \in \mathbb{K}^{m\times n}$, $D \in \mathbb{K}^{n\times p}$, sei 0 die Nullmatrix, jeweils in geeigneten Dimensionen. Sei $\lambda \in \mathbb{K}$. Es gilt:*

$$
\begin{aligned}
&(A \cdot B) \cdot D = A \cdot (B \cdot D) \quad \textit{Assoziativität},\\
&A \cdot (B + C) = A \cdot B + A \cdot C \quad \textit{Distributivität},\\
&(B + C) \cdot D = B \cdot D + C \cdot D \quad \textit{Distributivität},\\
&(\lambda A) \cdot B = \lambda(A \cdot B) = A \cdot (\lambda B),\\
&A \cdot 0 = 0, \quad 0 \cdot A = 0,\\
&A \cdot E_m = A, \quad E_k \cdot A = A.
\end{aligned}
$$

Bemerkung 4.2.18. Ist $\mathbb{K}$ ein Körper, und $n \in \mathbb{N}$, so ist $\mathbb{K}^{n\times n}$ mit der Matrizenaddition und der Matrixmultiplikation ein Ring mit Einselement E_n. Eine Struktur, die zugleich Ring und Vektorraum ist (wobei die Vektorraumaddition zugleich die Ringaddition ist, und sich die Multiplikation mit Skalaren mit der Ringmultiplikation verträgt, siehe vierte Gleichung von Satz 4.2.17) heißt eine Algebra[18].

In einem Mittelschullehrbuch fand ich einmal die folgende Passage: "... Richtig geraten! Die Menge der $n \times n$-Matrizen bildet einen Körper!" Was sagen Sie? Immer ein Körper? Nie ein Körper, oder vielleicht doch in bestimmten Fällen? Was alles könnte gegen einen Körper sprechen? — Lesen Sie alle Texte kritisch, auch dieses Skriptum! Wir alle, auch sehr gute AutorInnen, sind nicht gefeit gegen Fehler. Aber Sie als MathematikerInnen sind im Stande, Fehler zu erkennen und auszubessern!

4.2.3. *Transponierte und adjungierte Matrix.*

Definition 4.2.19. Sei $\mathbb{K}$ ein Körper, $A = (\alpha_{i,j})_{i=1,\cdots,m,\ j=1,\cdots,n} \in \mathbb{K}^{m\times n}$.

1) Die transponierte Matrix $A^T \in \mathbb{K}^{n\times m}$ entsteht, indem man die Zeilen von A als Spalten,und die Spalten von A demnach als Zeilen schreibt:

$$
\begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \cdots & \alpha_{m,n} \end{pmatrix}^T := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{m,1} \\ \alpha_{1,2} & \cdots & \alpha_{m,2} \\ \vdots & & \vdots \\ \alpha_{1,n} & \cdots & \alpha_{m,n} \end{pmatrix}.
$$

[17] Die Dimensionen der Matrizen sind einfach so festgelegt, dass es die Produkte, die unten gebildet werden, wirklich gibt.

[18] Das Wort Algebra kann also zwei ganz verschiedene Dinge bezeichnen: Einen Zweig der Mathematik, oder eine bestimmte Struktur.

2) Ist $\mathbb{K} = \mathbb{R}$, so ist die adjungierte Matrix $A^* := A^T$. Ist $\mathbb{K} = \mathbb{C}$, so entsteht die adjungierte Matrix A^* aus der transponierten Matrix A^T, indem man für jeden Koeffizienten die konjugiert komplexe Zahl nimmt:

$$\begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \cdots & \alpha_{m,n} \end{pmatrix}^* := \begin{pmatrix} \overline{\alpha_{1,1}} & \cdots & \overline{\alpha_{m,1}} \\ \overline{\alpha_{1,2}} & \cdots & \overline{\alpha_{m,2}} \\ \vdots & & \vdots \\ \overline{\alpha_{1,n}} & \cdots & \overline{\alpha_{m,n}} \end{pmatrix}.$$

Schreibweise 4.2.20. Statt A^T findet man auch A^t oder A'.

Beispiel 4.2.21.

$$\begin{aligned} \begin{pmatrix} 1+2i & 1-2i \\ 0 & 3i \\ 4-i & 5 \end{pmatrix}^T &= \begin{pmatrix} 1+2i & 0 & 4-i \\ 1-2i & 3i & 5 \end{pmatrix}, \\ \begin{pmatrix} 1+2i & 1-2i \\ 0 & 3i \\ 4-i & 5 \end{pmatrix}^* &= \begin{pmatrix} 1-2i & 0 & 4+i \\ 1+2i & -3i & 5 \end{pmatrix}. \end{aligned}$$

Die folgenden Rechenregeln sind keine Überraschungen. Achten Sie aber auf konjugiert komplexe Zahlen bei der adjungierten Matrix, und auf die Reihenfolgen beim Matrixprodukt!

Satz 4.2.22. *Sei $\mathbb{K}$ ein Körper, seien $A, B \in \mathbb{K}^{m\times n}$, $C \in \mathbb{K}^{k\times m}$ und $\lambda \in \mathbb{K}$. In jeder Regel für adjungierte Matrizen nehmen wir zusätzlich an $\mathbb{K} = \mathbb{C}$.*

$$\begin{aligned} &(A+B)^T = A^T + B^T, \quad (A+B)^* = A^* + B^*, \\ &(\lambda A)^T = \lambda(A^T), \quad (\lambda A)^* = \overline{\lambda}(A^*), \\ &(CA)^T = A^T C^T, \quad (CA)^* = A^* C^*. \end{aligned}$$

Beweis. Von diesen Regeln zeigen wir nur $(CA)^T = A^T C^T$.
Sei $A = (\alpha_{i,j})_{i=1,\cdots,m,\ j=1,\cdots,n}$, $C = (\gamma_{p,i})_{p=1,\cdots,k,\ i=1,\cdots,m}$. Sei $\epsilon_{j,p}$ der Koeffizient von $(CA)^T$ in Zeile j und Spalte p. Wir müssen zeigen, dass $\epsilon_{j,p}$ auch der entsprechende Koeffizient von $A^T C^T$ ist.
Es ist $\epsilon_{j,p}$ der Koeffizient von CA in Zeile p und Spalte j, das ist

$$\epsilon_{j,p} = \sum_{s=1}^{m} \gamma_{p,s}\alpha_{s,j}.$$

Andererseits ist die j-te Zeile von A^T die Zeile

$$\begin{pmatrix} \alpha_{1,j}, \cdots, \alpha_{m,j} \end{pmatrix}$$

und die p-te Spalte von C^T die Spalte

$$\begin{pmatrix} \gamma_{p,1} \\ \vdots \\ \gamma_{p,m} \end{pmatrix}.$$

Daher ist der Koeffizient an der Stelle (j,p) von $A^T C^T$

$$\sum_{s=1}^{m} \alpha_{s,j}\gamma_{p,s} = \epsilon_{j,p}.$$

□

4.2.4. *Hintereinanderausführung von linearen Abbildungen und Matrixprodukt.*
Zunächst mutet die Definition des Matrizenproduktes willkürlich und kompliziert an. Der folgende Satz erklärt, was das Matrizenprodukt eigentlich bedeutet, und warum es so wichtig ist.

Satz 4.2.23. *Sei $\mathbb{K}$ ein Körper, $k, m, n \in \mathbb{N}$, $A \in \mathbb{K}^{k\times m}$, $B \in \mathbb{K}^{m\times n}$. Seien f, g die Homomorphismen*

$$f : \begin{cases} \mathbb{K}^m & \to \mathbb{K}^k, \\ y & \mapsto Ay, \end{cases} \qquad g : \begin{cases} \mathbb{K}^n & \to \mathbb{K}^m, \\ x & \mapsto Bx. \end{cases}$$

Dann ist AB die Matrix der Hintereinanderausführung $f \circ g$, d.h. für alle $x \in \mathbb{K}^n$ ist

$$[f \circ g](x) = (AB)x.$$

Beweis. Sei C die Matrix von $f \circ g$. Da $f \circ g \in \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^k)$, ist $C \in \mathbb{K}^{k\times n}$. Wir schreiben

$$A = (\alpha_{p,i})_{p=1,\cdots,k,\ i=1,\cdots,m},\ B = (\beta_{i,j})_{i=1,\cdots,m,\ j=1,\cdots,n},\ C = (\gamma_{p,j})_{p=1,\cdots,k,\ j=1,\cdots,n}.$$

Nach Satz 4.2.2 sind die Spalten von C die Bilder der Einheitsvektoren in $\mathbb{K}^n$ unter $f \circ g$, und die Spalten von B sind die Bilder der Einheitsvektoren des $\mathbb{K}^n$ unter g. Wir bestimmen die j-te Spalte von C, dies ist also $f(g(e_j))$:

$$\begin{pmatrix} \gamma_{1,j} \\ \vdots \\ \gamma_{k,j} \end{pmatrix} = f(g(e_j)) = f(\begin{pmatrix} \beta_{1,j} \\ \vdots \\ \beta_{m,j} \end{pmatrix}) = A \begin{pmatrix} \beta_{1,j} \\ \vdots \\ \beta_{m,j} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m \alpha_{1,i}\beta_{i,j} \\ \vdots \\ \sum_{i=1}^m \alpha_{k,i}\beta_{i,j} \end{pmatrix}.$$

Dies ist aber zugleich die j-te Spalte von AB. □

Man kann aus Satz 4.2.23 auch einen einfachen Beweis für die Assoziativität des Matrixproduktes ableiten, denn wir wissen, dass die Hintereinanderausführung von Abbildungen assoziativ ist: $f \circ (g \circ h) = (f \circ g) \circ h$.

4.2.5. *Zeilentransformationen, Spaltentransformationen und Matrizen.*

Satz 4.2.24. *Sei $\mathbb{K}$ ein Körper, sei $A = (\alpha_{p,i})_{p=1,\cdots,k,\ i=1,\cdots,m} \in \mathbb{K}^{k\times m}$ und $B = (\beta_{i,j})_{i=1,\cdots,m,\ j=1,\cdots,n} \in \mathbb{K}^{m\times n}$. Mit $z_{B,i}$, $z_{AB,i}$ bezeichnen wir die Zeilen von B bzw. AB, mit $s_{A,j}$, $s_{AB,j}$ bezeichnen wir die Spalten von A bzw. AB. Dann gilt:*

1) *Die Zeilen von AB sind Linearkombinationen der Zeilen von B. Es gilt*
$$z_{AB,p} = \sum_{i=1}^m \alpha_{p,i}\, z_{B,i}.$$
2) *Die Spalten von AB sind Linearkombinationen der Spalten von A. Es gilt*
$$s_{AB,j} = \sum_{i=1}^m \beta_{i,j}\, s_{A,i}.$$
3) *Der Zeilenraum von AB ist ein Unterraum des Zeilenraums von B, und der Spaltenraum von AB ist ein Unterraum des Spaltenraumes von A.*

Beweis. Punkt (1): Wir gehen von der Definition des Matrixproduktes aus und berechnen die p-te Zeile von AB:

$$\begin{aligned} z_{AB,p} &= \begin{pmatrix} \sum_{i=1}^m \alpha_{p,i}\beta_{i,1} & \cdots & \sum_{i=1}^m \alpha_{p,i}\beta_{i,n} \end{pmatrix} \\ &= \sum_{i=1}^m \alpha_{p,i} \begin{pmatrix} \beta_{i,1} & \cdots & \beta_{i,n} \end{pmatrix} = \sum_{i=1}^m \alpha_{p,i}\, z_{B,i}. \end{aligned}$$

Punkt (2) geht ebenso, und Punkt (3) ist eine direkte Folge der Punkte (1) und (2). □

Definition 4.2.25. Sei $\mathbb{K}$ ein Körper, $m \in \mathbb{N}$, $\lambda \in \mathbb{K} \setminus \{0\}$, $i \neq j \in \{1, \cdots, m\}$. Wir definieren folgende Matrizen in $\mathbb{K}^{m \times m}$, die sogenannten Elementarmatrizen:
Wir haben zur besseren Orientierung die Zeilen- und Spaltennummern an den linken und oberen Rand der Matrizen geschrieben, und alle nicht eingefüllten Koeffizienten sind wie in der Einheitsmatrix. Unterstrichen sind die Koeffizienten, die sich von der Einheitsmatrix unterscheiden. Die Reihenfolge der Zeilen und Spalten von i und j kann auch umgekehrt sein, je nachdem ob $i < j$ oder $j < i$.

$$T_{\text{tausch}}(i,j) := \begin{pmatrix} & (1) & & (i) & & (j) & & (m) \\ (1) & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ (i) & 0 & \cdots & \underline{0} & \cdots & \underline{1} & \cdots & 0 \\ & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ (j) & 0 & \cdots & \underline{1} & \cdots & \underline{0} & \cdots & 0 \\ & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ (m) & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$$T_{\text{add}}(i,j,\lambda) := \begin{pmatrix} & (1) & & (i) & & (j) & & (m) \\ (1) & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ (i) & 0 & \cdots & 1 & \cdots & \underline{\lambda} & \cdots & 0 \\ & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ (j) & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ (m) & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$$T_{\text{mult}}(i,\lambda) := \begin{pmatrix} & (1) & & (i) & & (j) & & (m) \\ (1) & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ (i) & 0 & \cdots & \underline{\lambda} & \cdots & 0 & \cdots & 0 \\ & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ (j) & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ (m) & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

Satz 4.2.26. *Die elementaren Zeilenumformungen (siehe Definition 3.3.12) können mit Hilfe der Matrixmultiplikation mit Elementarmatrizen geschrieben werden: Ist $\mathbb{K}$ ein Körper, $m, n \in \mathbb{N}$, $A \in \mathbb{K}^{m \times n}$ und die $m \times m$ Elementarmatrizen wie in Definition 4.2.25 gegeben. Sei $\lambda \in \mathbb{K} \setminus \{0\}$, $i \neq j \in \mathbb{K}$. Dann gilt:*

1) *$T_{\text{tausch}}(i,j) \cdot A$ entsteht aus A indem man die i-te und die j-te Zeile von A vertauscht.*
2) *$T_{\text{add}}(i,j,\lambda) \cdot A$ entsteht aus A indem man das λ-fache der j-ten Zeile zur i-ten Zeile von A addiert.*
3) *$T_{\text{mult}}(i,\lambda) \cdot A$ entsteht aus A indem man die i-te Zeile von A mit λ multipliziert.*

Beweis. Nachrechnen mit Hilfe von Satz 4.2.24, Punkt (1). □

Multipliziert man zu $A \in \mathbb{K}^{m \times n}$ von rechts die $n \times n$-Elementarmatrizen: $A \cdot T_{\text{tausch}}$, $A \cdot T_{\text{add}}$, $A \cdot T_{\text{mult}}$, dann bekommt man analoge Transformationen für die Spalten von A.

4.3. **Inverse Matrix.**

4.3.1. *Reguläre Matrizen.*

Definition 4.3.1. Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. Die Matrix $B \in \mathbb{K}^{n\times n}$ heißt inverse Matrix zu A wenn gilt:

$$AB = BA = E_n.$$

Wir schreiben in diesem Fall: $A^{-1} := B$.

Der folgende einfache Satz rechtfertigt die Schreibweise A^{-1} für *die* inverse Matrix:

Satz 4.3.2.

1) *Zu jeder Matrix kann es höchstens eine Inverse geben.*
2) *Besitzt eine quadratische Matrix A eine Inverse A^{-1}, so ist die Inverse zu A^{-1} wiederum A, also $(A^{-1})^{-1} = A$.*

Beweis. Punkt (1): Seien B und C inverse Matrizen zu A. Dann ist

$$B = E_nB = (CA)B = C(AB) = CE_n = C.$$

Punkt (2): Trivial auf Grund der Definition der Inversen. □

Beispiel 4.3.3. Das folgende Beispiel zeigt, dass nicht jede quadratische Matrix eine Inverse besitzt: Sei $A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$. Dann besitzt A keine inverse Matrix.

Beweis. Angenommen, A besitzt eine inverse Matrix A^{-1}. Man kann leicht nachrechnen: $A^2 = 0$. Dann wäre aber

$$A = E_nA = (A^{-1}A)A = A^{-1}A^2 = A^{-1}0 = 0,$$

aber $A \neq 0$, und wir haben einen Widerspruch. □

Anders als in einem Körper ist also im Ring der $n \times n$-Matrizen das Nullelement nicht das einzige Element, das keine Inverse für die Multiplikation besitzt.

Definition 4.3.4. Eine quadratische Matrix, heißt regulär, wenn sie eine Inverse besitzt. Andernfalls heißt die Matrix singulär.

Satz 4.3.5. *Sei $\mathbb{K}$ ein Körper und seien $A, B \in \mathbb{K}^{n\times n}$ zwei reguläre Matrizen. Dann ist auch AB regulär, und $(AB)^{-1} = B^{-1}A^{-1}$. (Achtung, die Reihenfolge dreht sich um!)*

Beweis.

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AE_nA^{-1} = AA^{-1} = E_n, \\ (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B = B^{-1}E_nB = B^{-1}B = E_n. \end{aligned}$$

□

Bemerkung 4.3.6. Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$ eine reguläre Matrix. Dann ist auch die transponierte Matrix A^T regulär, und es gilt: $(A^T)^{-1} = (A^{-1})^T$.

Beweis. Übung. □

Satz 4.3.7. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix, und*

$$f : \begin{cases} \mathbb{K}^n & \to \mathbb{K}^n \\ x & \mapsto Ax \end{cases}$$

der dazugehörige Homomorphismus.
Dann ist A genau dann regulär, wenn f ein Isomorphismus ist. In diesem Fall gehört die Matrix A^{-1} zur Umkehrabbildung f^{-1} von f.

Beweis. Sei zunächst A regulär. Wir definieren den Homomorphismus

$$g : \begin{cases} \mathbb{K}^n & \to \mathbb{K}^n, \\ x & \mapsto A^{-1}x. \end{cases}$$

Dann ist für $x \in \mathbb{K}^n$

$$\begin{aligned} [f \circ g](x) &= A(A^{-1}x) = (AA^{-1})x = E_n x = x, \\ [g \circ f](x) &= A^{-1}(Ax) = (A^{-1}A)x = E_n x = x. \end{aligned}$$

Also ist

$$f \circ g = g \circ f = \mathrm{id}_{\mathbb{K}^n}$$

die Identität auf $\mathbb{K}^n$. Daher ist g die Umkehrabbildung zu f.
Sei nun f bijektiv und f^{-1} die Umkehrabbildung zu f. Wir wissen, dass zu f^{-1} eine Matrix B gehört, sodass für alle $x \in \mathbb{K}^n$ gilt $f^{-1}(x) = Bx$. Dann ist BA die Matrix zur Hintereinanderausführung $f^{-1} \circ f$, also zur Identität, also ist $BA = E_n$. Ebenso ist AB die Matrix zu $f \circ f^{-1} = \mathrm{id}_{\mathbb{K}^n}$, also $AB = E_n$. Damit ist B die inverse Matrix zu A. □

Satz 4.3.8. *Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$ eine $n\times n$-Matrix. Dann sind folgende Aussagen äquivalent:*

1) *Der Rang von A ist n.*
2) *Der Spaltenraum von A ist der ganze Raum $\mathbb{K}^n$.*
3) *Die Spalten von A sind eine Basis des $\mathbb{K}^n$.*
4) *Die Spalten von A sind linear unabhängig.*
5) *Der Zeilenraum von A ist der ganze Raum $\mathbb{K}^{1\times n}$.*
6) *Die Zeilen von A sind eine Basis des $\mathbb{K}^{1\times n}$.*
7) *Die Zeilen von A sind linear unabhängig.*
8) *Der Kern von A ist der Nullraum.*
9) *A ist regulär.*

Beweis. (1) ⇒ (2): Sei $\rho(A) = n$. Nach Definition des Ranges hat also der Spaltenraum von A die Dimension n. Weil aber $\mathrm{rg}(A)$ also ein n-dimensionaler Unterraum des n-dimensionalen Raumes $\mathbb{K}^n$ ist, gilt wegen Satz 3.4.7 Gleichheit: $\mathrm{rg}(A) = \mathbb{K}^n$.
(2) ⇒ (3): Ist $\mathrm{rg}(A) = \mathbb{K}^n$, so bilden die Spalten von A ein Erzeugendensystem des $\mathbb{K}^n$. Aus diesem läßt sich jedenfalls eine Basis auswählen (Satz 3.2.13). Da aber jede Basis des $\mathbb{K}^n$ aus genau n Vektoren besteht, müssen alle Spalten bleiben. Sie bilden also ein minimales Erzeugendensystem, eine Basis, des $\mathbb{K}^n$.
(3) ⇒ (4): Jede Basis ist linear unabhängig.
(4) ⇒ (1): Seien die Spalten von A linear unabhängig. Wegen Satz 3.4.2 ist die Dimension des Spaltenraumes von A also mindestens n. Da der Spaltenraum von A aber ein Teilraum von $\mathbb{K}^n$ ist, kann seine Dimension auch nicht größer als n sein. Wir haben also $\rho(A) = \dim(\mathrm{rg}(A)) = n$.
(1) ⇒ (5) ⇒ (6) ⇒ (7) ⇒ (1): Genau wie die eben bewiesenen Äquivalenzen, unter Berücksichtigung, dass der Rang von A zugleich die Dimension des Zeilenraumes von A ist (Satz 3.4.10).

(1) ⇔ (8): Nach Satz 3.4.10 hat der Kern von A die Dimension $\dim(\ker(A)) = n - \rho(A)$, und diese ist genau dann 0, wenn $\rho(A) = n$ ist.

(1) ⇔ (9): Sei f der zu A gehörige Homomorphismus

$$f : \begin{cases} \mathbb{K}^n & \to \mathbb{K}^n \\ x & \mapsto Ax. \end{cases}$$

Sei zunächst $\rho(A) = n$. Dann ist (wegen Punkt(2)) der Spaltenraum von A der ganze $\mathbb{K}^n$, also ist f surjektiv. Wegen Punkt (8) ist der Kern von f der Nullraum, also ist f injektiv (Satz 4.1.10). Also ist f bijektiv und nach Satz 4.3.7 ist A regulär. Ist umgekehrt A regulär, so ist f bijektiv, insbesondere surjektiv. Damit ist der Spaltenraum von A der ganze Raum $\mathbb{K}^n$. Das ist Punkt (2), und aus diesem folgt Punkt (1). □

Satz 4.3.9. *Sei $\mathbb{K}$ ein Körper, $f : \mathbb{K}^n \to \mathbb{K}^n$ ein Endomorphismus[19] des $\mathbb{K}^n$, und A die dazu gehörige Matrix. Dann sind äquivalent:*

1) *A ist regulär.*
2) *f ist ein Isomorphismus.*
3) *Für jedes $b \in \mathbb{K}^n$ ist das Gleichungssystem $Ax = b$ eindeutig lösbar.*
4) *f ist injektiv.*
5) *Für jedes $b \in \mathbb{K}^n$ hat das Gleichungssystem $Ax = b$ höchstens eine Lösung.*
6) *f ist surjektiv.*
7) *Für jedes $b \in \mathbb{K}^n$ hat das Gleichungssystem $Ax = b$ mindestens eine Lösung.*

In diesem Fall ist die Lösung des Gleichungssytems $Ax = b$ gegeben durch

$$x = A^{-1}b.$$

Beweis. (1) ⇔ (2): Satz 4.3.7.
(2) ⇔ (3): Ist einfach eine äquivalente Formulierung für Bijektivität: f ist bijektiv wenn für alle $b \in \mathbb{K}^n$ die Gleichung $f(x) = b$ eindeutig lösbar ist.
(2) ⇒ (4): Jede bijektive Abbildung ist injektiv.
(4) ⇔ (5): Äquivalente Formulierung der Injektivität.
(4) ⇒ (1): Ist f injektiv, so ist nach Satz 4.1.10 $\ker(f) = \ker(A) = \{0\}$. Nach Satz 4.3.8 folgt dass A regulär ist.
(2) ⇒ (6): Jede bijektive Abbildung ist surjektiv.
(6) ⇔ (7): Äquivalente Formulierung von Surjektivität.
(6) ⇒ (1): Ist f surjektiv, so ist der Spaltenraum von A der ganze Raum $\mathbb{K}^n$, und nach Satz 4.3.8 ist dann A regulär.
Sei nun A regulär, dann besitzt also $Ax = b$ eine eindeutige Lösung. Wir zeigen, dass $x = A^{-1}b$ das Gleichungssystem löst:

$$Ax = A(A^{-1}b) = (AA^{-1})b = E_n b = b.$$

□

Um ein einzelnes Gleichungssystem $Ax = b$ zu lösen, zahlt sich aber die Berechnung von A^{-1} nicht aus, sie macht mehr Arbeit als die Lösung des Gleichungssystems. Will man dagegen das System für viele verschiedene rechte Seiten lösen, ist es schon sinnvoll, die Inverse zu berechnen. Der Hauptnutzen des Satzes besteht aber in seiner Anwendung als Hilfsmittel für Beweise und zur Konstruktion weiterer Formeln und Rechenregeln.

4.3.2. *Berechnung der inversen Matrix.*

Algorithmus 4.3.10 (Berechnung der inversen Matrix)**.** Sei $\mathbb{K}$ ein Körper, A eine $n \times n$-Matrix über $\mathbb{K}$. Der folgende Algorithmus entscheidet, ob A regulär ist, und berechnet gegebenfalls die inverse Matrix A^{-1}:

1) Schreibe nebeneinander die Matrix A und die Einheitsmatrix E_n.
2) Forme die linke Seite A durch Pivotschritte in verallgemeinerte Zeilenstufenform um. Führe gleichzeitig alle Zeilenumformungen auch auf der rechten Seite durch.

[19] also eine lineare Abbildung von $\mathbb{K}^n$ in denselben Raum $\mathbb{K}^n$

3) Wenn auf der linken Seite eine Nullzeile entsteht, ist A singulär, und es gibt keine Inverse. Stop.
4) Wenn auf der linken Seite keine Nullzeile entsteht, waren alle Zeilen Pivotzeilen. Verwende nun Pivotschritte wie in der Rücksubstitution, um überall ausser auf den Pivotelementen Nullen zu erzeugen. (Wieder werden alle Zeilentransformationen zugleich auf die rechte Seite ausgeführt.)
5) Dividiere jede Zeile durch ihr Pivotelement. Auf der linken Seite stehen nun nur mehr Einsen und Nullen. (Wieder werden die Divisionen auch auf der rechten Seite ausgeführt.)
6) Ordne die Zeilen so um, dass auf der linken Seite die Einheitsmatrix steht. (Wieder werden die Zeilen rechts ebenso umgeordnet.) Rechts steht dann die Inverse zu A.

Beweis. Die Matrix A lässt sich jedenfalls auf verallgemeinerte Zeilenstufenform bringen. Entsteht dabei eine Nullzeile, so ist der Rang von A kleiner als n, folglich ist nach Satz 4.3.8 A singulär. Werden alle Zeilen als Pivotzeilen verwendet, kann man die Einheitsmatrix auf der linken Seite erzeugen. Wir müssen noch zeigen, dass auf der rechten Seite die Inverse von A entsteht, wenn A regulär ist.
Sei also A regulär. Wir beginnen mit $X_0 := A$, $Y_0 := E_n$. Nach der k-ten elementaren Zeilentransformation sei X_k die linke Seite, und Y_k die rechte Seite. Wir zeigen durch Induktion: $Y_k = X_k A^{-1}$.
Induktionsverankerung:

$$Y_0 = E_n = AA^{-1} = X_0 A^{-1}.$$

Induktionsschritt: Sei nun $Y_k = X_k A^{-1}$. Jede elementare Zeilenumformung lässt sich durch Multiplikation mit einer Elementarmatrix T beschreiben: $Y_{k+1} = TY_k$, $X_{k+1} = TX_k$. Dann ist also

$$Y_{k+1} = TY_k = T(X_k A^{-1}) = (TX_k)A^{-1} = X_{k+1}A^{-1}.$$

Damit ist die Induktion beendet.
Am Ende aller Zeilenumformungen, sagen wir, nach dem m-ten Schritt, steht links $X_m = E_n$. Damit steht rechts

$$Y_m = X_m A^{-1} = E_n A^{-1} = A^{-1}.$$

□

Beispiel 4.3.11. Ist die folgende Matrix A regulär? Wenn ja, berechne die Inverse:

$$A = \begin{pmatrix} 2 & 3 & 12 & 6 \\ 4 & 3 & 10 & 5 \\ 2 & 1 & 3 & 2 \\ 6 & 1 & 4 & 11 \end{pmatrix}$$

Lösung:
1.) Umformung auf verallgemeinerte Zeilenstufenform:

$$\begin{array}{cccc|cccc} 2 & 3 & 12 & 6 & 1 & 0 & 0 & 0 \\ 4 & 3 & 10 & 5 & 0 & 1 & 0 & 0 \\ 2* & 1 & 3 & 2 & 0 & 0 & 1 & 0 \\ 6 & 1 & 4 & 11 & 0 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{cccc|cccc} 0 & 2 & 9 & 4 & 1 & 0 & -1 & 0 \\ 0 & 1* & 4 & 1 & 0 & 1 & -2 & 0 \\ 2* & 1 & 3 & 2 & 0 & 0 & 1 & 0 \\ 0 & -2 & -5 & 5 & 0 & 0 & -3 & 1 \end{array}$$

$$\begin{array}{cccc|cccc} 0 & 0 & 1* & 2 & 1 & -2 & 3 & 0 \\ 0 & 1* & 4 & 1 & 0 & 1 & -2 & 0 \\ 2* & 1 & 3 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 7 & 0 & 2 & -7 & 1 \end{array}$$

$$\begin{array}{cccc|cccc} 0 & 0 & 1* & 2 & 1 & -2 & 3 & 0 \\ 0 & 1* & 4 & 1 & 0 & 1 & -2 & 0 \\ 2* & 1 & 3 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1* & -3 & 8 & -16 & 1 \end{array}$$

Der Rang ist 4, es gibt eine Inverse!
2.) Rücksubstitutionsschritte:

$$\begin{array}{cccc|cccc} 0 & 0 & 1* & 0 & 7 & -18 & 35 & -2 \\ 0 & 1* & 4 & 0 & 3 & -7 & 14 & -1 \\ 2* & 1 & 3 & 0 & 6 & -16 & 33 & -2 \\ 0 & 0 & 0 & 1* & -3 & 8 & -16 & 1 \end{array}$$

$$\begin{array}{cccc|cccc} 0 & 0 & 1* & 0 & 7 & -18 & 35 & -2 \\ 0 & 1* & 0 & 0 & -25 & 65 & -126 & 7 \\ 2* & 1 & 0 & 0 & -15 & 38 & -72 & 4 \\ 0 & 0 & 0 & 1* & -3 & 8 & -16 & 1 \end{array}$$

$$\begin{array}{cccc|cccc} 0 & 0 & 1* & 0 & 7 & -18 & 35 & -2 \\ 0 & 1* & 0 & 0 & -25 & 65 & -126 & 7 \\ 2* & 0 & 0 & 0 & 10 & -27 & 54 & -3 \\ 0 & 0 & 0 & 1* & -3 & 8 & -16 & 1 \end{array}$$

Dividieren durch Pivotelemente:

$$\begin{array}{cccc|cccc} 0 & 0 & 1* & 0 & 7 & -18 & 35 & -2 \\ 0 & 1* & 0 & 0 & -25 & 65 & -126 & 7 \\ 1* & 0 & 0 & 0 & 5 & -13,5 & 27 & -1,5 \\ 0 & 0 & 0 & 1* & -3 & 8 & -16 & 1 \end{array}$$

Umordnen der Zeilen:

$$\begin{array}{cccc|cccc} 1* & 0 & 0 & 0 & 5 & -13,5 & 27 & -1,5 \\ 0 & 1* & 0 & 0 & -25 & 65 & -126 & 7 \\ 0 & 0 & 1* & 0 & 7 & -18 & 35 & -2 \\ 0 & 0 & 0 & 1* & -3 & 8 & -16 & 1 \end{array}$$

Die Inverse von A ist

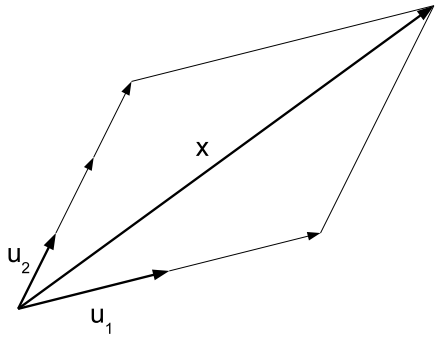
$$A^{-1} = \begin{pmatrix} 5 & -13,5 & 27 & -1,5 \\ -25 & 65 & -126 & 7 \\ 7 & -18 & 35 & -2 \\ -3 & 8 & -16 & 1 \end{pmatrix}$$

□

4.4. **Koordinaten.**

4.4.1. *Einführung von Koordinaten.*

Ist U ein endlichdimensionaler Vektorraum mit einer Basis $B = (u_1, \cdots, u_n)$, so lässt sich jedes $x \in U$ in eindeutigerweise in der Form $x = \sum_{i=1}^n \xi_i u_i$ schreiben. Wir nennen den Vektor $(\xi_1, \cdots, \xi_n)^T$ den Koordinatenvektor von x bezüglich der Basis B. In der folgenden Grafik hat der Vektor x bezüglich der Basis (u_1, u_2) den Koordinatenvektor $(2, 3)^T$.



Wir werden ein Koordinatensystem nun als Isomorphismus zwischen U und $\mathbb{K}^n$ interpretieren. Das sieht zunächst ein wenig umständlich aus, hilft uns aber später, die Übersicht zu behalten, wenn wir zwischen verschiedenen Koordinatensystemen hin und her wechseln.

Satz 4.4.1. *Sei U ein n-dimensionaler Vektorraum über $\mathbb{K}$ und $B = (u_1, \cdots, u_n)$ eine Basis von U. Seien $e_1, \cdots, e_n$ die Einheitsvektoren in $\mathbb{K}^n$.*

1. *Dann gibt es eine eindeutig bestimmte lineare Abbildung $\Phi_B : \mathbb{K}^n \to U$ sodass $\Phi_B(e_i) = u_i$.*
2. *Die Abbildung $\Phi_B : \mathbb{K}^n \to U$ ist ein Isomorphismus.*
3. *Jedes $x \in U$ läßt sich bekanntlich in eindeutiger Weise in der Form schreiben: $x = \sum_{i=1}^n \xi_i u_i$. Es gilt:*

$$\Phi_B^{-1}(x) = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

Beweis. Da $(e_1, \cdots, e_n)$ eine Basis $\mathbb{K}^n$ ist, gibt es nach Satz 4.1.12 eine eindeutige Abbildung $\Phi_B : \mathbb{K}^n \to V$ mit $\Phi_B(e_i) = u_i$. Diese Abbildung hat die Gestalt

$$\Phi_B(\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}) = \Phi_B(\sum_{i=1}^n \xi_i e_i) = \sum_{i=1}^n \xi_i u_i \,.$$

Weil $(u_1, \cdots, u_n)$ Basis von U ist, ist wegen Satz 4.1.12 die Abbildung Φ_B ein Isomorphismus, und es folgt

$$\Phi_B^{-1}(\sum_{i=1}^n \xi_i u_i) = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

□

Definition 4.4.2. Sei U ein Vektorraum über $\mathbb{K}$ und $B = (u_1, \cdots, u_n)$ eine Basis von U. Seien $e_1, \cdots, e_n$ die Einheitsvektoren in $\mathbb{K}^n$.
Wir definieren die Abbildung $\Phi_B : \mathbb{K}^n \to U$ als die eindeutig bestimmte lineare Abbildung, für die gilt:

$$(\forall i = 1, \cdots, n)\ \Phi_B(e_i) = u_i.$$

Ist $x \in U$, so heißt $\Phi_B^{-1}(x)$ der Koordinatenvektor von x bezüglich der Basis B.

Satz 4.4.3. *Seien U und V zwei endlichdimensionale Vektorräume über einem gemeinsamen Körper $\mathbb{K}$. Dann gilt: U ist genau dann isomorph zu V, wenn die Dimensionen von U und V gleich sind.*

Beweis. Seien zunächst U und V zwei isomorphe Vektorräume, sei $f : U \to V$ ein Isomorphismus, und sei $(u_1, \cdots, u_n)$ eine Basis von U. Wegen Satz 4.1.12 ist $(f(u_1), \cdots, f(u_n))$ eine Basis von V. Damit besitzen sowohl U als auch V je eine Basis aus n Vektoren, sie haben also dieselbe Dimension n.

Seien nun umgekehrt U und V zwei Vektorräume mit derselben Dimension n. Sei $B = (u_1, \cdots, u_n)$ eine Basis von U, und $C = (v_1, \cdots, v_n)$ eine Basis von V. Wir definieren die Koordinaten $\Phi_B : \mathbb{K}^n \to U$ und $\Phi_C : \mathbb{K}^n \to V$, diese Abbildungen sind Isomorphismen. Daher ist auch $\Phi_C \circ \Phi_B^{-1} : U \to V$ ein Isomorphismus. □

Wenn wir also die Theorie des $\mathbb{K}^n$ beherrschen, beherrschen wir in Wirklichkeit die Theorie aller endlichdimensionalen Vektorräume.

4.4.2. *Matrix eines Homomorphismus bezüglich gegebener Basen.*

Bemerkung 4.4.4. Seien V und W zwei endlichdimensionale Vektorräume über demselben Körper $\mathbb{K}$. Sei $f : V \to W$ eine lineare Abbildung. Es sei $B = (b_1, \cdots, b_n)$ eine Basis von V und $C = (c_1, \cdots, c_m)$ eine Basis von W. Mit Φ_B und Φ_C bezeichen wir wie üblich die Isomorphismen

$$\Phi_B : \begin{cases} \mathbb{K}^n & \to V \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} & \mapsto \sum_{j=1}^n \xi_j b_j \end{cases}, \qquad \Phi_C : \begin{cases} \mathbb{K}^m & \to W \\ \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} & \mapsto \sum_{i=1}^m \eta_i c_i \end{cases}.$$

Dann ist

$$\Phi_C^{-1} \circ f \circ \Phi_B : \mathbb{K}^n \to \mathbb{K}^m$$

der Homomorphismus, der für jedes $x \in V$ den Koordinatenvektor $\Phi_B^{-1}(x)$ in den Koordinatenvektor $\Phi_C^{-1}(f(x))$ des Bildes $f(x)$ überführt.

Definition 4.4.5. Seien V und W zwei endlichdimensionale Vektorräume über demselben Körper $\mathbb{K}$. Sei $f : V \to W$ eine lineare Abbildung. Es sei $B = (b_1, \cdots, b_n)$ eine Basis von V und $C = (c_1, \cdots, c_m)$ eine Basis von W, und Φ_B, Φ_C die entsprechenden Koordinatenabbildungen.
Dann bezeichnet $M_C^B(f) \in \mathbb{K}^{m \times n}$ die Matrix des Homomorphismus $\Phi_C^{-1} \circ f \circ \Phi_B$.
$M_C^B(f)$ heißt die Matrix von f bezüglich der Basen B und C.

Das folgende Diagramm macht diese Situation anschaulich:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \uparrow{\scriptstyle \Phi_B} & & \downarrow{\scriptstyle \Phi_C^{-1}} \\ \mathbb{K}^n & \xrightarrow{M_C^B(f)} & \mathbb{K}^m \end{array}$$

Jeder Pfeil stellt einen Homomorphismus dar. Wird ein Pfeil durch einen weiteren fortgesetzt, drückt das die Hintereinanderausführung der beiden Homomorphismen aus. Führen mehrere Wege (also mehrere Hintereinanderausführungen von Homomorphismen) von einem Raum zum anderen, so ergeben die verschiedenen Hintereinanderausführungen dieselbe Abbildung: Wir nennen ein solches Diagramm ein kommutatives Diagramm. Zum Beispiel ist oben: $M_C^B(f) = \Phi_C^{-1} \circ f \circ \Phi_B$.

Satz 4.4.6. *Seien V und W zwei endlichdimensionale Vektorräume über demselben Körper $\mathbb{K}$. Sei $f : V \to W$ eine lineare Abbildung. Es sei $B = (b_1, \cdots, b_n)$ eine Basis von V und $C = (c_1, \cdots, c_m)$ eine Basis von W, und Φ_B, Φ_C die entsprechenden Koordinatenabbildungen. Sei $M_C^B(f)$ die Matrix von f bezüglich der Basen B und C. Dann gilt:*
Die Spalten der Matrix $M_C^B(f)$ sind die Koordinaten der Vektoren $f(b_1), \cdots, f(b_n)$ bezüglich der Basis C.

Beweis. Die Spalten von $M_C^B(f)$ sind die Bilder $M_C^B(e_1), \cdots, M_C^B(e_n)$ der Einheitsvektoren $e_1, \cdots, e_n \in \mathbb{K}^n$. Nun ist $\Phi_B(e_j) = b_j$, und

$$M_C^B(f)e_j = \Phi_C^{-1} \circ f \circ \Phi_B(e_j) = \Phi_C^{-1}(f(b_j)).$$

Dies ist genau der Koordinatenvektor von $f(b_j)$. □

Satz 4.4.7. *Seien U, V, W drei Vektorräume über $\mathbb{K}$ mit den (endlichen) Dimensionen n, m, k. Seien $f : U \to V$ und $g : V \to W$ lineare Abbildungen. Seien B, C, D Basen von U, V, W und Φ_B, Φ_C, Φ_D die entsprechenden Koordinatenabbildungen. Es gilt:*

$$M_D^B(g \circ f) = M_D^C(g) M_C^B(f).$$

Beweis. Die Situation wird durch das folgende kommutative Diagramm verdeutlicht:

$$\begin{array}{ccccc} U & \xrightarrow{f} & V & \xrightarrow{g} & W \\ \uparrow{\scriptstyle \Phi_B} & & \uparrow{\scriptstyle \Phi_C} & & \uparrow{\scriptstyle \Phi_D} \\ \mathbb{K}^n & \xrightarrow{M_C^B(f)} & \mathbb{K}^m & \xrightarrow{M_D^C(g)} & \mathbb{K}^k \end{array}$$

Es ist

$$M_D^B(g \circ f) = \Phi_D^{-1} \circ (g \circ f) \circ \Phi_B = (\Phi_D^{-1} \circ g \circ \Phi^C) \circ (\Phi_C^{-1} \circ f \circ \Phi_B) = M_D^C(g) M_C^B(f).$$

□

Satz 4.4.8. *Seien V und W zwei endlichdimensionale Vektorräume über demselben Körper $\mathbb{K}$ mit den (endlichen) Dimensionen n und m. Sei $f : V \to W$ eine lineare Abbildung. Es seien B, C Basen von V, W, und Φ_B, Φ_C die entsprechenden Koordinatenabbildungen. Sei $M_C^B(f)$ die Matrix von f bezüglich der Basen B und C. Dann gilt:*
Die Abbildung f ist genau dann ein Isomorphismus, wenn $n = m$ und die Matrix $M_C^B(f)$ regulär ist. In diesem Fall ist

$$M_B^C(f^{-1}) = [M_C^B(f)]^{-1}.$$

Beweis. Wenn $n \neq m$, also wenn die Dimensionen von V und W verschieden sind, sind V und W bekanntlich nicht isomorph.

Sei nun $n = m$ und $M_C^B(f)$ regulär. Wir definieren den Homomorphismus $g : W \to V$ durch

$$g := \Phi_B \circ [M_C^B(f)]^{-1} \circ \Phi_C^{-1}.$$

Es gilt dann

$$\begin{aligned} & g \circ f = \Phi_B \circ [M_C^B(f)]^{-1} \circ \Phi_C^{-1} \circ f = \Phi_B \circ [M_C^B(f)]^{-1} \circ \Phi_C^{-1} \circ f \circ \Phi_B \circ \Phi_B^{-1} \\ = \quad & \Phi_B \circ [M_C^B(f)]^{-1} \circ M_C^B(f) \circ \Phi_B^{-1} = \mathrm{id}_V . \end{aligned}$$

Ebenso kann man zeigen:

$$f \circ g = \mathrm{id}_W .$$

Also besitzt f eine Inverse, nämlich $f^{-1} = g$.

Sei nun f ein Isomorphismus und $M_B^C(f^{-1})$ die Matrix der Umkehrabbildung bezüglich der Basen C und B. Es ist

$$M_B^C(f^{-1}) M_C^B(f) = \Phi_B^{-1} \circ f^{-1} \circ \Phi_C \circ \Phi_C^{-1} \circ f \circ \Phi_B = E_n.$$

Ebenso zeigt man

$$M_C^B(f) M_B^C(f^{-1}) = E_n.$$

Daher ist $M_B^C(f^{-1})$ die Inverse zu $M_C^B(f)$. □

Bemerkung 4.4.9. Seien V und W zwei Vektorräume über demselben Körper $\mathbb{K}$ mit (endlichen) Dimensionen n und m. Sei $f : V \to W$ eine lineare Abbildung. Seien B, C Basen von V, W und Φ_B, Φ_C die entsprechenden Koordinatenabbildungen. Sei $M_C^B(f)$ die Matrix von f bezüglich der Basen B und C. Dann gilt:

1) $\ker(f) = \Phi_B(\ker(M_C^B(f)))$, insbesondere ist f genau dann injektiv, wenn der Rang $\rho(M_C^B(f)) = n$ ist.
2) $\mathrm{rg}(f) = \Phi_C(\mathrm{rg}(M_C^B(f)))$, insbesondere ist f genau dann surjektiv, wenn der Rang $\rho(M_C^B(f)) = m$ ist, und es ist $\dim(\mathrm{rg}(f)) = \rho(M_C^B(f))$.

Beweis. Leichte (hoffentlich?) Übung. □

4.4.3. *Koordinatentransformation.*

Bemerkung 4.4.10. Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$. Seien $B = (b_1, \cdots, b_n)$ und $C = (c_1, \cdots, c_n)$ zwei Basen von V. Es seien Φ_B und Φ_C die Isomorphismen

$$\Phi_B : \begin{cases} \mathbb{K}^n & \to V \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} & \mapsto \sum_{j=1}^n \xi_j b_j \end{cases}, \qquad \Phi_C : \begin{cases} \mathbb{K}^n & \to V \\ \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} & \mapsto \sum_{i=1}^n \eta_i c_i \end{cases}.$$

Dann ist $\Phi_C^{-1} \Phi_B : \mathbb{K}^n \to \mathbb{K}^n$ jener Isomorphismus, der für jedes $v \in V$ den Koordinatenvektor $\Phi_B^{-1}(v)$ von v bezüglich B in den Koordinatenvektor $\Phi_C^{-1}(v)$ bezüglich C überführt.

Das folgende Diagramm macht die Situation anschaulich. Wir können $\Phi_C^{-1}\Phi_B$ auch als die Matrix $M_C^B(\mathrm{id}_V)$ auffassen.

$$\begin{array}{ccc} V & \xrightarrow{\mathrm{id}_V} & V \\ \uparrow{\scriptstyle \Phi_B} & & \downarrow{\scriptstyle \Phi_C^{-1}} \\ \mathbb{K}^n & \xrightarrow{T_C^B} & \mathbb{K}^n \end{array}$$

Definition 4.4.11. Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$. Seien $B = (b_1, \cdots, b_n)$ und $C = (c_1, \cdots, c_n)$ zwei Basen von V und Φ_B und Φ_C die entsprechenden Koordinatenabbildungen. Wir definieren

$$T_C^B = \Phi_C^{-1}\Phi_B.$$

Satz 4.4.12. *Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$. Seien $B = (b_1, \cdots, b_n)$ und $C = (c_1, \cdots, c_n)$ zwei Basen von V und Φ_B und Φ_C die entsprechenden Koordinatenabbildungen. Dann gilt:*

1) *Die Matrix T_C^B ist zugleich $M_C^B(\mathrm{id}_V)$.*
2) *Die Matrix T_C^B ist regulär, und ihre Inverse ist T_B^C.*
3) *Die Spalten von T_C^B sind die Basisvektoren $b_1, \cdots, b_n$, geschrieben in Koordinaten bezüglich C.*
4) *Die Spalten von $[T_C^B]^{-1}$ sind die Basisvektoren $c_1, \cdots, c_n$, geschrieben in Koordinaten bezüglich B.* [20]

Beweis. Punkt (1): Es ist

$$T_C^B = \Phi_C^{-1} \circ \Phi_B = \Phi_C^{-1} \circ \mathrm{id}_V \circ \Phi_B = M_C^B(\mathrm{id}_V).$$

Punkt (2): Da id_V ein Isomorphismus ist, ist nach Satz 4.4.8 $T_C^B = M_C^B(\mathrm{id}_V)$ regulär mit der Inversen

$$[T_C^B]^{-1} = M_B^C(\mathrm{id}_V^{-1}) = M_B^C(\mathrm{id}_V) = T_B^C.$$

Punkt (3): Nach Satz 4.4.6 die Spalten von $T_C^B = M_C^B(\mathrm{id}_V)$ die Koordinatenvektoren von

$$\mathrm{id}_V(b_1) = b_1, \ \cdots, \ \mathrm{id}_V(b_n) = b_n$$

bezüglich C.
Punkt (4): Punkt (3) und $[T_C^B]^{-1} = T_B^C$. □

Satz 4.4.13. *Sei V ein n-dimensionaler Vektorraum mit Basen A und B und dazugehörigen Koordinatenabbildungen $\Phi_A, \Phi_B : \mathbb{K}^n \to V$. Sei W ein m-dimensionaler Vektorraum mit Basen C und D und dazugehörigen Koordinatenabbildungen $\Phi_C, \Phi_D : \mathbb{K}^m \to W$. Sei $f : V \to W$ eine lineare Abbildung.*
Dann ist

$$M_D^A(f) = T_D^C M_C^B(f) T_B^A.$$

Beweis. Wir schreiben f als Hintereinanderausführung:

$$f = \mathrm{id}_W \circ f \circ \mathrm{id}_V,$$

wie das folgende kommutative Diagramm verdeutlicht:

$$\begin{array}{ccccccc} V & \xrightarrow{\mathrm{id}_V} & V & \xrightarrow{f} & W & \xrightarrow{\mathrm{id}_W} & W \\ \uparrow{\scriptstyle \Phi_A} & & \uparrow{\scriptstyle \Phi_B} & & \uparrow{\scriptstyle \Phi_C} & & \uparrow{\scriptstyle \Phi_D} \\ \mathbb{K}^n & \xrightarrow{T_B^A} & \mathbb{K}^n & \xrightarrow{M_C^B(f)} & \mathbb{K}^m & \xrightarrow{T_D^C} & \mathbb{K}^m \end{array}$$

[20] In manchen Aufgaben hat man die neuen Basisvektoren in alten Koordinaten gegeben und kann so zuerst $[T_C^B]^{-1}$ bestimmen und anschließend die Inverse berechnen.

Nach Satz 4.4.7 gilt

$$M_D^A(f) = M_D^C(\mathrm{id}_W) M_C^B(f) M_B^A(\mathrm{id}_V) = T_D^C M_C^B(f) T_B^A.$$

□

Korollar 4.4.14. *Sei V ein n-dimensionaler Vektorraum über $\mathbb{K}$, seien B und C zwei Basen von V mit den entsprechenden Koordinatenabbildungen $\Phi_B, \Phi_C : \mathbb{K}^n \to V$. Sei $f : V \to V$ ein Endomorphismus. Es gilt:*

$$M_C^C(f) = T_C^B M_B^B(f) T_B^C = [T_B^C]^{-1} M_B^B(f) T_B^C.$$

Beweis. Satz 4.4.13 und Satz 4.4.12. □

Definition 4.4.15. Seien $m, n \in \mathbb{N}$, $P, Q \in \mathbb{K}^{m \times n}$ Matrizen gleicher Dimension über einem Körper $\mathbb{K}$.

1) P und Q heißen äquivalente Matrizen, wenn es reguläre Matrizen $S \in \mathbb{K}^{m \times m}$ und $T \in \mathbb{K}^{n \times n}$ gibt, so dass $Q = S^{-1} P T$.
2) Sei nun zusätzlich $m = n$. Die Matrizen P und Q heißen ähnliche Matrizen, wenn es eine reguläre Matrix $T \in \mathbb{K}^{n \times n}$ gibt, so dass $Q = T^{-1} P T$.

Bemerkung 4.4.16. Seien $m, n \in \mathbb{N}$, $P, Q \in \mathbb{K}^{m \times n}$ Matrizen gleicher Dimension über einem Körper $\mathbb{K}$.

1) P und Q sind genau dann äquivalent, wenn es einen n-dimensionalen Vektorraum V mit zwei Basen A, B und einen m-dimensionalen Vektorraum W mit zwei Basen C, D gibt, sowie eine lineare Abbildung $f : V \to W$ gibt, so dass P und Q die Matrizen derselben Abbildung f bezüglich verschiedener Basen in V und W sind:
$$P = M_C^A(f), \quad Q = M_D^B(f).$$
2) Sei zusätzlich $m = n$. Die Matrizen P und Q sind genau dann ähnlich, wenn es einen n-dimensionalen Vektorraum V mit zwei Basen A, B und einen Endomorphismus $f : V \to V$ gibt, so dass P und Q die Matrizen derselben Abbildung f bezüglich verschiedener Basen in V sind:
$$P = M_A^A(f), \quad Q = M_B^B(f).$$

Beweis. Übung! Können Sie das? Tipp: Für V und W kann man $\mathbb{K}^n$ und $\mathbb{K}^m$ nehmen. Für die Basen A und C kann man die Einheitsbasen heranziehen. □

Man kann beweisen, dass zwei Matrizen gleicher Dimension genau dann äquivalent sind, wenn sie denselben Rang haben, daher liefert dieser Begriff nichts Neues und wird selten verwendet. Dagegen ist Ähnlichkeit eine viel stärkere Bedingung und spielt eine große Rolle, denn durch die Wahl einer geeigneten Basis kann man lineare Abbildungen durch besonders einfache Matrizen darstellen. Mehr darüber folgt im zweiten Semester.

Beispiel 4.4.17. Die Matrizen

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

sind äquivalent, aber nicht ähnlich.

Beweis. Wir betrachten den Homomorphismus

$$f : \begin{cases} \mathbb{R}^2 & \to \mathbb{R}^2 \\ x & \mapsto Px. \end{cases}$$

Wir betrachten zwei Basen, welche aus den beiden Einheitsvektoren e_1, e_2 bestehen und sich nur durch die Reihenfolge unterscheiden: $E = (e_1, e_2)$, die Einheitsbasis, und $B = (e_2, e_1)$. Beachten Sie, dass $f(e_1) = e_1$, $f(e_2) = 0$. Offensichtlich ist $M_E^E(f) = P$. Die Spalten von $M_E^B(f)$ sind die Vektoren $f(e_2), f(e_1)$ geschrieben in Koordinaten bezüglich der Einheitsbasis E, also ist $M_E^B(f) = Q$. Damit sind P und Q äquivalent.

Wären P und Q ähnlich, so wäre also auch $Q = M_C^C(f)$ bezüglich einer geeigneten Basis C. Es ist aber

$$\begin{aligned} M_B^B(f \circ f) &= M_B^B(f) M_B^B(f) = P^2 = P, \\ M_C^C(f \circ f) &= M_C^C(f) M_C^C(f) = Q^2 = 0. \end{aligned}$$

Daher ist einerseits $f \circ f = f$ und andererseits $f \circ f = 0$. Widerspruch! □

Beispiel 4.4.18. Gegeben sei der Endomorphismus

$$f : \begin{cases} \mathbb{R}^3 & \to \mathbb{R}^3, \\ x & \mapsto Px, \end{cases} \quad \text{mit } P = \begin{pmatrix} 4 & 0 & -2 \\ 0 & 2 & 2 \\ -2 & 2 & 3 \end{pmatrix}.$$

(P ist also die Matrix von f bezüglich der Einheitsbasis E.)
Wir betrachten in $\mathbb{R}^3$ die neue Basis

$$C = (\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}).$$

Wie sieht die Matrix $M_C^C(f)$ von f bezüglich der Basis C aus?

Lösung: Wir berechnen zunächst die Matrix zur Koordinatentransformation: Die Spalten von T_E^C sind die Basisvektoren von C in Einheitskoordinaten, also ist

$$T_E^C = \begin{pmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{pmatrix}.$$

Als nächstes berechnen wir $T_C^E = [T_E^C]^{-1}$:

$$\begin{array}{ccc|ccc} 2 & -2 & 1 & 1 & 0 & 0 \\ 2 & 1 & -2 & 0 & 1 & 0 \\ 1* & 2 & 2 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 0 & -6 & -3 & 1 & 0 & -2 \\ 0 & -3 & -6 & 0 & 1 & -2 \\ 1* & 2 & 2 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 0 & -2 & -1 & 1/3 & 0 & -2/3 \\ 0 & 1* & 2 & 0 & -1/3 & 2/3 \\ 1* & 2 & 2 & 0 & 0 & 1 \end{array}$$

$$\begin{array}{ccc|ccc} 0 & 0 & 3 & 1/3 & -2/3 & 2/3 \\ 0 & 1* & 2 & 0 & -1/3 & 2/3 \\ 1* & 0 & -2 & 0 & 2/3 & -1/3 \end{array}$$

$$\begin{array}{ccc|ccc} 0 & 0 & 1* & 1/9 & -2/9 & 2/9 \\ 0 & 1* & 2 & 0 & -1/3 & 2/3 \\ 1* & 0 & -2 & 0 & 2/3 & -1/3 \end{array}$$

$$\begin{array}{ccc|ccc} 0 & 0 & 1* & 1/9 & -2/9 & 2/9 \\ 0 & 1* & 0 & -2/9 & 1/9 & 2/9 \\ 1* & 0 & 0 & 2/9 & 2/9 & 1/9 \end{array}$$

$$\begin{array}{ccc|ccc} 1* & 0 & 0 & 2/9 & 2/9 & 1/9 \\ 0 & 1* & 0 & -2/9 & 1/9 & 2/9 \\ 0 & 0 & 1* & 1/9 & -2/9 & 2/9 \end{array}$$

Also ist

$$T_C^E = \frac{1}{9}\begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix}.$$

Nun ist

$$\begin{aligned} M_C^C(f) &= T_C^E P T_E^C \\ &= \frac{1}{9}\begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix}\begin{pmatrix} 4 & 0 & -2 \\ 0 & 2 & 2 \\ -2 & 2 & 3 \end{pmatrix}\begin{pmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{pmatrix} \\ &= \frac{1}{9}\begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix}\begin{pmatrix} 6 & -12 & 0 \\ 6 & 6 & 0 \\ 3 & 12 & 0 \end{pmatrix} \\ &= \frac{1}{9}\begin{pmatrix} 27 & 0 & 0 \\ 0 & 54 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Bezüglich der Basis C hat f die Diagonalmatrix

$$M_C^C(f) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

□

Im zweiten Semester werden wir lernen, wie man systematisch Basen finden kann, in denen sich Endomorphismen besonders einfach schreiben lassen.

4.5. **Faktorräume.**

4.5.1. *Äquivalenzrelationen.*

Definition 4.5.1. Sei M eine Menge. Eine Relation auf M ist eine Teilmenge von $M \times M$. Ist R eine Relation, und sind $a, b \in M$, so sagen wir: A steht zu b in der Relation R genau dann, wenn $(a, b) \in R$.

Für unseren Hausgebrauch stellen wir uns vor: Eine Relation ist eine Beziehung zwischen den Elementen von M, zwei Elemente stehen entweder in dieser Beziehung, oder nicht. (Ich hoffe, Sie merken, dass der eben geschriebene Satz zu einer mathematischen Definition nicht taugt!)

Die formale Definition 4.5.1 erlaubt, Relationen durch den Begriff der Menge einzuführen, womit man alle logischen und mengentheoretischen Schwierigkeiten, die man bei der axiomatischen Grundlegung des Mengenbegriffs überwinden muss, nicht noch einmal aufs Neue für die Einführung des Relationsbegriffs bewältigen muss.

Zum Beispiel: Die Gleichheit auf einer Menge ist eine Relation. Entweder zwei Elemente sind gleich (nämlich, wenn es in Wirklichkeit dasselbe Element ist) oder nicht. Auf der Menge $\mathbb{R}$ sind $\leq$ oder $<$ zwei weitere Relationen. In der Menge $\mathbb{Z}$ gibt es zum Beispiel auch die Relationen "a ist ein Teiler von b" oder "$a - b$ ist eine gerade Zahl".

Definition 4.5.2. Sei M eine Menge und $\simeq$ eine Relation auf M mit den folgenden drei Eigenschaften:

$$\begin{array}{lll} (\forall a \in M) & a \simeq a & \text{Reflexivität,} \\ (\forall a, b \in M) & a \simeq b \Rightarrow b \simeq a & \text{Symmetrie,} \\ (\forall a, b, c \in M) & [a \simeq b \wedge b \simeq c] \Rightarrow a \simeq c & \text{Transitivität.} \end{array}$$

Dann heißt $\simeq$ eine Äquivalenzrelation auf M.

Für Äquivalenzrelationen nimmt man gerne Symbole, welche dem Gleichheitszeichen ähnlich sehen. Gleichheit ist die einfachste Äquivalenzrelation.

Beispiel 4.5.3.

1) Sei M eine Menge. Die Gleichheit auf M ist eine Äquivalenzrelation.
2) Seien M, N Mengen und $f : M \to N$ eine Abbildung. Die Relation:
$$x \simeq y :\Leftrightarrow f(x) = f(y)$$
ist eine Äquivalenzrelation.
3) Sei $k \in \mathbb{N}$. Die Relation
$$x \simeq y :\Leftrightarrow k \text{ teilt } (y - x)$$
ist eine Äquivalenzrelation auf $\mathbb{Z}$.
4) Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Auf $\mathbb{K}^{n \times n}$ ist die Relation
$$A \simeq B :\Leftrightarrow A \text{ und } B \text{ sind ähnliche Matrizen}$$
eine Äquivalenzrelation.

Satz 4.5.4. *Sei M eine Menge.*

1) *Sei $\mathcal{U}$ eine Partition auf M*[21]*. Dann ist die folgende Relation eine Äquivalenzrelation auf M:*
$$x \simeq y :\Leftrightarrow (\exists U \in \mathcal{U})\ [x \in U \wedge y \in U].$$
Die folgende Definition erklärt eine surjektive Abbildung $\pi : M \to \mathcal{U}$:
$$\pi(x) = U :\Leftrightarrow x \in U.$$
Diese Abbildung nennt man die kanonische Projektion von M auf $\mathcal{U}$.
2) *Sei $\simeq$ eine Äquivalenzrelation auf M. Ist $x \in M$, so definieren wir die Äquivalenzklasse von x durch*
$$[x] := \{y \in M \mid x \simeq y\}.$$
Dann ist die Menge $\mathcal{U} := \{[x] \mid x \in M\}$ eine Partition auf M, und die Abbildung
$$\pi : \begin{cases} M & \to \mathcal{U} \\ x & \mapsto [x] \end{cases}$$
ist die dazugehörige kanonische Projektion.

Beweis. Punkt (1): Sei $\mathcal{U}$ eine Partition auf M. Sei $x \in M$. Nach Definition 2.3.10 (1) gibt es mindestens ein $U \in \mathcal{U}$ mit $x \in U$. Sind $U, V \in \mathcal{U}$ mit $x \in U$ und $x \in V$, so gilt $U = V$ nach Definition 2.3.10 (2). Daher gibt es genau ein $U \in \mathcal{U}$ mit $x \in U$, und damit ist π eine Abbildung von M nach $\mathcal{U}$. Weil nach Definition 2.3.10 (3) jedes $U \in \mathcal{U}$ mindestens ein Element enthält, ist π surjektiv.
Wir betrachten nun die Relation $\simeq$ und überprüfen die drei Eigenschaften der Äquivalenzrelationen. Zunächst gibt es für jedes $x \in M$ eine Klasse $U \in \mathcal{U}$ mit $x \in U$. Somit ist $x \in U \wedge x \in U$, also $x \simeq x$.

[21] siehe Definition 2.3.10

Sind $x, y \in M$ mit $x \simeq y$, also gibt es eine Klasse $U \in \mathcal{U}$ mit $x \in U \wedge y \in U$, das ist dasselbe wie $y \in U \wedge x \in U$. Also ist auch $y \simeq x$.
Sind $x, y, z \in M$ mit $x \simeq y$ und $y \simeq z$. Also gibt es Klassen $U, V \in \mathcal{U}$ mit $x, y \in U$, $y, z \in V$. Da $y \in U \cap V$, ist $U = V$ wegen Definition 2.3.10 (2) und wir haben $x, z \in U$. Also gilt $x \simeq z$.
Punkt (2) Sei nun $\simeq$ eine Äquivalenzrelation auf M und zu jedem $x \in M$ sei $[x]$ die Äquivalenzklasse von x. Wir zeigen die drei Eigenschaften einer Partition. Zunächst gilt für jedes $x \in M$ wegen der Reflexivität $x \simeq x$, daher ist $x \in [x]$. Also ist jedes x in mindestens einer Klasse enthalten.
Sind $x, y \in M$ so, dass $[x] \cap [y] \neq \emptyset$, das heißt, es gibt ein $z \in M$ mit $x \simeq z$ und $y \simeq z$. Wir zeigen: $[x] \subset [y]$: Sei $u \in [x]$, also $u \simeq x$. Dann gelten also (unter Ausnützung der Symmetrie) folgende Relationen:

$$y \simeq z,\ z \simeq x,\ x \simeq u.$$

Also ist wegen der Transitivität $y \simeq u$ und folglich $u \in [y]$. Ebenso zeigt man $[y] \subset [x]$, also ist $[y] = [x]$.
Da $x \in [x]$, ist keine Klasse $[x]$ leer. □

Schreibweise 4.5.5. Es gibt keine normierte Schreibweise für die Äquivalenzklasse von x, jedoch werden die Schreibweisen $\pi(x)$, $[x]$ und $\overline{x}$ häufig verwendet.

4.5.2. *Nebenklassen und Faktorräume.*
Wir zeigen zunächst mit Zeichnungen einige Nebenklassen eines eindimensionalen Unterraumes von $\mathbb{R}^2$, und Nebenklassen von Unterräumen des $\mathbb{R}^3$.

Satz 4.5.6. *Sei V ein Vektorraum über $\mathbb{K}$, und W ein Unterraum von V. Wir definieren die folgende Relation auf V:*

$$x \underset{W}{\equiv} y :\Leftrightarrow y - x \in W.$$

Dann gilt:

1) *Die Relation $\underset{W}{\equiv}$ ist eine Äquivalenzrelation auf V.*
2) *Die Relation $\underset{W}{\equiv}$ ist im folgenden Sinn verträglich mit den Vektorraumoperationen: Seien $x_1, x_2, y_1, y_2 \in V$, $\lambda \in \mathbb{K}$ so, dass $x_1 \underset{W}{\equiv} x_2$ und $y_1 \underset{W}{\equiv} y_2$. Dann ist*
$$x_1 + y_1 \underset{W}{\equiv} x_2 + y_2 \quad \text{und} \quad \lambda x_1 \underset{W}{\equiv} \lambda x_2.$$
3) *Die Äquivalenzklasse $[x]$ von x bezüglich der Relation $\underset{W}{\equiv}$ ist*
$$[x] = x + W = \{x + w \mid w \in W\}.$$

Beweis. Punkt (1): Wir zeigen die drei Eigenschaften der Äquivalenzrelation. Wir werden immer wieder verwenden, dass W ein Unterraum ist. Zunächst gilt $x - x = 0 \in W$, also $x \underset{W}{\equiv} x$ für jedes $x \in V$. Sind $x, y \in V$ mit $x \underset{W}{\equiv} y$, also $v := y - x \in W$. Dann ist $x - y = -v \in W$, und damit $y \underset{W}{\equiv} x$. Sind $x, y, z \in V$ mit $x \underset{W}{\equiv} y$ und $y \underset{W}{\equiv} z$, so ist also $v := y - x \in W$, und $w = z - y \in W$. Damit ist $z - x = w + v \in W$ und folglich $x \underset{W}{\equiv} z$.
Punkt (2): Sei $x_1 \underset{W}{\equiv} x_2$, d.h., $v := x_2 - x_1 \in W$, und sei $y_1 \underset{W}{\equiv} y_2$, d.h. $w := y_2 - y_1 \in W$. Sei $\lambda \in \mathbb{K}$. Es ist dann $(x_2 + y_2) - (x_1 + y_1) = v + w \in W$, also $x_1 + y_1 \underset{W}{\equiv} x_2 + y_2$, und $\lambda x_2 - \lambda x_1 = \lambda v \in W$, also $\lambda x_1 \underset{W}{\equiv} \lambda x_2$.
Punkt (3): Es ist $y \in [x]$ genau dann, wenn $y \underset{W}{\equiv} x$, also, wenn es ein $w \in W$ gibt mit $y - x = w$, d.h. aber, $y = x + w$. □

Definition 4.5.7. Sei V ein Vektorraum über $\mathbb{K}$, und W ein Unterraum von V. Sei $\underset{W}{\equiv}$ und $[x]$ definiert wie in Satz 4.5.6.

1) Zu jedem $x \in V$ heißt die Äquivalenzklasse $[x] = x + W$ die Nebenklasse oder Faser von x bezüglich W.
2) Die Partition $V/W := \{[x] \mid x \in V\}$ heißt der Faktorraum von V bezüglich W.

Satz 4.5.8.

1) *Die folgenden Operationen auf V/W sind wohldefiniert: Seien $x, y \in V$ und $[x], [y] \in V/W$ ihre Nebenklassen. Sei $\lambda \in \mathbb{K}$.*
$$[x] + [y] := [x + y], \quad \lambda[x] := [\lambda x].$$
2) *Mit diesen Operationen ist V/W ein Vektorraum über $\mathbb{K}$. Der Nullvektor in V/W ist $[0] = W$.*
3) *Die kanonische Projektion*
$$\pi : \begin{cases} V & \to V/W \\ x & \mapsto [x] \end{cases}$$
ist eine surjektive lineare Abbildung mit Kern W.

Beweis. Punkt (1): Jede Nebenklasse $[x] \in V/W$ lässt sich als Nebenklasse verschiedener Elemente aus V zeigen. Wir müssen zeigen, dass die Definition von der Wahl dieser Elemente unabhängig ist. Sei also $[x_1] = [x_2]$, $[y_1] = [y_2]$. Wir müssen zeigen $[x_1 + y_1] = [x_2 + y_2]$. Es gilt aber $x_1 \underset{W}{\equiv} x_2$ weil $x_2 \in [x_2] = [x_1]$ und $y_1 \underset{W}{\equiv} y_2$, also nach Satz 4.5.6 auch $x_1 + y_1 \underset{W}{\equiv} x_2 + y_2$ und $\lambda x_1 \underset{W}{\equiv} \lambda x_2$. Damit ist aber $[x_1 + y_1] = [x_2 + y_2]$ und $[\lambda x_1] = [\lambda x_2]$.

Punkt (2): Hier muss man die Vektorraumaxiome Punkt für Punkt durchrechnen. Wir zeigen zum Beispiel eines der Distributivgesetze: Seien $[x], [y] \in V/W$, $\lambda \in \mathbb{K}$. Es ist dann (unter Benützung von Punkt (1) und des Distributivgesetzes in V):

$$\lambda([x] + [y]) = \lambda[x + y] = [\lambda(x + y)] = [\lambda x + \lambda y] = [\lambda x] + [\lambda y] = \lambda[x] + \lambda[y].$$

Die anderen Axiome gehen ähnlich. Es ist $[-x] = -[x]$ und der Nullvektor in V/W ist $[0]$. Es ist aber $[0] = 0 + W = W$.

Punkt (3): Als kanonische Projektion auf die Klassen einer Äquivalenzrelation ist π eine surjektive Abbildung. Es ist wegen Punkt (2):

$$\pi(\lambda x + \mu y) = [\lambda x + \mu y] = \lambda[x] + \mu[y],$$

also ist π linear. Letztlich ist

$$x \in \ker(\pi) \Leftrightarrow [x] = [0] \Leftrightarrow x \in W.$$

□

Bemerkung 4.5.9. Sei V ein Vektorraum über $\mathbb{K}$ und W ein Unterraum von V.

1) Dann gibt es einen Vektorraum Z und eine lineare Abbildung $f : V \to Z$ so, dass $\ker(f) = W$.
2) Ist zusätzlich $\dim(V) < \infty$, dann gilt
$$\dim(V/W) = \dim(V) - \dim(W).$$

Beweis. Punkt (1): Wähle $Z = V/W$ und $f = \pi$.
Punkt (2): Es ist $W = \ker(\pi)$ und $V/W = \mathrm{rg}(\pi)$. Nun folgt die Bemerkung aus Satz 4.1.11, angewendet auf die lineare Abbildung π. □

Die Konstruktion von Faktorstrukturen wird in der Algebra oft verwendet: Analog kann man für Gruppen oder Ringe Faktorstrukturen definieren. Während aber jeder Unterraum eines Vektorraums die Bildung eines Faktorraumes erlaubt, und damit Kern einer linearen Abbildung sein kann, kann bei nicht-kommutativen Gruppen nur bezüglich ganz speziellen Untergruppen, den sogenannten Normalteilern, eine Faktorgruppe gebildet werden. Auch nicht jeder Unterring eines Ringes eignet sich zur Bildung eines Faktorringes, sondern nur die sogenannten Ideale.

4.5.3. *Der Faktorisierungssatz.*

Satz 4.5.10. *Seien U, V Vektorräume über $\mathbb{K}$, sei $f : V \to U$ eine lineare Abbildung. Es seien*

$$\pi : \begin{cases} V & \to V/\ker(f) \\ x & \mapsto [x] \end{cases}, \quad j : \begin{cases} \mathrm{rg}(f) & \to U \\ y & \mapsto y \end{cases}.$$

Dann gibt es einen Isomorphismus $\tilde{f} : V/\ker(f) \to \mathrm{rg}(f)$ so, dass $f = j \circ \tilde{f} \circ \pi$.

Das folgende Diagramm wird also kommutativ:

$$\begin{array}{ccc} V & \xrightarrow{f} & U \\ {\scriptstyle\pi}\big\downarrow{\scriptstyle\text{surjektiv}} & & {\scriptstyle j}\big\uparrow{\scriptstyle\text{injektiv}} \\ V/\ker(f) & \xrightarrow[\text{bijektiv}]{\tilde{f}} & \mathrm{rg}(f) \end{array}$$

Beweis. Wir definieren

$$\tilde{f} : \begin{cases} V/\ker(f) & \to \mathrm{rg}(f), \\ [x] & \mapsto f(x). \end{cases}$$

Wir müssen zunächst zeigen, dass $\tilde{f}$ wohldefiniert ist. Jede Nebenklasse aus $V/\ker(f)$ lässt sich in der Form $[x]$ mit $x \in V$ schreiben. Ist $[x_1] = [x_2]$, so müssen wir zeigen, dass $f(x_1) = f(x_2)$. Es gilt aber

$$[x_1] = [x_2] \Leftrightarrow x_2 - x_1 \in \ker(f) \Leftrightarrow f(x_2) - f(x_1) = 0 \Leftrightarrow f(x_1) = f(x_2).$$

Wir haben damit gezeigt, dass $\tilde{f}$ wohldefiniert ist, und zugleich erhalten, dass $\tilde{f}$ injektiv ist. Da jedes $y \in \mathrm{rg}(f)$ sich in der Form $y = f(x)$ mit einem $x \in V$ schreiben lässt, ist

$$y = f(x) = \tilde{f}([x]),$$

also ist $\tilde{f}$ surjektiv. Wir zeigen, dass $\tilde{f}$ linear ist:

$$\tilde{f}(\lambda[x] + \mu[y]) = \tilde{f}([\lambda x + \mu y]) = f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) = \lambda \tilde{f}([x]) + \mu \tilde{f}([y]).$$

□

Korollar 4.5.11. *Seien U, V, W Vektorräume über $\mathbb{K}$, seien $f : V \to U$ und $g : V \to W$ lineare Abbildungen, und sei $\ker(f) = \ker(g)$. Dann sind die Bilder $\mathrm{rg}(f)$ und $\mathrm{rg}(g)$ isomorph.*

Beweis. Beide Bilder sind isomorph zu $V/\ker(f)$. □

5. Determinanten

5.1. **Definition der Determinante.** Wir erklären zunächst an Zeichnungen anschaulich Inhalt und Orientierung

Definition 5.1.1. Sei $\mathbb{K}$ ein Körper, $n \in \mathbb{N}$ und $\det : \mathbb{K}^{n\times n} \to \mathbb{K}$ eine Funktion. Die Funktion det heißt eine Determinante, wenn die folgenden Eigenschaften erfüllt sind:

1) Die Funktion det ist linear in den Zeilen, d.h., sind
$$A = \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & a_i & - \\ & \vdots & \\ - & a_n & - \end{pmatrix}, \quad B = \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & b_i & - \\ & \vdots & \\ - & a_n & - \end{pmatrix}$$
zwei Matrizen mit Zeilen a_k, b_k, die sich nur in der i-ten Zeile unterscheiden, und sind $\lambda, \mu \in \mathbb{K}$, so ist
$$\det \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & \lambda a_i + \mu b_i & - \\ & \vdots & \\ - & a_n & - \end{pmatrix} = \lambda \det \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & a_i & - \\ & \vdots & \\ - & a_n & - \end{pmatrix} + \mu \det \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & b_i & - \\ & \vdots & \\ - & a_n & - \end{pmatrix}.$$
(Man sagt auch, die Determinante ist multilinear.)
2) Hat A zwei gleiche Zeilen, so ist $\det(A) = 0$:
$$\det \begin{pmatrix} & \vdots & \\ - & a_i & - \\ & \vdots & \\ - & a_i & - \\ & \vdots & \end{pmatrix} = 0.$$
3) Die Determinante der Einheitsmatrix ist $\det(E_n) = 1$. (Man sagt auch, det ist normiert.)

Bemerkung 5.1.2. Wir werden im Zuge von Kapitel 5 zeigen, dass es auf $\mathbb{K}^{n\times n}$ genau eine Determinante gibt. Zunächst werden wir zeigen, dass — wenn es überhaupt eine Determinante gibt — diese mit Hilfe von Pivottransformationen zu berechnen geht. Das zeigt auch, dass es höchstens eine Determinante gibt.

Erst in Unterabschnitt 5.3 werden wir, mit Hilfe einer komplizierten, für die praktische Rechnung wenig nützlichen Formel beweisen können, dass es eine Determinante gibt. Damit ist dann gezeigt, dass es genau eine Determinante gibt.

Erst damit wird die folgende Schreibweise endgültig gerechtfertigt, die wir bereits jetzt verwendet haben: $\det : \mathbb{K}^{n\times n} \to \mathbb{K}$ bezeichnet die Determinante. Statt $\det(A)$ schreibt man auch oft $|A|$, oder die Matrix A zwischen senkrechten Strichen statt Klammern.

Bemerkung 5.1.3. Die Determinante ist linear in den einzelnen Zeilen, aber im Allgemeinen gilt
$$\det(\lambda A) \neq \lambda \det(A), \quad \det(A + B) \neq \det(A) + \det(B).$$

Statt dessen gilt für $A \in \mathbb{K}^{n\times n}$:

$$\det(\lambda A) = \lambda^n \det(A).$$

5.1.1. *Berechnung der Determinante mit Pivottransformationen.*

Der folgende Satz zeigt, wie sich eine Determinante bei Anwendung von elementaren Zeilentransformationen verhält (vgl. Definition 3.3.12).

Satz 5.1.4. *Sei $\mathbb{K}$ ein Körper und $\det : \mathbb{K}^{n\times n} \to \mathbb{K}$ eine Determinante. Dann gilt:*

1) *Entsteht B aus A, indem eine Zeile von A mit einem Skalar λ multipliziert wird, so ist $\det(B) = \lambda \det(A)$.*
2) *Entsteht B aus A, indem ein Vielfaches einer Zeile von A zu einer anderen Zeile von A addiert wird, so ist $\det(B) = \det(A)$.*
3) *Entsteht B aus A durch Vertauschung von zwei Zeilen, so ist $\det(B) = -\det(A)$. (Man sagt auch, die Determinante ist alternierend.)*
4) *Sind die Zeilen von A linear abhängig, so ist $\det(A) = 0$.*

Beweis. Punkt (1): ist eine unmittelbare Folge von Definition 5.1.1 (1).

Punkt (2): Wir benützen erst Definition 5.1.1 (1), dann Definition 5.1.1 (2):

$$\begin{vmatrix} \vdots \\ - \quad a_i + \lambda a_j \quad - \\ \vdots \\ - \quad a_j \quad - \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ - \quad a_i \quad - \\ \vdots \\ - \quad a_j \quad - \\ \vdots \end{vmatrix} + \lambda \begin{vmatrix} \vdots \\ - \quad a_j \quad - \\ \vdots \\ - \quad a_j \quad - \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ - \quad a_i \quad - \\ \vdots \\ - \quad a_j \quad - \\ \vdots \end{vmatrix} + 0.$$

Punkt (3) Wir benützen merhmals den eben bewiesenen Punkt (2), und ganz am Ende die Linearität in den Zeilen:

$$\begin{vmatrix} \vdots \\ - \quad a_j \quad - \\ \vdots \\ - \quad a_i \quad - \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ - \quad a_j + a_i \quad - \\ \vdots \\ - \quad a_i \quad - \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ - \quad a_j + a_i \quad - \\ \vdots \\ - \quad a_i - (a_j + a_i) \quad - \\ \vdots \end{vmatrix}$$

$$= \begin{vmatrix} \vdots \\ - \quad a_j + a_i \quad - \\ \vdots \\ - \quad -a_j \quad - \\ \vdots \end{vmatrix} = \begin{vmatrix} \vdots \\ - \quad a_i \quad - \\ \vdots \\ - \quad -a_j \quad - \\ \vdots \end{vmatrix} = - \begin{vmatrix} \vdots \\ - \quad a_i \quad - \\ \vdots \\ - \quad a_j \quad - \\ \vdots \end{vmatrix}$$

Punkt (4) Sei A eine Matrix mit den Zeilen $a_1, \cdots, a_n$. Sind die Zeilen von A linear abhängig, so gibt es einen Index i und Koeffizienten $\lambda_k \in \mathbb{K}$ (für $k \neq i$), sodass

$$a_i = \sum_{k\neq i} \lambda_k a_k.$$

Wir verwenden nun die Multilinearität und anschließend Definition 5.1.1 (2):

$$\begin{vmatrix} - & a_1 & - \\ & \vdots & \\ - & \sum_{k\neq i}\lambda_k a_k & - \\ & \vdots & \\ - & a_n & - \end{vmatrix} = \sum_{k\neq i} \begin{vmatrix} & \vdots & \\ - & \lambda_k a_k & - \\ & \vdots & \\ - & a_k & - \\ & \vdots & \end{vmatrix} = \sum_{k\neq i} \lambda_k \begin{vmatrix} & \vdots & \\ - & a_k & - \\ & \vdots & \\ - & a_k & - \\ & \vdots & \end{vmatrix} = 0.$$

□

Bemerkung 5.1.5. Wir wissen jetzt: Ist det eine Determinantenfunktion, und ist A eine singuläre Matrix, so ist $\det(A) = 0$. Später werden wir beweisen: A ist genau dann singulär, wenn $\det(A) = 0$.

Algorithmus 5.1.6. Sei $\mathbb{K}$ ein Körper, $\det : \mathbb{K}^{n\times n} \to \mathbb{K}$ eine Determinantenfunktion. Sei $A \in \mathbb{K}^{n\times n}$. Das folgende Verfahren liefert $\det(A)$:

1) Setze $D = 1$.
2) Führe A durch Pivotschritte in verallgemeinerte Zeilenstufenform über. Dabei ändere bei Bedarf D:
   - Wird das Vielfache einer Zeile zu einer anderen addiert, wird D nicht verändert.
   - Wird eine Zeile mit $\lambda \neq 0$ multipliziert, so wird D durch λ dividiert.
   - Werden zwei Zeilen vertauscht, so wird das Vorzeichen von D umgekehrt.
3) Wenn bei der Umformung eine Nullzeile entsteht, so ist $\det(A) = 0$. Stop.
4) Wenn bei der Umformung keine Nullzeile entsteht, wurden alle Zeilen als Pivotzeilen verwendet. Ordne die Zeilen durch Vertauschung so um, dass die Pivotelemente auf der Diagonalen stehen. Bei jeder Vertauschung ändere das Vorzeichen von D. (Hinweis: Die Nicht-Pivotelemente spielen keine weitere Rolle, man darf sie ohne Weiteres durch 0 ersetzen.)
5) Multipliziere D mit dem Produkt der Pivotelemente.
6) $\det(A) = D$.

Beweis. Durch Schritt (2) entsteht aus A eine neue Matrix B. Wegen Satz 5.1.4 ist $\det(A) = D\det(B)$.

Ist $\rho(A) < n$, also A singulär, so enthält B eine Nullzeile. In diesem Fall ist wegen Satz 5.1.4 (4) die Determinante $\det(A) = 0$.

Ist A regulär, so entsteht keine Nullzeile. Daher kommen alle Zeilen als Pivotzeilen und alle Spalten als Pivotspalten vor. Wir führen in Gedanken — nicht rechnerisch — die Rücksubstitutionsschritte aus, mit denen die Nicht-Pivotelemente auf 0 reduziert werden. Diese ändern die Determinante nicht. Daher dürfen wir alle Nicht-Pivotelemente durch Null ersetzen. Anschließend können wir die Zeilen vertauschen, bis alle Pivotelemente auf der Diagonalen stehen. Es entsteht eine Diagonalmatrix C. Mit jeder Vertauschung ändern wir das Vorzeichen von D, daher gilt nun nach Satz 5.1.4 (3) die Determinante $\det(A) = D\det(C)$. Letztlich dividieren wir (in Gedanken) jede Zeile durch das Pivotlement, das in dieser Zeile steht, und multiplizieren dafür D mit dem Pivotelement. Es entsteht die Einheitsmatrix, und wegen Satz 5.1.4 und Definition 5.1.1 (3) ist $\det(A) = D\det(E_n) = D$. □

Beispiel 5.1.7. Berechnen Sie die Determinante

$$\begin{vmatrix} -8 & 3 & 2 & -3 \\ 1 & -5 & -6 & -21 \\ 13 & 6 & -10 & -27 \\ 2 & 4 & 2 & 9 \end{vmatrix}$$

Lösung: Pivotschritte, welche die Determinante nicht verändern:

$$\begin{matrix} -8 & 3 & 2 & -3 \\ 1 & -5 & -6 & -21 \\ 13 & 6 & -10 & -27 \\ 2 & 4 & 2* & 9 \end{matrix}$$

$$\begin{matrix} -10 & -1 & 0 & -12 \\ 7 & 7 & 0 & 6* \\ 23 & 26 & 0 & 18 \\ 2 & 4 & 2* & 9 \end{matrix}$$

$$\begin{matrix} 4 & 13 & 0 & 0 \\ 7 & 7 & 0 & 6* \\ 2* & 5 & 0 & 0 \\ 2 & 4 & 2* & 9 \end{matrix}$$

$$\begin{matrix} 0 & 3* & 0 & 0 \\ 7 & 7 & 0 & 6* \\ 2* & 5 & 0 & 0 \\ 2 & 4 & 2* & 9 \end{matrix}$$

Wir wissen jetzt: Die Matrix hat vollen Rang, ist also regulär. Ab jetzt spielen die Nicht-Pivotelemente keine Rolle mehr, wir interessieren uns nur mehr für die Pivotelemente. Wir vertauschen Zeilen so lange paarweise, bis die Pivotelemente auf der Diagonalen liegen.

$$\begin{matrix} 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{matrix}$$

$$\begin{matrix} 0 & 3 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 2 & 0 \end{matrix}$$

$$\begin{matrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 2 & 0 \end{matrix}$$

$$\begin{matrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{matrix}$$

3 Vertauschungen, eine ungerade Zahl, daher wird das Vorzeichen umgedreht. Also ist $\det(A) = -2 \cdot 3 \cdot 2 \cdot 6 = -72$. □

5.1.2. *Dreiecksmatrizen und Blockmatrizen.*

Definition 5.1.8. Sei $A = (\alpha_{i,j})_{i,j=1,\cdots,n}$ eine quadratische Matrix.

1) A heißt obere Dreiecksmatrix, wenn unter der Diagonalen nur Nullen stehen, d.h., falls $i > j$, dann ist $\alpha_{i,j} = 0$.
2) A heißt untere Dreiecksmatrix, wenn über der Diagonalen nur Nullen stehen, d.h., falls $i < j$, dann ist $\alpha_{i,j} = 0$.
3) A heißt Diagonalmatrix, wenn außerhalb der Diagonalen nur Nullen stehen, d.h. falls $i \neq j$, dann ist $\alpha_{i,j} = 0$.

Korollar 5.1.9. *Die Determinante einer Dreiecksmatrix ist das Produkt der Diagonalelemente.*

Beweis. Eine Dreiecksmatrix ist bereits in Zeilenstufenform, und die Pivotelemente sitzen bereits auf der Diagonalen. □

Definition 5.1.10. Sei $\mathbb{K}$ ein Körper, $m_1, m_2, n_1, n_2 \in \mathbb{N}$. Seien

$$\begin{aligned} A &= (\alpha_{i,j})_{i=1,\cdots,m_1,\, j=1,\cdots,n_1} \in \mathbb{K}^{m_1 \times n_1}, \\ B &= (\beta_{i,j})_{i=1,\cdots,m_1,\, j=1,\cdots,n_2} \in \mathbb{K}^{m_1 \times n_2}, \\ C &= (\gamma_{i,j})_{i=1,\cdots,m_2,\, j=1,\cdots,n_1} \in \mathbb{K}^{m_1 \times n_1}, \\ D &= (\delta_{i,j})_{i=1,\cdots,m_2,\, j=1,\cdots,n_2} \in \mathbb{K}^{m_2 \times n_2}. \end{aligned}$$

Wir definieren die Blockmatrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n_1} & \beta_{1,1} & \cdots & \beta_{1,n_2} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{m_1,1} & \cdots & \alpha_{m_1,n_1} & \beta_{m_1,1} & \cdots & \beta_{m_1,n_2} \\ \gamma_{1,1} & \cdots & \gamma_{1,n_1} & \delta_{1,1} & \cdots & \delta_{1,n_2} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{m_2,1} & \cdots & \gamma_{m_2,n_1} & \delta_{m_2,1} & \cdots & \delta_{m_2,n_2} \end{pmatrix}.$$

Ist $m_1 = n_1$ und $m_2 = n_2$ (also A und D quadratisch) und einer der Blöcke B oder C eine Nullmatrix, so heißt die Blockmatrix eine Blockdreiecksmatrix. Sind A und D quadratisch, und beide anderen Blöcke Nullmatrizen, so heißt die Blockmatrix eine Blockdiagonalmatrix.

Bemerkung 5.1.11. Definition 5.1.10 lässt sich analog auch für Matrizen mit mehr als 2×2 Blöcken verallgemeinern.

Satz 5.1.12. *Sei $\mathbb{K}$ ein Körper, seien $A \in \mathbb{K}^{m\times m}$, $D \in \mathbb{K}^{n\times n}$ quadratisch und $B \in \mathbb{K}^{m\times n}$, $C \in \mathbb{K}^{n\times m}$. Seien $\det_m, \det_n, \det_{m+n}$ Determinanten auf $\mathbb{K}^{m\times m}$, $\mathbb{K}^{n\times n}$ und $\mathbb{K}^{(m+n)\times(m+n)}$. (Wir werden die Indizes $m, n, m+n$ weglassen, aus dem Zusammenhang ist klar, welche der drei Determinanten jeweils gemeint ist.) Dann ist*

$$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = \det(A) \det(D).$$

Beweis. Der Kürze halber schreiben wir

$$T := \begin{pmatrix} A & B \\ 0 & D \end{pmatrix}.$$

Sei A singulär. Dann sind die Spalten von A, und folglich die ersten m Spalten von T, linear abhängig, und damit ist die Blockmatrix singulär. In diesem Fall ist $\det(T) = 0$ und auch $\det(A)\det(D) = 0 \cdot \det(D) = 0$.
Sei D singulär. Dann sind die Zeilen von D, und folglich die letzten n Zeilen von T,

linear abhängig, und daher ist T singulär. Wieder folgt $0 = \det(T) = \det(A)\det(D)$. Nun seien sowohl A als auch D regulär. Durch Pivotschritte und Zeilenvertauchungen lässt sich A in eine Diagonalmatrix A_1 umwandeln. Wir führen dieselben Schritte mit den ersten m Zeilen von T durch und erhalten eine Matrix

$$\tilde{T} = \begin{pmatrix} A_1 & B_1 \\ 0 & D \end{pmatrix}.$$

Auch D läßt sich durch Zeilentransformationen in eine Diagonalmatrix D_1 umwandeln. Dieselben Transformationen, angewendet auf die letzten n Zeilen von $\tilde{T}$, ergeben eine Matrix

$$T_1 = \begin{pmatrix} A_1 & B_1 \\ 0 & D_1 \end{pmatrix}.$$

Sei nun a das Produkt der Diagonalelemente von A_1 und d das Produkt der Diagonalelemente D_1, sei p die Anzahl der Zeilenvertauschungen unter den Transformationen, die A in A_1 umwandeln, und q die Anzahl der Zeilenvertauschungen bei der Umwandlung von D in D_1. Es ist dann ad das Produkt der Diagonalelemente von T_1 und $p+q$ die Anzahl der Zeilenvertauschungen bei der Umwandlung von T in T_1. Dann ist

$$\det(A) = (-1)^p a,\ \det(D) = (-1)^q d,\ \det(T) = (-1)^{p+q} ad = \det(A)\det(D).$$

Der Beweis für die andere Blockdreiecksmatrix geht ebenso. □

5.1.3. *Die Eindeutigkeit der Determinante.*

Satz 5.1.13. *Auf $\mathbb{K}^{n\times n}$ gibt es höchstens eine Determinante.*

Beweis. Wenn es überhaupt eine Determinante gibt, ist sie das eindeutige Ergebnis der Berechnung durch Algorithmus 5.1.6. □

5.2. **Multiplikationssatz für Determinanten.**

5.2.1. *Matrixmultiplikation und Determinante.*

Lemma 5.2.1. *Sei $\mathbb{K}$ ein Körper, und seien $T_{\text{tausch}}(i,j)$, $T_{\text{add}}(i,j,\lambda)$, $T_{\text{mult}}(i,\lambda)$ die $n\times n$-Elementarmatrizen aus Definition 4.2.25. Sei* det *eine Determinante auf $\mathbb{K}^{n\times n}$. Dann gilt für alle Matrizen $A \in \mathbb{K}^{n\times n}$:*

$$\begin{aligned} \det(T_{\text{tausch}}(i,j)A) &= -\det(A), \\ \det(T_{\text{add}}(i,j,\lambda)A) &= \det(A), \\ \det(T_{\text{mult}}(i,\lambda)A) &= \lambda\det(A). \end{aligned}$$

Insbesondere ist

$$\det(T_{\text{tausch}}(i,j)) = -1, \quad \det(T_{\text{add}}(i,j,\lambda)) = 1, \quad \det(T_{\text{mult}}(i,\lambda)) = \lambda.$$

Beweis. Die Linksmultiplikation mit Elementarmatrizen ist eine elementare Zeilenumformung. Daher ist das Lemma nichts anderes als eine Umformulierung von Satz 5.1.4. Die Determinanten der Elementarmatrizen ergeben sich, wenn man für A die Einheitsmatrix $A = E_n$ einsetzt. □

Lemma 5.2.2. *Jede reguläre Matrix in $\mathbb{K}^{n\times n}$ lässt sich als Produkt von Elementarmatrizen schreiben.*

Beweis. Ist A regulär, so haben wir in Algorithmus 4.3.10 gesehen, dass sich A durch elementare Umformungen auf die Einheitsmatrix reduzieren lässt, also $E_n = T_m T_{m-1} \cdots T_1 A$, d.h., $A = T_1^{-1} \cdots T_m^{-1} E_n$. Nach Lemma 3.3.13 sind auch die Inversen der Elementarmatrizen wieder Elementarmatrizen. □

Satz 5.2.3. *Sei* $\mathbb{K}$ *ein Körper und* det *eine Determinantenfunktion auf* $\mathbb{K}^{n\times n}$. *Seien* $A, B \in \mathbb{K}^{n\times n}$. *Dann ist*

$$\det(AB) = \det(A) \cdot \det(B).$$

Beweis. Ist A singulär, so ist auch AB singulär. Also gilt dann nach Satz 5.1.4 (4)

$$\det(A) = \det(AB) = 0.$$

Insbesondere ist also

$$\det(AB) = 0 = 0 \det(B) = \det(A) \det(B).$$

Ist A regulär, so lässt sich A nach Lemma 5.2.2 in der Form $A = T_1 \cdots T_m$ mit Elementarmatrizen T_k schreiben. Nach Lemma 5.2.1 gilt $\det(T_i C) = \det(T_i) \det(C)$ für jedes $C \in \mathbb{K}^{n\times n}$. Damit gilt (mittels vollständiger Induktion nach m):

$$\begin{aligned} \det(A) &= \det(T_1 \cdots T_m E_n) = \Pi_{i=1}^m \det(T_i) \cdot 1, \\ \det(AB) &= \det(T_1 \cdots T_m B) = \Pi_{i=1}^m \det(T_i) \cdot \det(B). \end{aligned}$$

Also ist $\det(AB) = \det(A) \det(B)$. □

Satz 5.2.4. *Sei* det *eine Determinante auf* $\mathbb{K}^{n\times n}$ *und* $A \in \mathbb{K}^{n\times n}$. *Dann hat die Transponierte* A^T *dieselbe Determinante wie* A*:*

$$\det(A^T) = \det(A).$$

Beweis. Ist A singulär, also $\rho(A) < n$, so ist auch $\rho(A^T) = \rho(A) < n$. Daher ist auch A^T singulär, und beide Determinanten sind Null: $\det(A) = \det(A^T) = 0$.

Ist A regulär, so ist nach Lemma 5.2.2 $A = T_1 \cdots T_m$ mit Elementarmatrizen T_i. Man sieht leicht, dass die Transponierten der Elementarmatrizen wieder Elementarmatrizen mit derselben Determinante sind. Also ist

$$\det(A^T) = \det(T_m^T \cdots T_1^T) = \Pi_{i=1}^m \det(T_i^T) = \Pi_{i=1}^m \det(T_i) = \det(A).$$

□

Korollar 5.2.5. *Algorithmus 5.1.6 zur Berechnung der Determinante funktioniert ebenso, wenn man statt Zeilenumformungen Spaltenumformungen verwendet.*

Beweis. Spaltenumformungen an A entsprechen Zeilenumformungen an A^T. □

5.2.2. *Determinante als Kriterium für Regularität.*

Satz 5.2.6. *Sei* det *eine Determinante auf* $\mathbb{K}^{n\times n}$ *und* $A \in \mathbb{K}^{n\times n}$. *Es gilt:*
A ist genau dann regulär, wenn $\det A \neq 0$. *In diesem Fall ist*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Beweis. Ist A singulär, so sind die Zeilen von A linear abhängig, folglich ist nach Satz 5.1.4 die Determinante $\det(A) = 0$. Ist A regulär, so ist nach Satz 5.2.3

$$1 = \det(E_n) = \det(AA^{-1}) = \det(A) \det(A^{-1}).$$

□

5.3. **Existenz der Determinante.**

5.3.1. *Permutationen.*

Definition 5.3.1. Sei $n \in \mathbb{N}$ und $\sigma : \{1, \cdots, n\} \to \{1, \cdots, n\}$ eine bijektive Abbildung. Dann heißt σ eine Permutation auf der Menge $\{1, \cdots, n\}$.
Mit $\mathcal{S}_n$ bezeichnen wir die Menge aller Permutationen auf $\{1, \cdots, n\}$.

Schreibweise 5.3.2. Eine Permutation $\sigma \in \mathcal{S}_n$ können wir anschreiben, in dem wir in der ersten Zeile die Zahlen $1, \cdots, n$, und darunter die Werte $\sigma(1), \cdots, \sigma(n)$ anschreiben. Es ist üblich, diesen Ausdruck in Klammern zusammenzufassen. Aus dem Zusammenhang müssen Sie erkennen, dass dies keine Matrix, sondern die Beschreibung einer Permutation ist.

Beispiel 5.3.3. Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} \in \mathcal{S}_4$$

bildet folgendermaßen ab:

$$\sigma(1) = 2, \quad \sigma(2) = 4, \quad \sigma(3) = 3, \quad \sigma(4) = 1.$$

Sei

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

Berechnen Sie die Hintereinanderausführungen $\sigma \circ \tau$ und $\tau \circ \sigma$, und σ^{-1}.

Lösung: Wir berechnen nacheinander $[\sigma \circ \tau](i)$ für $i = 1, 2, 3, 4$ und tragen die Werte in die Beschreibung von $\sigma \circ \tau$ ein. Es ist $\tau(1) = 4$, also $[\sigma \circ \tau](1) = \sigma(4) = 1$. Es ist $\tau(2) = 3$, daher $[\sigma \circ \tau](2) = \sigma(3) = 3$, u.s.w.:

$$\begin{aligned} \sigma \circ \tau &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}, \\ \tau \circ \sigma &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}. \end{aligned}$$

Zur Berechnung von σ^{-1} beachten wir zunächst: $\sigma(1) = 2$, daher ist $\sigma^{-1}(2) = 1$. Wir tragen in der Beschreibung von σ^{-1} den Wert 1 unter der Stelle 2 ein, u.s.w.:

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}.$$

□

Satz 5.3.4. *Sei $n \in \mathbb{N}$.*

1) *$\mathcal{S}_n$ bildet mit der Hintereinanderausführung eine Gruppe, die sogenannte symmetrische Gruppe. Dabei ist das neutrale Element die Identität* id*, das inverse Element zu $\sigma \in \mathcal{S}_n$ die Umkehrabbildung σ^{-1}.*
2) *Ist $n \geq 3$, dann ist $\mathcal{S}_n$ nicht kommutativ.*
3) *$\mathcal{S}_n$ enthält $n! := n \cdot (n-1) \cdots 2 \cdot 1$ Elemente. (Die Zahl $n!$ bezeichnet man als n factorielle.)*

Beweis. Punkt (1): Da die Hintereinanderausführung von bijektiven Abbildungen wieder bijektiv ist, ist $\circ$ eine innere Verknüpfung auf $\mathcal{S}_n$. Bekanntlich ist die Hintereinanderausführung von Funktionen assoziativ. Es ist $\mathrm{id} \circ \sigma = \sigma \circ \mathrm{id} = \sigma$, daher ist id das neutrale Element. Für die Umkehrabbildung σ^{-1} gilt $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \mathrm{id}$, also ist σ^{-1} das inverse Element zu σ.

Punkt (2) Es ist

$$\begin{pmatrix}1&2&3&4\cdots\\2&1&3&4\cdots\end{pmatrix}\circ\begin{pmatrix}1&2&3&4\cdots\\1&3&2&4\cdots\end{pmatrix} = \begin{pmatrix}1&2&3&4\cdots\\2&3&1&4\cdots\end{pmatrix},$$
$$\begin{pmatrix}1&2&3&4\cdots\\1&3&2&4\cdots\end{pmatrix}\circ\begin{pmatrix}1&2&3&4\cdots\\2&1&3&4\cdots\end{pmatrix} = \begin{pmatrix}1&2&3&4\cdots\\3&1&2&4\cdots\end{pmatrix}.$$

Punkt (3): Wenn wir eine Permutation σ auf $\{1,\cdots,n\}$ zusammenbauen, haben wir zunächst n Möglichkeiten, das Bild $\sigma(1)$ festzulegen. Nun ist einer der Werte aus $1\cdots n$ vergeben, und wir haben noch $n-1$ Möglichkeiten, das Bild $\sigma(2)$ festzulegen. Wir haben also bisher $n(n-1)$ verschiedene Möglichkeiten, $\sigma(1)$ und $\sigma(2)$ zu fixieren. Es bleiben $n-2$ Möglichkeiten für $\sigma(3)$, u.s.w. □

Definition 5.3.5.

1) Sei $\sigma\in\mathcal{S}_n$ eine Permutation. Ein Fehlstand von σ ist ein Paar (i,j) mit $i<j\in\{1,\cdots,n\}$ so, dass $\sigma(j)<\sigma(i)$.
2) Sei m die Anzahl der Fehlstände von σ. Dann ist das Vorzeichen oder Signum von σ
$$\operatorname{sgn}(\sigma) := (-1)^m.$$
3) Ist $\operatorname{sgn}(\sigma)=1$, so heißt σ eine gerade Permutation, sonst heißt σ eine ungerade Permutation.

Beispiel 5.3.6. Welche Fehlstände und welches Vorzeichen hat die Permutation

$$\tau=\begin{pmatrix}1&2&3&4&5&6\\5&3&1&6&4&2\end{pmatrix}?$$

Lösung: Fehlstände sind überall da zu finden, wo in der zweiten Zeile eine größere Zahl vor einer kleineren steht: Es steht:
5 vor den kleineren Zahlen 3, 1, 4, 2. Daher sind die Paare $(1,2)$, $(1,3)$, $(1,5)$ und $(1,6)$ Fehlstände. Das sind einmal 4 Fehlstände.
3 vor 1 und 2, das gibt 2 weitere Fehlstände.
1 vor keiner kleineren Zahl, kein weiterer Fehlstand.
6 vor 4 und 2, gibt 2 weitere Fehlstände.
4 vor 2, noch ein Fehlstand.
Insgesamt haben wir 9 Fehlstände, eine ungerade Zahl. Also ist $\operatorname{sgn}(\tau)=-1$, und τ ist eine ungerade Permutation. □

Definition 5.3.7. Sei $1\le i<j\le n$. Die Permutation

$$\tau(k) := \begin{cases} j & \text{wenn } k=i,\\ i & \text{wenn } k=j,\\ k & \text{wenn } k\notin\{i,j\},\end{cases}$$

heißt eine Vertauschung. Ist insbesondere $j=i+1$, so heißt τ eine Nachbarvertauschung oder Transposition.

Lemma 5.3.8. *Jede Vertauschung ist ungerade.*

Beweis. Wir suchen die Fehlstände in der Vertauschung

$$\tau=\begin{pmatrix}\cdots&(i-1)&i&(i+1)&\cdots&(j-1)&j&(j+1)&\cdots\\ \cdots&(i-1)&j&(i+1)&\cdots&(j-1)&i&(j+1)&\cdots\end{pmatrix}.$$

Es steht in der zweiten Zeile j vor den kleineren Zahlen $(i+1),\cdots,(j-1),i$, das gibt $j-i$ Fehlstände.
Es steht in der zweiten Zeile i nach den größeren Zahlen $j,(i+1),\cdots,(j-1)$. Das

wären wieder $j-i$ Fehlstände, aber der Fehlstand (i,j) wurde schon oben gezählt. τ hat daher insgesamt $2(j-i)-1$ Fehlstände und ist daher ungerade. □

Satz 5.3.9.

1) *Sei $\tau \in \mathcal{S}_n$ eine Transposition, und $\sigma \in \mathcal{S}_n$. Dann hat $\sigma \circ \tau$ entweder genau einen Fehlstand mehr oder einen Fehlstand weniger als σ.*
2) *Sei $\sigma \in \mathcal{S}_n$ eine Permutation mit m Fehlständen. Dann gibt es (nicht unbedingt eindeutig bestimmte) Transpositionen $\tau_1, \cdots, \tau_m$ so, dass*
$$\sigma = \tau_1 \circ \cdots \circ \tau_m \circ \mathrm{id}\,.$$
3) *Lässt sich eine Permutation $\sigma \in \mathcal{S}_n$ als Hintereinanderausführung von m Transpositionen schreiben, so ist m ungerade, falls σ ungerade ist, und gerade, falls σ gerade ist.*
4) *Sind $\sigma, \rho \in \mathcal{S}_n$, so ist*
$$\mathrm{sgn}(\sigma \circ \rho) = \mathrm{sgn}(\sigma) \cdot \mathrm{sgn}(\rho).$$

Beweis. Punkt (1): Sei τ die Nachbarvertauschung von i und $i+1$. Es ist
$$\sigma = \begin{pmatrix} 1 & \cdots & i & (i+1) & \cdots & n \\ \sigma(1) & \cdots & \sigma(i) & \sigma(i+1) & \cdots & \sigma(n) \end{pmatrix},$$
$$\sigma \circ \tau = \begin{pmatrix} 1 & \cdots & i & (i+1) & \cdots & n \\ \sigma(1) & \cdots & \sigma(i+1) & \sigma(i) & \cdots & \sigma(n) \end{pmatrix}.$$
Im Vergleich zu σ haben in der unteren Zeile nur die zwei benachbarten Elemente $\sigma(i)$ und $\sigma(i+1)$ Platz getauscht. Daher ändert sich in der Stellung der Elemente der unteren Zeile zueinander nur das Paar $(\sigma(i), \sigma(i+1))$. War $(i, i+1)$ kein Fehlstand in σ, so ist es nun in $\sigma\circ\tau$ ein Fehlstand, und umgekehrt. Es gibt also einen Fehlstand mehr, oder einen Fehlstand weniger.

Punkt (2): Vollständige Induktion nach der Anzahl der Fehlstände m. Ist $m = 0$, so ist σ die Identität, und es braucht keine Transpositionen. Wir nehmen nun an, dass die Behauptung für Permutationen mit $m-1$ Fehlständen gilt. Sei σ eine Permutation mit $m > 0$ Fehlständen. Dann gibt es ein i sodass $\sigma(i) > \sigma(i+1)$. Sonst müsste für alle $i < j$ gelten: $\sigma(i) < \sigma(i+1) < \cdots < \sigma(j)$, und es könnte gar keinen Fehlstand geben. Sei nun τ_m die Vertauschung des Paares $(i, i+1)$. Dann hat $\sigma \circ \tau_m$ einen Fehlstand weniger als σ, denn nun ist das Paar $(i, i+1)$ kein Fehlstand mehr. Nach Induktionsannahme gibt es Transpositionen $\tau_1, \cdots, \tau_{m-1}$ so, dass
$$\sigma \circ \tau_m = \tau_1 \circ \cdots \circ \tau_{m-1}.$$
Beachten wir noch, dass für jede Transposition τ gilt $\tau \circ \tau = \mathrm{id}$, dann ist
$$\sigma = \sigma \circ \tau_m \circ \tau_m = \tau_1 \circ \cdots \circ \tau_m.$$

Punkt (3): Wegen Punkt (1) ist für jede Permutation σ und jede Transposition τ
$$\mathrm{sgn}(\sigma \circ \tau) = -\mathrm{sgn}(\sigma).$$
Daher ist
$$\mathrm{sgn}(\tau_1 \circ \cdots \circ \tau_m) = (-1)^m.$$

Punkt (4): Wegen Punkt (2) lassen sich σ und ρ als Hintereinanderausführung von Transpositionen schreiben:
$$\sigma = \tau_1 \circ \cdots \circ \tau_m,\ \rho = \tilde{\tau}_1 \circ \cdots \circ \tilde{\tau}_k.$$
Dann ist
$$\sigma \circ \rho = \tau_1 \circ \cdots \circ \tau_m \circ \tilde{\tau}_1 \circ \cdots \circ \tilde{\tau}_k$$

eine Hintereinanderausführung von $m+k$ Transpositionen. Nach Punkt (3) ist

$$\operatorname{sgn}(\sigma) = (-1)^m, \quad \operatorname{sgn}(\rho) = (-1)^k, \quad \operatorname{sgn}(\sigma\circ\rho) = (-1)^{k+m} = (-1)^m \cdot (-1)^k.$$

□

5.3.2. *Die Formel von Leibniz.*

Satz 5.3.10. *Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Die folgende Funktion* $\det : \mathbb{K}^{n\times n} \to \mathbb{K}$ *ist eine Determinante.*

$$\det \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} := \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\, \alpha_{1,\sigma(1)}\alpha_{2,\sigma(2)}\cdots\alpha_{n,\sigma(n)} \ .$$

Beweis. Linearität in den Zeilen:

$$\begin{aligned}
&\det \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \lambda\alpha_{i,1}+\mu\beta_{i,1} & \cdots & \lambda\alpha_{i,n}+\mu\beta_{i,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} \\
= &\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\, \alpha_{1,\sigma(1)}\cdots(\lambda\alpha_{i,\sigma(i)}+\mu\beta_{i,\sigma(i)})\cdots\alpha_{n,\sigma(n)} \\
= &\lambda \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\, \alpha_{1,\sigma(1)}\cdots\alpha_{i,\sigma(i)}\cdots\alpha_{n,\sigma(n)} + \mu \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\, \alpha_{1,\sigma(1)}\cdots\beta_{i,\sigma(i)}\cdots\alpha_{n,\sigma(n)} \\
= &\lambda \det \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{i,1} & \cdots & \alpha_{i,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} + \mu \det \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \beta_{i,1} & \cdots & \beta_{i,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix}.
\end{aligned}$$

Matrizen mit zwei gleichen Zeilen haben Determinante 0: Wir zeigen zuerst: Bei Vertauschung zweier Zeilen ändert sich das Vorzeichen der Determinante. Sei τ die Vertauschung

$$\tau = \begin{pmatrix} 1 & \cdots & i & \cdots & j & \cdots & n \\ 1 & \cdots & j & \cdots & i & \cdots & n \end{pmatrix}$$

Nach Lemma 5.3.8 ist $\operatorname{sgn}(\tau) = -1$. Nach Satz 5.3.9 gilt daher $\operatorname{sgn}(\sigma\circ\tau) = -\operatorname{sgn}(\sigma)$. Damit gilt für die Matrix mit vertauschten Zeilen i, j:

$$\begin{aligned}
&\det \begin{pmatrix} \vdots & & \vdots \\ \alpha_{j,1} & \cdots & \alpha_{j,n} \\ \vdots & & \vdots \\ \alpha_{i,1} & \cdots & \alpha_{i,n} \\ \vdots & & \vdots \end{pmatrix} \\
= &\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\, \alpha_{1,\sigma(1)}\cdots\alpha_{j,\sigma(i)}\cdots\alpha_{i,\sigma(j)}\cdots\alpha_{n,\sigma(n)} \\
= &\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\, \alpha_{1,\sigma\circ\tau(1)}\cdots\alpha_{j,\sigma\circ\tau(j)}\cdots\alpha_{i,\sigma\circ\tau(i)}\cdots\alpha_{n,\sigma\circ\tau(n)}
\end{aligned}$$

$$
\begin{aligned}
&= -\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma\circ\tau)\,\alpha_{1,\sigma\circ\tau(1)}\cdots\alpha_{i,\sigma\circ\tau(i)}\cdots\alpha_{j,\sigma\circ\tau(j)}\cdots\alpha_{n,\sigma\circ\tau(n)}\\
&= -\sum_{\rho\in\mathcal{S}_n} \operatorname{sgn}(\rho)\,\alpha_{1,\rho(1)}\cdots\alpha_{i,\rho(i)}\cdots\alpha_{j,\rho(j)}\cdots\alpha_{n,\rho(n)}\\
&= -\det\begin{pmatrix} \vdots & & \vdots \\ \alpha_{i,1} & \cdots & \alpha_{i,n} \\ \vdots & & \vdots \\ \alpha_{j,1} & \cdots & \alpha_{j,n} \\ \vdots & & \vdots \end{pmatrix}.
\end{aligned}
$$

Im letzten Schritt haben wir ausgenützt, dass wenn σ alle Permutationen durchläuft, dann auch $\rho := \sigma\circ\tau$ alle Permutationen durchläuft.
Nun sei A eine Matrix, deren Zeilen Nummer i und Nummer j gleich sind. Wenn man die Zeilen i und j vertauscht, entsteht wieder A. Daher gilt $\det(A) = -\det(A)$, und folglich ist $\det(A) = 0$.
Determinante der Einheitsmatrix ist Eins: Die Koeffizienten der Einheitsmatrix sind $\delta_{i,j}$, dabei bedeutet δ das Kronecker-Delta:

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

Nun gibt es für jede Permutation $\sigma \neq \mathrm{id}$ einen Index i mit $\sigma(i) \neq i$. Damit ist

$$\operatorname{sgn}(\sigma)\,\delta_{1,\sigma(1)}\cdots\delta_{i,\sigma(i)}\cdots\delta_{n,\sigma(n)} = 0.$$

Nur für die Indentität ist

$$\operatorname{sgn}(id)\,\delta_{1,1}\cdots\delta_{n,n} = 1.$$

Daher ist

$$\det(E_n) = \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\,\delta_{1,\sigma(1)}\cdots\delta_{i,\sigma(i)}\cdots\delta_{n,\sigma(n)} = 1.$$

□

Satz 5.3.11. *Es gibt genau eine Determinante* $\mathbb{K}^{n\times n} \to \mathbb{K}$.

Beweis. Durch die Leibnizsche Formel wird eine Determinante definiert, daher gibt es mindestens eine Determinante. Wenn det eine Determinante ist, lässt sich sich mit Algorithmus 5.1.6 berechnen. Daher kann es höchstens eine Determinante geben. □

5.4. **Entwicklungssatz von Laplace.**

5.4.1. *Entwicklungssatz von Laplace.*

Definition 5.4.1. Sei $\mathbb{K}$ ein Körper, $A = (\alpha_{i,j}) \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. Für jedes Indexpaar (i, j) entsteht die Matrix $A'_{i,j}$ durch Streichen der i-ten Zeile und der j-ten Spalte:

$$A'_{i,j} := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,j-1} & \alpha_{1,j+1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{i-1,1} & \cdots & \alpha_{i-1,j-1} & \alpha_{i-1,j+1} & \cdots & \alpha_{i-1,n} \\ \alpha_{i+1,1} & \cdots & \alpha_{i+1,j-1} & \alpha_{i+1,j+1} & \cdots & \alpha_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,j-1} & \alpha_{n,j+1} & \cdots & \alpha_{n,n} \end{pmatrix}$$

Wir definieren die komplementäre Matrix $A^\sharp \in \mathbb{K}^{n\times n}$ zu A:

$$\begin{aligned} A^\sharp &= \begin{pmatrix} \alpha^\sharp_{1,1} & \cdots & \alpha^\sharp_{1,n} \\ \vdots & & \vdots \\ \alpha^\sharp_{n,1} & \cdots & \alpha^\sharp_{n,n} \end{pmatrix} \quad \text{mit} \\ \alpha^\sharp_{i,j} &:= (-1)^{i+j} \det(A'_{j,i}). \end{aligned}$$

(Achtung auf die Umkehrung der Reihenfolge der Indizes in der obigen Formel bei $\det(A'_{j,i})$!)

Bemerkung 5.4.2 (Schachbrettregel)**.** Die Vorzeichen $(-1)^{i+j}$ liegen wie ein Schachbrett auf der Matrix. Auf der Diagonalen liegt immer Plus:

$$\begin{pmatrix} + & - & + & \cdots & \pm \\ - & + & - & \cdots & \mp \\ + & - & + & \cdots & \pm \\ \vdots & \vdots & \vdots & & \vdots \\ \pm & \mp & \pm & \cdots & + \end{pmatrix}.$$

Satz 5.4.3. *Sei $\mathbb{K}$ ein Körper und $A = (\alpha_{i,j}) \in \mathbb{K}^{n\times n}$.*

1) *("Entwicklung nach einer Zeile":) Sei $i \in \{1, \cdots, n\}$. Dann ist*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \alpha_{i,j} \det(A'_{i,j}).$$

2) *("Entwicklung nach einer Spalte":) Sei $j \in \{1, \cdots, n\}$. Dann ist*

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \alpha_{i,j} \det(A'_{i,j}).$$

Beweis. Wir zeigen die Entwicklung nach einer Zeile. Die Entwicklung nach einer Spalte folgt dann aus $\det(A) = \det(A^T)$.
Sei zunächst τ_i die Permutation

$$\tau_i = \begin{pmatrix} 1 & 2 & \cdots & i-1 & i & i+1 & \cdots & n-1 & n \\ i & 1 & \cdots & i-2 & i-1 & i+1 & \cdots & n-1 & n \end{pmatrix}$$

Offensichtlich sind die Fehlstände von τ_i die Paare $(1,k)$ mit $k = 2, \cdots, i$. Das sind $i-1$ Fehlstände. Daher ist

$$\operatorname{sgn}(\tau_i) = (-1)^{i-1}.$$

Wir betrachten nun die Matrix

$$B_{i,j} := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,j-1} & \alpha_{1,j} & \alpha_{1,j+1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{i-1,1} & \cdots & \alpha_{i-1,j-1} & \alpha_{i-1,j} & \alpha_{i-1,j+1} & \cdots & \alpha_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \alpha_{i+1,1} & \cdots & \alpha_{i+1,j-1} & \alpha_{i+1,j} & \alpha_{i+1,j+1} & \cdots & \alpha_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,j-1} & \alpha_{n,j} & \alpha_{n,j+1} & \cdots & \alpha_{n,n} \end{pmatrix}.$$

Wir tauschen die j-te Spalte ans linke Ende, und die i-te Zeile ans obere Ende der Matrix. Es ergibt sich eine Blockdreiecksmatrix, wenn man die erste Zeile und

erste Spalte als einen 1×1-Block, und die anderen Zeilen und Spalten als einen $(n-1) \times (n-1)$-Block ansieht:

$$\det(B_{i,j}) = \operatorname{sgn}(\tau_j)\operatorname{sgn}(\tau_i) \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \alpha_{1,j} & \alpha_{1,1} & \cdots & \alpha_{1,j-1} & \alpha_{1,j+1} & \cdots & \alpha_{1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{i-1,j} & \alpha_{i-1,1} & \cdots & \alpha_{i-1,j-1} & \alpha_{i-1,j+1} & \cdots & \alpha_{i-1,n} \\ \alpha_{i+1,j} & \alpha_{i+1,1} & \cdots & \alpha_{i+1,j-1} & \alpha_{i+1,j+1} & \cdots & \alpha_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{n,j} & \alpha_{n,1} & \cdots & \alpha_{n,j-1} & \alpha_{n,j+1} & \cdots & \alpha_{n,n} \end{vmatrix}.$$

Weil $\operatorname{sgn}(\tau_i)\operatorname{sgn}(\tau_j) = (-1)^{i-1}(-1)^{j-1} = (-1)^{i+j}$, und nach Satz 5.1.12, ergibt sich nun

$$\det(B_{i,j}) = (-1)^{i+j} \det(A'_{i,j}).$$

Nun sind wir in der Lage, den Entwicklungssatz fertig zu beweisen. Sind $e_1^T, \cdots, e_n^T$ die Einheits-Zeilenvektoren, so lässt sich die i-te Zeile von A schreiben:

$$\alpha_{i,1}e_1^T + \cdots + \alpha_{i,n}e_n^T.$$

Wegen der Linearität in der i-ten Zeile ist

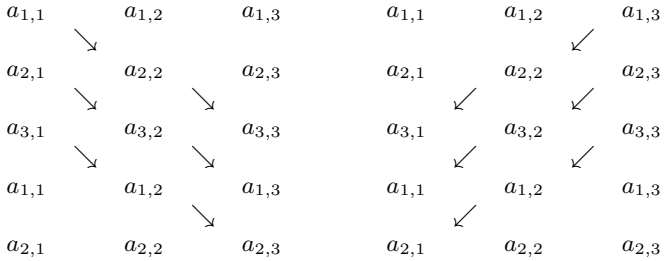
$$\det(A) = \sum_{j=1}^{n} \alpha_{i,j} \det(B_{i,j}) = \sum_{j=1}^{n} (-1)^{i+j} \alpha_{i,j} \det(A'_{i,j}).$$

□

Korollar 5.4.4. *Für 2×2-Matrizen gilt:*

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

Korollar 5.4.5 (Regel von Sarrus)**.** *Um die Determinante einer 3×3 Matrix zu berechnen, schreibt man die ersten zwei Zeilen unter der Matrix noch einmal an. Man addiert dann die drei Produkte, die sich von links oben nach rechts unten ergeben, und subtrahiert die drei Produkte, die sich von rechts oben nach links unten ergeben:*



$$[a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3}] \quad - \quad [a_{1,3}a_{2,2}a_{3,1} + a_{2,3}a_{3,2}a_{1,1} + a_{3,3}a_{1,2}a_{2,1}].$$

Beispiel 5.4.6. Der Laplacesche Entwicklungssatz eignet sich für Determinanten in Matrizen, die viele Nullen enthalten. Zum Beispiel berechnen wir die Determinante

$$\begin{vmatrix} \lambda - 4 & 0 & 1 & 2 \\ -2 & \lambda + 3 & 3 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 1 & 1 & \lambda - 2 \end{vmatrix}.$$

Lösung: Wir entwickeln zunächst nach der dritten Zeile, anschließend die beiden 3×3-Determinanten nach der ersten Spalte. Die 2×2-Determinanten lösen wir mittels Korollar 5.4.4 auf.

$$\begin{vmatrix} \lambda-4 & 0 & 1 & 2 \\ -2 & \lambda+3 & 3 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 1 & 1 & \lambda-2 \end{vmatrix} = \lambda \begin{vmatrix} \lambda-4 & 0 & 2 \\ -2 & \lambda+3 & 0 \\ 0 & 1 & \lambda-2 \end{vmatrix} - 1 \begin{vmatrix} \lambda-4 & 0 & 1 \\ -2 & \lambda+3 & 3 \\ 0 & 1 & 1 \end{vmatrix}$$

$$\begin{aligned} &= \lambda(\lambda-4) \begin{vmatrix} \lambda+3 & 0 \\ 1 & \lambda-2 \end{vmatrix} + 2\lambda \begin{vmatrix} 0 & 2 \\ 1 & \lambda-2 \end{vmatrix} - (\lambda-4) \begin{vmatrix} \lambda+3 & 3 \\ 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} \\ &= \lambda(\lambda-4)(\lambda+3)(\lambda-2) - 4\lambda - (\lambda-4)\lambda + 2 \\ &= \lambda^4 - 3\lambda^3 - 11\lambda^2 + 24\lambda + 2. \end{aligned}$$

□

5.4.2. *Die Komplementärmatrix und die Inverse.*

Satz 5.4.7. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{n\times n}$ und $A^\sharp$ die Komplementärmatrix zu A. Dann ist*

$$AA^\sharp = A^\sharp A = \det(A)\, E_n.$$

Insbesondere gilt für jede reguläre Matrix A

$$A^{-1} = \frac{1}{\det(A)}\, A^\sharp.$$

Beweis. Wir zeigen $AA^\sharp = \det(A)\, E_n$. Das Produkt $A^\sharp A = \det(A)\, E_n$ geht genauso.
Sei also $A = (\alpha_{i,j})$, $A^\sharp = (\alpha^\sharp_{i,j})$ und $AA^\sharp = (\gamma_{i,j})$. Wir bilden eine Matrix $\tilde{A}$, indem wir die j-te Zeile von A durch die i-te Zeile ersetzen. Für $i = j$ bleibt $\tilde{A} = A$, für $i \neq j$ kommt die i-te Zeile von A in der Matrix $\tilde{A}$ zweimal vor. Daher ist

$$\det(\tilde{A}) = \begin{cases} \det(A) & \text{wenn } i = j \\ 0 & \text{wenn } i \neq j \end{cases} = \delta_{i,j}\, \det(A)$$

(mit dem Kronecker-δ). Wir entwickeln die Determinante von $\tilde{A}$ nach der j-ten Zeile:

$$\delta_{i,j}\, \det(A) = \det(\tilde{A}) = \begin{vmatrix} \text{(Zeile 1:)} & \alpha_{1,1} & \cdots & \alpha_{1,n} \\ & \vdots & & \vdots \\ \text{(Zeile } j\text{:)} & \alpha_{i,1} & \cdots & \alpha_{i,n} \\ & \vdots & & \vdots \\ \text{(Zeile } n\text{:)} & \alpha_{n,1} & \cdots & \alpha_{n,n} \end{vmatrix}$$

$$= \sum_{k=1}^{n} (-1)^{j+k}\, \alpha_{i,k}\, \det(A'_{j,k}) = \sum_{k=1}^{n} \alpha_{i,k}\alpha^\sharp_{k,j} = \gamma_{i,j}\,.$$

Damit ist der Satz bewiesen. □

Bemerkung 5.4.8. Besonders einfach ist dieser Satz zur Invertierung von 2×2-Matrizen anzuwenden:

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

5.4.3. *Cramersche Regel.*

Satz 5.4.9 (Cramersche Regel)**.** *Sei $A \in \mathbb{K}^{n\times n}$ eine reguläre Matrix mit den Spalten $a_1, \cdots, a_n$. Sei $b \in \mathbb{K}^n$ und $x = (\xi_1, \cdots, \xi_n)^T$ die eindeutige Lösung des Gleichungssystems $Ax = b$. Dann gilt*

$$\xi_i = \frac{1}{\det(A)} \begin{vmatrix} a_1 & \cdots & a_{i-1} & b & a_{i+1} & \cdots & a_n \end{vmatrix}$$

Beweis. Wir setzen

$$B_i := \begin{pmatrix} a_1 & \cdots & a_{i-1} & a_i & a_{i+1} & \cdots & a_n & b \\ 0 & & 0 & 1 & 0 & & 0 & \xi_i \end{pmatrix}$$

Dann ist

$$B_i \begin{pmatrix} x \\ -1 \end{pmatrix} = 0.$$

Daher ist B_i singulär, und $\det(B_i) = 0$. Wir entwickeln nach der letzten Zeile und vertauschen am Ende die Spalten:

$$\begin{aligned} 0 &= (-1)^{n+1+i} \det \begin{pmatrix} a_1 & \cdots & a_{i-1} & a_{i+1} & \cdots & a_n & b \end{pmatrix} \\ &\quad + \xi_i \det \begin{pmatrix} a_1 & \cdots & a_{i-1} & a_i & a_{i+1} & \cdots & a_n \end{pmatrix} \\ &= -\det \begin{pmatrix} a_1 & \cdots & a_{i-1} & b & a_{i+1} & \cdots & a_n \end{pmatrix} \\ &\quad + \xi_i \det \begin{pmatrix} a_1 & \cdots & a_{i-1} & a_i & a_{i+1} & \cdots & a_n \end{pmatrix} \end{aligned}$$

Zur Bestimmung des Vorzeichens im letzten Schritt haben wir verwendet, dass die Spalte b, um an die Stelle i zu gelangen, die Spalten $a_n, \cdots, a_{i+1}$ überspringen muss. Das sind $(n - i)$ Transpositionen. □

Die Cramersche Regel eignet sich gut zur Lösung von Systemen von 2 Gleichungen in 2 Unbekannten. Sonst ist die Bestimmung jeder einzelnen Determinante so aufwendig wie die Lösung des ganzen Systems, und wenn die Systemmatrix nicht regulär ist, scheitert die Cramersche Regel, während Pivottransformationen die Lösungsmenge ergeben. Die Cramersche Regel ist aber von großem Wert für die Theorie, weil sie eine geschlossene Formel liefert, aus der sich Eigenschaften der Lösung ableiten lassen.

Beispiel 5.4.10. Wir lösen mit der Cramerschen Regel

$$\begin{array}{rcrcl} 3x & + & 5y & = & 7 \\ 2x & - & 7y & = & 4 \end{array}$$

Lösung:

$$\begin{aligned} x &= \frac{\left|\begin{smallmatrix} 7 & 5 \\ 4 & -7 \end{smallmatrix}\right|}{\left|\begin{smallmatrix} 3 & 5 \\ 2 & -7 \end{smallmatrix}\right|} = \frac{-69}{-31} = \frac{69}{31}, \\ y &= \frac{\left|\begin{smallmatrix} 3 & 7 \\ 2 & 4 \end{smallmatrix}\right|}{\left|\begin{smallmatrix} 3 & 5 \\ 2 & -7 \end{smallmatrix}\right|} = \frac{-2}{-31} = \frac{2}{31}. \end{aligned}$$

6. Vektorräume mit innerem Produkt

6.1. **Der euklidische Raum.**

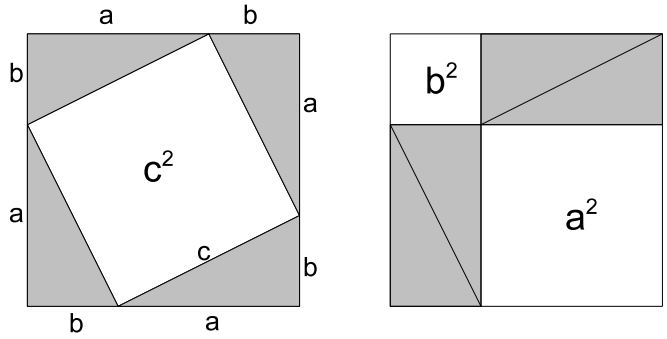
6.1.1. *Etwas Elementargeometrie.*

In diesem Abschnitt, der nur dazu dient, Sie an einige wichtige Fakten der Elementargeometrie zu erinnern, gehen wir nicht streng axiomatisch vor, sondern motivieren die Resultate und Begriffe nur anschaulich. Eine sorgfältige Definition der Begriffe und eine strenge Beweisführung würde viel mehr Arbeit benötigen.

Satz 6.1.1 (Satz von Pythagoras)**.** *Gegeben sei ein rechtwinkeliges Dreieck mit Seitenlängen a, b für die Katheten und c für die Hypotenuse. Dann gilt*

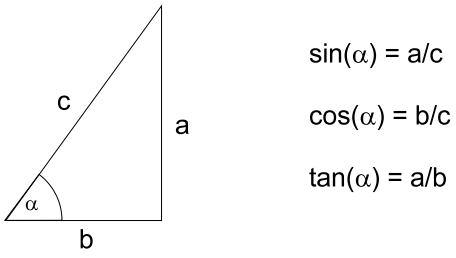
$$a^2 + b^2 = c^2.$$

Beweisidee: In einem Quadrat der Seitenlänge $a + b$ decken wir einen Teil der Fläche durch 4 kongruente rechtwinkelige Dreiecke mit Seitenlängen a, b, c zu, siehe untenstehende Grafik. Je nach Anordnung der Dreiecke ergibt sich für den nicht abgedeckten Teil entweder ein Quadrat der Seitenlänge c oder zwei Quadrate mit Seitenlängen a und b. Daher ist die nicht abgedeckte Fläche $a^2 + b^2 = c^2$.



□

Definition 6.1.2. Gegeben sei ein rechtwinkeliges Dreieck mit den Seitenlängen a und b für die Katheten und c für die Hypotenuse. Der Kathete a gegenüber liege der Winkel α. Wir definieren die Winkelfunktionen:



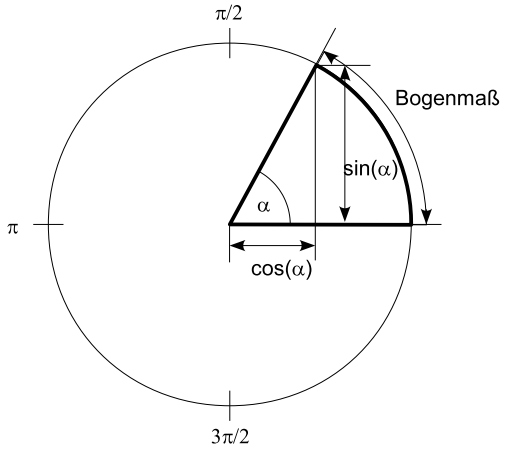
Eine axiomatisch hieb- und stichfeste Definition würde benötigen, dass diese Winkelfunktionen wirklich nur vom Winkel α und nicht von weiteren Eigenschaften des gewählten Dreiecks abhängt.

Bemerkung 6.1.3. Unter dem Einheitskreis in $\mathbb{R}^2$ versteht man die Kreislinie mit Mittelpunkt 0 und Radius 1. Ein Strahl im Einheitskreis, der mit dem Vektor $e_1 = (0,1)^T$ einen Winkel von α einschließt, symbolisiert den Winkel α. Positive Winkel werden linksdrehend, negative rechtsdrehend eingetragen. Trifft der Strahl den Einheitskreis im Punkt $x = (\xi, \eta)^T$, so sieht man folgendes:

1) Das Bogenmaß von α ist die Länge des Bogens von e_1 nach x auf dem Einheitskreis. (Negativ, falls der Bogen rechtsdrehend durchlaufen wird.) Insbesondere haben die Winkel 90°, 180°, 270° und 360° die Bogenmaße $\pi/2$, π, $3\pi/2$ und 2π. Winkel, die sich im Bogenmaß um geradzahlige Vielfache von π unterscheiden, gelten als gleich. Die Umrechnungsformel lautet

$$\text{Bogenmaß} = \text{Gradmaß} \cdot \frac{\pi}{180}.$$

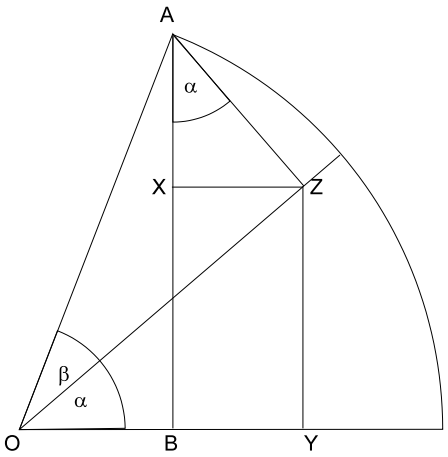
2) Die Hypotenuse des rechtwinkeligen Dreiecks mit den Ecken 0, $(\xi, 0)$ und $(\xi, \eta)^T$ hat die Länge 1. Daher sieht man: $\xi = \cos(\alpha)$, $\eta = \sin(\alpha)$. Die Vorzeichen der Winkelfunktionen hängen davon ab, in welchem Quadranten der Winkel liegt, und lassen sich in der Einheitskreis-Darstellung leicht erkennen.
3) Da in diesem rechtwinkeligen Dreieck mit den Ecken der Satz von Pythagoras gilt, folgt die Formel $\cos^2(\alpha) + \sin^2(\alpha) = 1$.



Satz 6.1.4 (Additionstheoreme der Winkelfunktionen)**.** *Seien $\alpha, \beta \in \mathbb{R}$. Für den Summenwinkel $\alpha + \beta$ gilt*

$$\begin{aligned} \sin(\alpha+\beta) &= \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta), \\ \cos(\alpha+\beta) &= \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta). \end{aligned}$$

Beweisidee: Die folgende Graphik zeigt den Einheitskreis und eingezeichnet die beiden Winkel α und β. Sinus und Cosinus des Summenwinkels lassen sich aus dieser Grafik als Strecken ablesen.



Es sind

$$\begin{aligned}\sin(\alpha+\beta) &= \overline{AB},\\ \cos(\alpha+\beta) &= \overline{OB}.\end{aligned}$$

Auflösung des Dreiecks OAZ liefert

$$\begin{aligned}\overline{AZ} &= \sin(\beta),\\ \overline{OZ} &= \cos(\beta).\end{aligned}$$

Es folgt nun

$$\begin{aligned}\overline{AX} &= \overline{AZ}\cos(\alpha) = \sin(\beta)\cos(\alpha),\\ \overline{BX} &= \overline{OZ}\sin(\alpha) = \sin(\alpha)\cos(\beta),\\ \sin(\alpha+\beta) &= \overline{AX}+\overline{BX} = \sin(\beta)\cos(\alpha)+\sin(\alpha)\cos(\beta).\end{aligned}$$

Ebenso ist

$$\begin{aligned}\overline{BY} &= \overline{AZ}\sin(\alpha) = \sin(\beta)\sin(\alpha),\\ \overline{OY} &= \overline{OZ}\cos(\alpha) = \cos(\alpha)\cos(\beta),\\ \sin(\alpha+\beta) &= \overline{OY}-\overline{BY} = \cos(\beta)\cos(\alpha)-\sin(\alpha)\sin(\beta).\end{aligned}$$

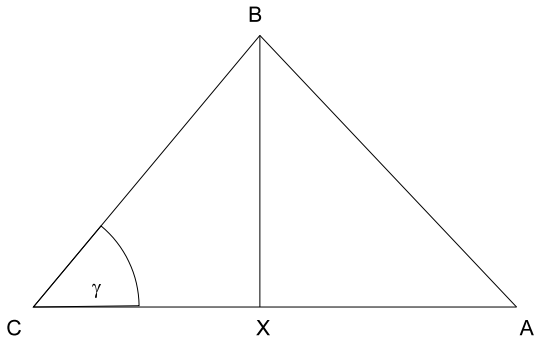
□

Satz 6.1.5 (Cosinussatz)**.** *Gegeben sei ein Dreieck mit den Seitenlängen a, b, c. Der Seite mit Länge c gegenüber liege der Winkel γ. Dann gilt*

$$c^2 = a^2 + b^2 - 2ab\cos(\gamma).$$

Beweisidee Die folgende Grafik zeigt das Dreieck ABC. Dabei sind die Seitenlängen die folgenden Strecken: $a = \overline{BC}$, $b = \overline{AC}$, $c = \overline{AB}$. Durch den Punkt B wurde eine

Hilfslinie gezogen, welche auf die Seite $\overline{CA}$ orthogonal steht.



Auflösung des Dreiecks ABX liefert

$$\begin{aligned} \overline{CX} &= \overline{CB}\cos\gamma, \\ \overline{BX} &= \overline{CB}\sin\gamma. \end{aligned}$$

Nach dem Satz von Pythagoras ergibt sich

$$\begin{aligned} \overline{AB}^{\,2} &= \overline{AX}^{\,2} + \overline{XB}^{\,2} \\ &= (\overline{CA} - \overline{CX})^2 + \overline{BX}^{\,2} \\ &= (\overline{CA} - \overline{CB}\cos(\gamma))^2 + \overline{CB}^{\,2}\sin^2(\gamma) \\ &= \overline{CB}^{\,2}(\sin^2(\gamma) + \cos^2(\gamma)) - 2\,\overline{CA}\,\overline{CB}\cos(\gamma) + \overline{CA}^2 \\ &= \overline{CB}^{\,2} + \overline{CA}^{\,2} - 2\,\overline{CA}\,\overline{CB}\cos(\gamma). \end{aligned}$$

□

6.1.2. *Euklidische Norm und inneres Produkt in $\mathbb{R}^n$.*

Beachten Sie: Wir haben nie definiert, was eine Länge ist, auch nie, was ein Winkel oder ein rechter Winkel ist. Daher steht (für uns) alles, was wir im vorigen Unterabschnitt entwickelt haben, auf keiner axiomatischen Basis. Wir wollen nun geeignete Begriffe definieren, mit denen Längen, Orthogonalität und Winkel erfasst werden können, ohne die ganze Elementargeometrie axiomatisch entwickeln zu müssen. Dazu zäumen wir das Pferd vom Schwanz her auf: Wie verwenden unsere Kenntnisse von Pythagoras und Winkelfunktionen als Motivation für unsere Definitionen.

Unsere erste Definition schafft einen Begriff für die Länge eines Vektors:

Definition 6.1.6. Sei $a = (\alpha_1, \cdots, \alpha_n)^T \in \mathbb{R}^n$. Die euklidische Norm von a ist definiert durch

$$\|a\|_2 := \sqrt{\alpha_1^2 + \cdots + \alpha_n^2}.$$

Schreibweise 6.1.7. Üblicherweise schreibt man einfach $\|a\|$ ohne Subskript für die euklidische Norm, wenn keine Verwechslungen zu erwarten sind. Wir können aber auch andere Normen auf $\mathbb{R}^n$ definieren, nur haben diese keine so anschauliche Interpretation als Länge. Was man unter einer Norm im Allgemeinen versteht, werden wir erst in Definition 6.2.1 festlegen.
Manchmal werden die Normen auf $\mathbb{R}^n$ auch mit einfachen Strichen geschrieben: $|a|$. Statt des Wortes Norm wird auch oft das Wort Betrag oder Absolutbetrag verwendet.

Satz 6.1.8. *Die euklidische Norm hat folgende Eigenschaften:*

1) *Für alle $x \in \mathbb{R}^n$ ist $\|x\|_2 \geq 0$. Der Vektor x ist genau dann der Nullvektor, wenn $\|x\|_2 = 0$.*
2) *Für alle $x \in \mathbb{R}^n$ und alle $\lambda \in \mathbb{R}$ gilt $\|\lambda x\|_2 = |\lambda| \ \|x\|_2$.*
3) *Für alle $x, y \in \mathbb{R}^n$ gilt: $\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$.*

Beweis. Die meisten der Eigenschaften oben lassen sich leicht nachrechnen. Nur die sogenannte Dreiecksungleichung (3) ist etwas schwieriger. Wir beweisen sie in abstrakterem Rahmen im nächsten Abschnitt, siehe Satz 6.2.15. □

Bemerkung 6.1.9. Ist $x \in \mathbb{R}^n \setminus \{0\}$ ein Vektor, so hat der Vektor

$$y = \frac{1}{\|x\|_2} \, x$$

dieselbe Richtung wie x und es ist $\|y\|_2 = 1$.
Man nennt y den normierten Vektor zu x.

Definition 6.1.10. Seien $a = (\alpha_1, \cdots, \alpha_n)^T$ und $b = (\beta_1, \cdots, \beta_n)^T$ zwei Vektoren des $\mathbb{R}^n$. Das euklidische innere Produkt von a und b ist definiert durch

$$\langle a, b \rangle := \sum_{i=1}^{n} \alpha_i \beta_i.$$

Schreibweise 6.1.11. Andere Schreibweisen fuer das euklidische innere Produkt sind (a, b), $a \cdot b$, $\langle a \mid b \rangle$.
Es gibt auch andere innere Produkte auf $\mathbb{R}^n$. Man nennt das innere Produkt auch skalares Produkt. Wie man ein inneres Produkt im Allgemeinen definiert, sehen wir in Definition 6.2.12.

Der folgende Trick ist immer wieder nützlich:

Bemerkung 6.1.12. Seien $a, b \in \mathbb{R}^n$. Das euklidische innere Produkt lässt sich durch Matrixmultiplikation schreiben:

$$\langle a, b \rangle = a^T b.$$

Satz 6.1.13. *Das euklidische innere Produkt hat folgende Eigenschaften: Seien $a, b, c \in \mathbb{R}^n$, $\lambda, \mu \in \mathbb{R}$. Dann gilt*

1) $\langle a, a \rangle \geq 0$. *Es ist* $\langle a, a \rangle = 0$ *genau dann, wenn* $a = 0$.
2) $\langle a, b \rangle = \langle b, a \rangle$,
3) $\langle \lambda a + \mu b, c \rangle = \lambda \langle a, c \rangle + \mu \langle b, c \rangle$,
4) $\langle c, \lambda a + \mu b \rangle = \lambda \langle c, a \rangle + \mu \langle c, b \rangle$.
5) $\|a\|_2^2 = \langle a, a \rangle$.
6) $|\langle a, b \rangle| \leq \|a\|_2 \, \|b\|_2$.

Beweis. Die meisten Eigenschaften lassen sich leicht nachrechnen. Nur die sogenannte Cauchy-Schwarzsche Ungleichung (6) ist trickreicher. Weil wir diese Ungleichung in abstrakter Form in der nächsten Sektion beweisen werden, lassen wir hier den Beweis aus - siehe Satz 6.2.13. □

Die folgende Bemerkung gibt erst die Motivation für die Untersuchung des euklidischen inneren Produkte. Wenn wir auf einander orthogonale Vektoren damit charakterisieren wollen, dass der Satz von Pythagoras gilt, ergibt sich eine sehr einfache Charakterisierung von Orthogonalität durch das innere Produkt:

Satz 6.1.14. *Seien $a = (\alpha_1, \cdots, \alpha_n)^T$ und $b = (\beta_1, \cdots, \beta_n)^T$ zwei Vektoren aus $\mathbb{R}^n$. Dann gilt*

$$\|a \pm b\|_2^2 = \|a\|^2 + \|b\|^2 \pm 2\langle a, b \rangle,$$
$$\|a\|_2^2 + \|b\|_2^2 = \|a \pm b\|_2^2 \ \Leftrightarrow \ \langle a, b \rangle = 0.$$

Beweis.

$$\begin{aligned}\|a \pm b\|_2^2 &= \sum_{i=1}^{n}(\alpha_i \pm \beta_i)^2 = \sum_{i=1}^{n}(\alpha_i^2 + \beta_i^2 \pm 2\alpha_i\beta_i) \\ &= \sum_{i=1}^{n}\alpha_i^2 + \sum_{i=1}^{n}\beta_i^2 \pm 2\sum_{i=1}^{n}\alpha_i\beta_i = \|a\|_2^2 + \|b\|_2^2 \pm 2\langle a, b\rangle.\end{aligned}$$

□

Und nun drehen wir wieder den Spieß um: Statt davon auszugehen, was ein rechter Winkel ist, und dann den Pythagoräischen Lehrsatz zu beweisen, definieren wir mit Hilfe des inneren Produktes was Orthogonalität heißt.

Definition 6.1.15. Seien $a, b \in \mathbb{R}^n$, U, V zwei Teilmengen des $\mathbb{R}^n$. Dann sagt man

1) Die Vektoren a und b stehen aufeinander orthogonal, wenn gilt $\langle a, b\rangle = 0$.
2) Der Vektor a steht orthogonal auf die Menge U, wenn gilt: $(\forall u \in U)\, \langle a, u\rangle = 0$.
3) Die Mengen U und V stehen orthogonal aufeinander, wenn gilt: $(\forall u \in U)\, (\forall v \in V)\, \langle u, v\rangle = 0$.
4) Das orthogonale Komplement von U ist definiert durch

$$U^{\perp} := \{x \in \mathbb{R}^n \mid x \text{ steht orthogonal auf } U\}.$$

Bemerkung 6.1.16. Besonders leicht ist das Auffinden von Normalvektoren in $\mathbb{R}^2$: Ist $a = (\alpha_1, \alpha_2)^T \in \mathbb{R}^2 \setminus \{0\}$, so sind die beiden Vektoren $(\pm\alpha_2, \mp\alpha_1)^T$ die einzigen Vektoren des $\mathbb{R}^2$, welche auf a orthogonal stehen, und die gleiche euklidische Norm wie a haben.

Beweis. Leichte Rechnung! Man muss allerdings die Fälle $\alpha_1 = 0$ und $\alpha_1 \neq 0$ unterscheiden. □

Eine geometrisch befriedigende Definition des Winkels würde einen viel tieferen, axiomatischen Zugang zur Elementargeometrie bedeuten. Für die lineare Algebra genügt es aber, zur Definition des Winkels den Cosinussatz heranzuziehen: Demnach müsste für den Winkel γ zwischen zwei Vektoren a, b gelten

$$\|a - b\|_2^2 = \|a\|_2^2 + \|b\|_2^2 - 2\|a\|_2\,\|b\|_2\,\cos(\gamma).$$

Ein Vergleich mit Satz 6.1.14 würde ergeben

$$\cos(\gamma) = \frac{\langle a, b\rangle}{\|a\|_2\ \|b\|_2}$$

Wieder verwenden wir diese Formel als die Definition des Winkels:

Definition 6.1.17. Seien $a, b \in \mathbb{R}^n \setminus \{0\}$. Der Winkel zwischen a und b ist definiert durch

$$\measuredangle(a, b) := \arccos\left(\frac{\langle a, b\rangle}{\|a\|_2\ \|b\|_2}\right).$$

Diese Definition braucht eine Rechtfertigung: Der Arcuscosinus kann nur von Zahlen genommen werden, die im Intervall $[-1, 1]$ liegen. Nun ist aber wegen Satz 6.1.13 (6)

$$\left|\frac{\langle a, b\rangle}{\|a\|_2\ \|b\|_2}\right| \leq 1.$$

Damit derArcuscosinus eindeutig bestimmt ist, lassen wir nur Winkel im Intervall $[0, \pi]$ zu.

6.1.3. *Einige geometrische Aufgaben.*

Das sind einige Beispiele von einfachen geometrischen Übungsaufgaben. Es ist nicht beabsichtigt, eine Liste von Rezepten zur Lösung geometrischer Standardaufgaben aufzustellen und zu beweisen. Statt dessen sollen Ihnen diese Beispiel die Denkweise und den rechnerischen Umgang mit euklidischer Norm und Skalarprodukt nahe bringen. Wenn man denken kann, muss man keine Rezepte auswendig lernen, sie ergeben sich dann ganz von selbst mit etwas Hausverstand. Erfinden Sie selbst ähnliche Aufgaben und systematische Lösungsverfahren.

Beispiel 6.1.18. Bestimmen Sie alle Vektoren des $\mathbb{R}^4$, die auf die Vektoren $a = (1,2,1,3)^T$, $b = (5,6,2,10)^T$ und $c = (6,8,3,13)^T$ orthogonal stehen.

Lösung: Wir suchen die Vektoren $x = (\xi_1, \cdots, \xi_4)^T$, welche erfüllen:

$$\langle a, x\rangle = 0,\ \langle b, x\rangle = 0,\ \langle c, x\rangle = 0.$$

Mit der Definition des euklidischen inneren Produktes ergibt sich das homogene lineare Gleichungssystem:

$$\begin{pmatrix} 1 & 2 & 1 & 3 \\ 5 & 9 & 2 & 10 \\ 6 & 11 & 3 & 13 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Wir bestimmen also den Kern einer Matrix. Durch Pivotschritte transformiert man die Matrix:

$$\begin{pmatrix} 1* & 2 & 1 & 3 \\ 5 & 9 & 2 & 10 \\ 6 & 11 & 3 & 13 \end{pmatrix} \to \begin{pmatrix} 1* & 2 & 1 & 3 \\ 0 & -1* & -3 & -5 \\ 0 & -1 & -3 & -5 \end{pmatrix} \to \begin{pmatrix} 1* & 0 & -5 & -7 \\ 0 & 1* & 3 & 5 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Durch Nullsetzen aller nichtbasischen Variablen außer jeweils einer erhalten wir eine Basis des Kernes:

$$x_1 = \begin{pmatrix} 5 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 7 \\ -5 \\ 0 \\ 1 \end{pmatrix}.$$

Die Menge der Vektoren, die auf a und b orthogonal stehen, ist der von x_1 und x_2 aufgespannte Unterraum. □

Beispiel 6.1.19. Beschreiben Sie die Gerade

$$g = \{\begin{pmatrix} 3 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 4 \end{pmatrix} \mid \lambda \in \mathbb{R}\}$$

durch eine Gleichung:

$$g = \{\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \in \mathbb{R}^2 \mid \alpha_1\xi_1 + \alpha_2\xi_2 = \gamma\}.$$

Lösung: Wir wollen einen 1-dimensionalen affinen Unterraum des $\mathbb{R}^2$ beschreiben, daher brauchen wir ein Gleichungssystem mit einer Systemmatrix von Rang 1. Wir wählen einen beliebigen Normalvektor (ungleich 0) zum Richtungsvektor der Geraden, zum Beispiel $(4,-3)^T$. Die Matrix $(4,-3)$ hat offensichtlich Rang 1. Nun ist für alle $x \in g$:

$$\begin{aligned} (4,-3)x &= \langle \begin{pmatrix} 4 \\ -3 \end{pmatrix}, x\rangle = \langle \begin{pmatrix} 4 \\ -3 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 4 \end{pmatrix}\rangle \\ &= \langle \begin{pmatrix} 4 \\ -3 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}\rangle + \langle \begin{pmatrix} 4 \\ -3 \end{pmatrix}, \lambda \begin{pmatrix} 3 \\ 4 \end{pmatrix}\rangle = 9 + \lambda\, 0 = 9. \end{aligned}$$

Also ist g eine Teilmenge der Lösungsmenge der Gleichung

$$4\xi_1 - 3\xi_2 = 9.$$

Da die Lösungsmenge dieser Gleichung selbst ein eindimensionaler affiner Unterraum des $\mathbb{R}^2$ ist, kann g keine echte Teilmenge davon sein. Daher ist

$$g = \{x \in \mathbb{R} \mid \langle \begin{pmatrix} 4 \\ -3 \end{pmatrix}, x\rangle = 9\}.$$

□

Bemerkung 6.1.20. Sei $g = \{x_0 + \lambda v \mid \lambda \in \mathbb{R}\}$ eine Gerade in $\mathbb{R}^2$. Sei $n \neq 0$ ein beliebiger Normalvektor auf g. Dann ist

$$g = \{x \in \mathbb{R}^2 \mid \langle n, x\rangle = \langle n, x_0\rangle\}.$$

Man nennt diese Darstellung einer Geraden des $\mathbb{R}^2$ die Normalvektorform.

Beispiel 6.1.21. Beschreiben Sie die Gerade

$$g = \{\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R}\}$$

durch ein System von Gleichungen:

$$g = \{\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \in \mathbb{R}^3 \mid \alpha_1\xi_1 + \alpha_2\xi_2 + \alpha_3\xi_3 = \gamma,$$
$$\beta_1\xi_1 + \beta_2\xi_2 + \beta_3\xi_3 = \delta\}.$$

Lösung: Um im $\mathbb{R}^3$ einen eindimensionalen affinen Raum als Lösung eines linearen Gleichungssystems zu erhalten, brauchen wir eine Systemmatrix mit Rang 2. Wir wählen also zwei linear unabhängige Normalvektoren auf den Richtungsvektor $(2, 2, 1)^T$ der Geraden. Suchen wir zunächst einen Normalvektor $a = (\alpha_1, \alpha_2, \alpha_3)^T$ mit dem Ansatz $\alpha_1 = 0$. Es muß also dann gelten

$$0 = \langle \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}\rangle = 2\alpha_2 + \alpha_3.$$

Das ist erfüllt mit $a = (0, 1, -2)^T$. Suchen wir nun einen Normalvektor $b = (\beta_1, 0, \beta_3)^T$, so erhalten wir ebenso (zum Beispiel) $b = (1, 0, -2)^T$. Sei nun $x \in g$. Dann ist

$$\begin{aligned} \langle a, x\rangle &= \langle \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}\rangle = \langle \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}\rangle + \lambda\, 0 = -4, \\ \langle b, x\rangle &= \langle \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}\rangle = \langle \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}\rangle + \lambda\, 0 = -3. \end{aligned}$$

Daher ist g eine Teilmenge der Lösungsmenge des folgenden Systems:

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

Da die Lösungsmenge dieses Gleichungssystems ein eindimensionaler affiner Raum ist, und g selbst eindimensional ist, ist

$$g = \{x \in \mathbb{R}^3 \mid \begin{pmatrix} 0 & 1 & -2 \\ 1 & 0 & -2 \end{pmatrix} x = \begin{pmatrix} -4 \\ -3 \end{pmatrix}\}.$$

□

Bemerkung 6.1.22. Einen Normalvektor auf einen Vektor $x \in \mathbb{R}^n$ kann man erhalten, indem man alle Komponenten außer zweien mit Null ansetzt, anschließend die beiden verbleibenden Komponenten von x vertauscht, und bei einer von beiden das Vorzeichen umdreht.

Beispiel 6.1.23. Sei g die Gerade aus Beispiel 6.1.21. Sei $x = (0, -10, 6)^T$. Bestimmen Sie den Punkt auf g, der am nächsten an x liegt. Wie groß ist der Normalabstand von x zu g?

Lösung: Sei y der Punkt auf g, der am nächsten an x liegt. Dann muß der Vektor $y-x$ auf den Richtungsvektor der Geraden orthogonal stehen. Als Punkt der Geraden läßt sich y durch die Parameterform schreiben:

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

Nun ist

$$\begin{aligned} 0 &= \langle x - y, \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \rangle = \langle \begin{pmatrix} 0 \\ -10 \\ 6 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - \lambda \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \rangle \\ &= \langle \begin{pmatrix} -1 \\ -10 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \rangle - \lambda \langle \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \rangle = -18 - 9\lambda. \end{aligned}$$

Also ist $\lambda = -2$ und

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

Der Normalabstand von x zur Geraden ist die Norm von $x - y$, also

$$\| \begin{pmatrix} 0 \\ -10 \\ 6 \end{pmatrix} - \begin{pmatrix} -3 \\ -4 \\ 0 \end{pmatrix} \|_2 = \| \begin{pmatrix} 3 \\ -6 \\ 6 \end{pmatrix} \|_2 = \sqrt{9 + 36 + 36} = 9.$$

□

Wir haben verwendet, dass der nächste Punkt y zu x auf der Geraden dort liegt, wo $x - y$ normal auf die Gerade steht. Das ist uns anschaulich klar. In Unterabschnitt 6.3 werden wir das aber beweisen.

Beispiel 6.1.24. Beschreiben Sie die Kreislinie mit Mittelpunkt $(2, 1)^T$ und Radius 4 durch eine Gleichung.

Lösung: Die Kreislinie ist die Menge aller Punkte, die von $(2, 1)^T$ den Abstand 4 haben. Ist $x = (\xi_1, \xi_2)$, so ist

$$\begin{aligned} &\|x - \begin{pmatrix} 2 \\ 1 \end{pmatrix}\|_2^2 = \| \begin{pmatrix} \xi_1 - 2 \\ \xi_2 - 1 \end{pmatrix} \|_2^2 \\ = \; &(\xi_1 - 2)^2 + (\xi_2 - 1)^2 = \xi_1^2 + \xi_2^2 - 4\xi_1 - 2\xi_2 + 5. \end{aligned}$$

Die Gleichung dieser Kreislinie ist also

$$\|x - \begin{pmatrix} 2 \\ 1 \end{pmatrix}\|_2^2 = 16,$$

oder, ausgeschrieben in Koordinaten

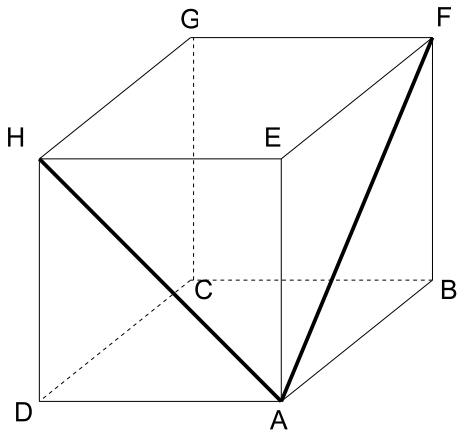
$$\xi_1^2 + \xi_2^2 - 4\xi_1 - 2\xi_2 - 11 = 0.$$

□

Bemerkung 6.1.25. Die Gleichung eines Kreises in $\mathbb{R}^2$ mit Mittelpunkt x_0 und Radius r ist

$$\|x - x_0\|^2 = r^2.$$

Beispiel 6.1.26. Die folgende Grafik zeigt einen Würfel mit den Ecken $A, \cdots, H$. Bestimmen Sie den Winkel zwischen den Geraden $\overline{AF}$ und $\overline{AH}$.



Lösung: Wir können ohne Beschränkung der Allgemeinheit annehmen, dass unser Würfel Seitenlänge 1 hat, und legen unser Koordinatensystem so an, dass der Nullpunkt im Punkt A liegt, und die drei Einheitsvektoren die Vektoren $e_1 = \vec{AB}$, $e_2 = \vec{AD}$ und $e_3 = \vec{AE}$ sind. Dann ist

$$\vec{AF} = e_1 + e_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{AH} = e_2 + e_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Man berechnet leicht:

$$\|\vec{AF}\|_2 = \|\vec{AH}\|_2 = \sqrt{2}, \quad \langle \vec{AF}, \vec{AH} \rangle = 1.$$

Daher ist

$$\cos(\measuredangle(\vec{AF}, \vec{AH})) = \frac{1}{\sqrt{2}\sqrt{2}} = \frac{1}{2},$$

und damit ist der Winkel zwischen $\vec{AF}$ und $\vec{AH}$ gleich $\pi/6$ oder $60°$. □

Sehen Sie die Grafik genau an. Gibt es einen kürzeren, elementargeometrischen Beweis, um zu sehen, dass dieser Winkel $60°$ sein muß?

6.1.4. *Äußeres Produkt in* $\mathbb{R}^3$.

Das äußere Produkt oder Kreuzprodukt gibt es nur in $\mathbb{R}^3$. Es spielt eine große Rolle in der Mechanik, z.B. zur mathematischen Formulierung von Drehimpuls und Drehmoment, und in der Elektrodynamik bei der Beschreibung der Wechselwirkung zwischen magnetischen und elektrischen Feldern. In dieser Vorlesung streifen wir es nur soweit, dass Sie damit rechnen können.

Definition 6.1.27. Seien $a = (\alpha_1, \alpha_2, \alpha_3)^T$ und $b = (\beta_1, \beta_2, \beta_3)^T$ zwei Vektoren des $\mathbb{R}^3$. Dann ist das äußere Produkt (Kreuzprodukt, vektorielle Produkt) von a und b definiert mit Hilfe von Determinanten:

$$\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} \times \begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\end{pmatrix} := \begin{pmatrix} \begin{vmatrix}\alpha_2 & \beta_2\\ \alpha_3 & \beta_3\end{vmatrix} \\ -\begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_3 & \beta_3\end{vmatrix} \\ \begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_2 & \beta_2\end{vmatrix} \end{pmatrix}$$

Achtung auf das negative Vorzeichen bei der mittleren Determinante!

Satz 6.1.28. *Seien $a, b, c \in \mathbb{R}^3$, $\lambda \in \mathbb{R}$. Dann gilt:*

1) *Für alle $x \in \mathbb{R}^3$ ist $\det(a, b, x) = \langle a \times b, x\rangle$.*
2) *$a \times b$ steht orthogonal zu a und orthogonal zu b.*
3) *Es ist $\det(a, b, a \times b) = \|a \times b\|_2^2$.*
4) *Es gelten die Rechenregeln*

$$\begin{aligned} &a \times (b + c) = a \times b + a \times c,\\ &a \times (\lambda b) = \lambda(a \times b) = (\lambda a) \times b,\\ &a \times b = -b \times a. \end{aligned}$$

Beweis. Teil (1): Wir entwickeln $\det(a, b, x)$ nach der letzten Spalte:

$$\begin{aligned} \begin{vmatrix}\alpha_1 & \beta_1 & \xi_1\\ \alpha_2 & \beta_2 & \xi_2\\ \alpha_3 & \beta_3 & \xi_3\end{vmatrix} &= \xi_1 \begin{vmatrix}\alpha_2 & \beta_2\\ \alpha_3 & \beta_3\end{vmatrix} - \xi_2 \begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_3 & \beta_3\end{vmatrix} + \xi_3 \begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_2 & \beta_2\end{vmatrix}\\ &= \left\langle \begin{pmatrix}\xi_1\\ \xi_2\\ \xi_3\end{pmatrix}, \begin{pmatrix} \begin{vmatrix}\alpha_2 & \beta_2\\ \alpha_3 & \beta_3\end{vmatrix} \\ -\begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_3 & \beta_3\end{vmatrix} \\ +\begin{vmatrix}\alpha_1 & \beta_1\\ \alpha_2 & \beta_2\end{vmatrix} \end{pmatrix} \right\rangle. \end{aligned}$$

Teil (2): folgt direkt aus (1), weil

$$\det(a, b, a) = \det(a, b, b) = 0.$$

Teil (3): folgt direkt aus (1) mit $x = a \times b$.
Teil (4): Direktes Nachrechnen. □

Bemerkung 6.1.29. Seien a, b zwei Vektoren des $\mathbb{R}^3$. Dann ist $a \times b$ jener Vektor, der folgende Eigenschaften hat:

1) Er steht sowohl auf a als auch b orthogonal.
2) Sein Betrag ist die Fläche des von a und b aufgespannten Parallelogramms.
3) Das System $a, b, a \times b$ ist positiv orientiert, falls $a \times b \neq 0$.

Beweis. Teil (1): Siehe Satz 6.1.28 (2).
Teil (2): Wir haben die Fläche eines Parallelogramms in $\mathbb{R}^3$ nicht definiert, unser "Beweis" kann daher nur eine Heuristik sein. Da $a \times b$ orthogonal auf a und b steht, ist das Volumen des von a, b und $a \times b$ aufgespannten Parallelepipeds die Fläche F

des Parallelogramms von a und b, multipliziert mit der Norm von $a \times b$. Es ist also (unter Benützung von Satz 6.1.28 (3))

$$F\|a \times b\|_2 = \det(a, b, a \times b) = \|a \times b\|_2^2.$$

Daher ist $F = \|a \times b\|_2$.
Teil (3): Folgt aus $\det(a, b, a \times b) = \|a \times b\|^2 > 0$. □

Beispiel 6.1.30. Beschreiben Sie die Ebene

$$E = \{ \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ 6 \\ 1 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \}$$

durch eine Gleichung

$$E = \{ \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} \mid \alpha_1\xi_1 + \alpha_2\xi_2 + \alpha_3\xi_3 = \gamma \}.$$

Lösung: Um einen zweidimensionalen affinen Unterraum des $\mathbb{R}^3$ zu beschreiben, braucht man ein lineares Gleichungssystem mit Rang 1. Wir suchen einen Vektor, der auf die ganze Ebene normal steht, und bekommen diesen durch das Kreuzprodukt der beiden Richtungsvektoren in der Parameterform:

$$u := \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} \times \begin{pmatrix} 4 \\ 6 \\ 1 \end{pmatrix} = \begin{pmatrix} \begin{vmatrix} 1 & 6 \\ -2 & 1 \end{vmatrix} \\ -\begin{vmatrix} 3 & 4 \\ -2 & 1 \end{vmatrix} \\ \begin{vmatrix} 3 & 4 \\ 1 & 6 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} 13 \\ -11 \\ 14 \end{pmatrix}.$$

Sei nun $x \in E$. Dann ist

$$\begin{aligned} \langle \begin{pmatrix} 13 \\ -11 \\ 14 \end{pmatrix}, x \rangle &= \langle \begin{pmatrix} 13 \\ -11 \\ 14 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ 6 \\ 1 \end{pmatrix} \rangle \\ &= \langle \begin{pmatrix} 13 \\ -11 \\ 14 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix} \rangle + \lambda 0 + \mu 0 = 85\,. \end{aligned}$$

Die Gleichung der Ebene in Normalvektorform ist also

$$\langle \begin{pmatrix} 13 \\ -11 \\ 14 \end{pmatrix}, x \rangle = 85\,.$$

□

6.1.5. *Norm und inneres Produkt in $\mathbb{C}^n$.*

Wir erwarten uns von einer Norm, dass sie die Länge eines Vektors, also eine nichtnegative reelle Zahl ergibt. Würde man die Norm eines Vektors über $\mathbb{C}$ durch

$$\| \begin{pmatrix} \xi_1^2 \\ \vdots \\ \xi_n^2 \end{pmatrix} \|_2^2 := \sum_{k=1}^{n} \xi_k^2$$

definieren, können sich auch negative und komplexe Zahlen ergeben. Wir erinnern uns, dass für komplexe Zahlen $z \in \mathbb{C}$ gilt

$$|z|^2 = \overline{z}\, z,$$

wobei $\overline{z}$ die zu z konjugiert komplexe Zahl bedeutet. Damit erhalten wir eine positive Norm und ein dazu passendes inneres Produkt durch folgende Definition:

Definition 6.1.31. Seien $x = (\xi_1, \cdots, xi_n)^T$, $y = (\eta_1, \cdots, \eta_n)^T \in \mathbb{C}^n$. Dann definieren wir

$$\|x\|_2 := \sqrt{\sum_{k=1}^{n} |\xi_k|^2}\,,$$

$$\langle x, y\rangle := \sum_{k=1}^{n} \overline{\xi_k}\ \eta_k\,.$$

Achtung: viele Autoren legen die konjugiert komplexe Zahl auf die rechte Seite und definieren:

$$\langle x, y\rangle := \sum_{k=1}^{n} \xi_k\ \overline{\eta_k}\,.$$

Die Regeln für Norm und inneres Produkt bleiben wie in Satz 6.1.8 und Satz 6.1.13, mit den folgenden wichtigen Ausnahmen:

Bemerkung 6.1.32. Sind $a, b, c \in \mathbb{C}^n$, $\lambda, \mu \in \mathbb{C}$, so ist

$$\begin{aligned}
\langle a, \lambda b + \mu c\rangle &= \lambda\langle a, b\rangle + \mu\langle a, c\rangle,\\
\langle \lambda b + \mu c, a\rangle &= \overline{\lambda}\langle b, a\rangle + \overline{\mu}\langle c, a\rangle,\\
\langle a, b\rangle &= \overline{\langle b, a\rangle}.
\end{aligned}$$

Bemerkung 6.1.33. Das innere Produkt zweier Vektoren $a, b \in \mathbb{C}^n$ lässt sich mit Hilfe der adjungierten Matrix a^* als Matrixprodukt schreiben:

$$\langle a, b\rangle = a^* b.$$

Beweis. Die Adjungierte einer Matrix entsteht, indem man die Matrix transponiert und von allen Einträgen die konjugierte Zahl nimmt. Damit ist

$$\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}^* \begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix} = \begin{pmatrix}\overline{\alpha_1} & \cdots & \overline{\alpha_n}\end{pmatrix} \begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix} = \sum_{k=1}^{n} \overline{\alpha_k}\beta_k = \left\langle \begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix}, \begin{pmatrix}\beta_1\\ \vdots\\ \beta_n\end{pmatrix} \right\rangle.$$

□

6.2. **Norm und inneres Produkt in abstrakten Vektorräumen.**

6.2.1. *Definition einer Norm.*

Definition 6.2.1. Sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Eine Norm auf V ist eine Abbildung

$$\begin{cases} V & \to \mathbb{R}\\ x & \mapsto \|x\| \end{cases}$$

mit folgenden Eigenschaften:

1) Für alle $x \in V$ ist $\|x\| \geq 0$. Die Norm $\|x\|$ ist genau dann Null, wenn $x = 0$.
2) Für alle $x \in V$ und alle $\lambda \in \mathbb{K}$ gilt $\|\lambda x\| = |\lambda|\,\|x\|$.
3) ("Dreiecksungleichung") Für alle $x, y \in V$ gilt $\|x + y\| \leq \|x\| + \|y\|$.

Ist V ein Vektorraum mit einer Norm $\|\cdot\|$, so heißt $(V, \|\cdot\|)$ ein normierter Vektorraum.

Bemerkung 6.2.2. Beachten Sie, dass wir nun nicht mehr über beliebigen Körpern arbeiten können, da wir in Definition 6.2.1 (2) den Absolutbetrag eines Skalars benötigen, und in Punkt (3) auf eine Ordnungsrelation $\leq$ auf der Menge der Skalare Bezug nehmen.

Die folgende Umkehrung der Dreiecksungleichung wird sehr oft verwendet:

Satz 6.2.3. *Sei $V, \|\cdot\|$ ein normierter Vektorraum. Dann gilt für alle $x, y \in V$:*

$$\big|\, \|x\| - \|y\| \,\big| \leq \|x - y\| \,.$$

Beweis. Wegen der Dreiecksungleichung ist

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\| \,,$$

daher ist

$$\|x\| - \|y\| \leq \|x - y\| \,.$$

Ebenso ist

$$\|y\| - \|x\| \leq \|y - x\| = \|x - y\|.$$

□

Um ein erstes Beispiel von Normen vorzubereiten, brauchen wir zunächst ein Lemma, das für sich selbst genommen bedeutend ist:

Lemma 6.2.4. *Seien $p, q \in (1, \infty)$ so, dass $p^{-1} + q^{-1} = 1$.*

1) *Seien $\xi, \eta \in [0, \infty)$. Dann ist*[22]

$$\xi\eta \leq \frac{1}{p}\,\xi^p + \frac{1}{q}\,\eta^q.$$

2) *("Höldersche Ungleichung in $\mathbb{C}^n$") Seien $\alpha_1, \cdots, \alpha_n, \beta_1, \cdots, \beta_n \in \mathbb{C}$. Dann ist*

$$\sum_{j=1}^{n} |\alpha_j \beta_j| \leq \left(\sum_{j=1}^{n} |\alpha_j|^p \right)^{1/p} \left(\sum_{j=1}^{n} |\beta_j|^q \right)^{1/q} .$$

Beweis. Punkt (1): Trivial falls $\xi = 0$ oder $\eta = 0$. Fixieren wir nun ein $\eta \in (0, \infty)$. Durch Differentialrechnung suchen wir das Minimum der Funktion

$$f : \begin{cases} (0,\infty) & \to (0,\infty) \\ \xi & \mapsto \frac{1}{p}\,\xi^p + \frac{1}{q}\,\eta^q - \xi\eta. \end{cases}$$

Wir beobachten zuerst, dass $\lim_{\xi\to\infty} f(\xi) = \infty$. Ferner ist $f(0) = q^{-1}\eta^q > 0$. Die erste Ableitung ist

$$f'(\xi) = \xi^{p-1} - \eta.$$

Nullsetzen ergibt, dass die einzige mögliche Extremstelle in $(0, \infty)$ bei $\xi = \eta^{1/(p-1)}$ liegt. Beachten Sie, dass $q = p/(p-1)$. An der Extremstelle ist daher

$$f(\eta^{1/(p-1)}) = \frac{1}{p}\,\eta^{p/(p-1)} + \frac{1}{q}\eta^q - \eta^{1/(p-1)}\eta = (\frac{1}{p} + \frac{1}{q} - 1)\,\eta^q = 0.$$

Daher gilt für alle $\xi, \eta \in (0, \infty)$:

$$\xi\eta \leq \frac{1}{p}\,\xi^p + \frac{1}{q}\,\eta^q.$$

[22]Das ist ein Spezialfall der Youngschen Ungleichung: Ist $f : [0, \infty) \to [0, \infty)$ streng monoton wachsend und stetig mit $f(0) = 0$, ist f^{-1} die Umkehrfunktion von f, und sind $\xi, \eta \geq 0$, so ist

$$\xi\eta \leq \int_0^\xi f(x)\,dx + \int_0^\eta f^{-1}(y)\,dy \,.$$

Die Ungleichung in Punkt (1) erhält man mit $f(\xi) := \xi^{p-1}$.

Punkt (2): OBdA seien nicht alle α_j und nicht alle β_j gleich Null, sonst ist die Aussage ohnehin trivial. Zur Abkürzung schreiben wir

$$A := \sum_{j=1}^{n} |\alpha_j|^p > 0,\ B := \sum_{j=1}^{n} |\beta_j|^q > 0.$$

Es gilt wegen Punkt (1) für jedes $\epsilon > 0$

$$\begin{aligned} \sum_{j=1}^{n} |\alpha_j \beta_j| &= \sum_{j=1}^{n} |\epsilon \alpha_j \frac{\beta_j}{\epsilon}| \leq \sum_{j=1}^{n} \left(\frac{1}{p} |\epsilon \alpha_j|^p + \frac{1}{q} |\frac{\beta_j}{\epsilon}|^q \right) \\ = \quad & \frac{1}{p} \epsilon^p \sum_{j=1}^{n} |\alpha_j|^p + \frac{1}{q} \epsilon^{-q} \sum_{j=1}^{n} |\beta_j|^q = \frac{1}{p} \epsilon^p A + \frac{1}{q} \epsilon^{-q} B \,. \end{aligned}$$

Wir wählen nun

$$\epsilon := A^{-1/(pq)} B^{1/(pq)}.$$

Dann ist

$$\epsilon^p A = A^{1-1/q} B^{1/q} = A^{1/p} B^{1/q}, \quad \epsilon^{-q} B = A^{1/p} B^{1-1/p} = A^{1/p} B^{1/q}.$$

Folglich ist

$$\sum_{j=1}^{n} |\alpha_j \beta_j| \leq \frac{1}{p} A^{1/p} B^{1/q} + \frac{1}{q} A^{1/p} B^{1/q} = A^{1/p} B^{1/q}.$$

□

Beispiel 6.2.5. Die folgenden Definitionen geben Normen in $\mathbb{R}^n$ bzw. $\mathbb{C}^n$: Sei $a = (\alpha_1, \cdots, \alpha_n)^T$:

$$\begin{aligned} \|a\|_p &:= \left(\sum_{i=1}^{n} |\alpha_i|^p \right)^{1/p} \quad \text{für } 1 \leq p < \infty, \\ \|a\|_\infty &:= \max(\{|\alpha_i| \ \mid \ i = 1, \cdots, n\}). \end{aligned}$$

Beweis. Wir überlassen dem Leser (das sind Sie!) die einfachen Beweise, dass $\|\cdot\|_1$ und $\|\cdot\|_\infty$ Normen sind. Ebenso einfach ist der Beweis für Punkte (1) und (2) von Definition 6.2.1 für $\|\cdot\|_p$. Etwas trickreicher ist nur die Dreiecksungleichung für $\|\cdot\|_p$. Seien $a = (\alpha_1, \cdots, \alpha_n)^T$ und $b = (\beta_1, \cdots, \beta_n)^T$ in $\mathbb{C}^n$. Wir setzen $q = p/(p-1)$, so

dass $p^{-1} + q^{-1} = 1$. In der folgenden Rechnung benützen wir Lemma 6.2.4 (2):

$$\begin{aligned}
\sum_{j=1}^{n} |\alpha_j + \beta_j|^p &= \sum_{j=1}^{n} |\alpha_j + \beta_j|\,|\alpha_j + \beta_j|^{p-1} \\
&\leq \sum_{j=1}^{n} |\alpha_j|\,|\alpha_j + \beta_j|^{p-1} + \sum_{j=1}^{n} |\beta_j|\,|\alpha_j + \beta_j|^{p-1} \\
&\leq \left(\sum_{j=1}^{n} |\alpha_j|^p\right)^{1/p} \left(\sum_{j=1}^{n} |\alpha_j + \beta_j|^{q(p-1)}\right)^{1/q} \\
&\quad + \left(\sum_{j=1}^{n} |\beta_j|^p\right)^{1/p} \left(\sum_{j=1}^{n} |\alpha_j + \beta_j|^{q(p-1)}\right)^{1/q} \\
&= \left(\sum_{j=1}^{n} |\alpha_j|^p\right)^{1/p} \left(\sum_{j=1}^{n} |\alpha_j + \beta_j|^p\right)^{1-1/p} \\
&\quad + \left(\sum_{j=1}^{n} |\beta_j|^p\right)^{1/p} \left(\sum_{j=1}^{n} |\alpha_j + \beta_j|^p\right)^{1-1/p}.
\end{aligned}$$

Wir kürzen nun durch $\|a + b\|_p^{1-1/p}$ und erhalten die Dreiecksungleichung:

$$\left(\sum_{j=1}^{n} |\alpha_j + \beta_j|^p\right)^{1/p} \leq \left(\sum_{j=1}^{n} |\alpha_j|^p\right)^{1/p} + \left(\sum_{j=1}^{n} |\beta_j|^p\right)^{1/p}.$$

□

Beispiel 6.2.6 (Gewichtete Normen)**.** Seien $\gamma_1, \cdots, \gamma_n > 0$. Für Vektoren $a = (\alpha_1, \cdots, \alpha_n) \in \mathbb{C}^n$ definieren wir folgende Normen:

$$\max_{i=1,\cdots,n} \gamma_i|\alpha_i|, \quad \sum_{i=1}^{n} \gamma_i|\alpha_i|, \quad \left(\sum_{i=1}^{n} \gamma_i|\alpha_i|^p\right)^{1/p}.$$

Beispiel 6.2.7. Sei $\mathcal{C}([0,1], \mathbb{R})$ der Vektorraum der stetigen Funktionen von $[0,1]$ nach $\mathbb{R}$. Aus der Analysis wissen wir, dass jedes $f \in \mathcal{C}([0,1], \mathbb{R})$ ein (endliches) Maximum besitzt und sich integrieren lässt. Die folgenden Definitionen ergeben Normen auf $\mathcal{C}([0,1], \mathbb{R})$:

$$\begin{aligned}
\|f\|_{L^p} &:= \left(\int_0^1 |f(t)|^p\, dt\right)^{1/p} \quad \text{für } 1 \leq p < \infty, \\
\|f\|_{L^\infty} &:= \max_{t\in[0,1]} |f(t)|.
\end{aligned}$$

Beweis. $\|\cdot\|_{L^\infty}$ *ist Norm:* Zunächst ist wegen der Definition jedenfalls $\|f\|_{L^\infty} \geq 0$. Ist das Maximum aller $|f(t)|$ gleich Null, so ist natürlich $f(t) = 0$ für alle $t \in [0,1]$. Es ist

$$\|\lambda f\|_{L^\infty} = \max_{t\in[0,1]} |\lambda f(t)| = |\lambda| \max_{t\in[0,1]} |f(t)| = |\lambda|\,\|f\|_{L^\infty}.$$

Seien nun $f, g \in \mathcal{C}([0,1], \mathbb{R})$, und sei t_0 eine Maximumstelle für $|f(t) + g(t)|$. Es gilt dann natürlich

$$|f(t_0)| \leq \max_{t\in[0,1]} |f(t)| = \|f\|_{L^\infty}, \quad |g(t_0)| \leq \|g\|_{L^\infty}.$$

Daher gilt

$$\|f+g\|_{L^\infty} = |f(t_0)+g(t_0)| \le |f(t_0)| + |g(t_0)| \le \|f\|_{L^\infty} + \|g\|_{L^\infty}.$$

$\|\cdot\|_{L^1}$ *ist Norm:* Es ist klar, dass das Integral der Nullfunktion wieder Null ergibt. Andererseits sagt ein Satz der Analysis: Wenn das Integral einer *stetigen* nichtnegativen Funktion Null ergibt, so ist diese Funktion überall Null. Ferner ist

$$\int_0^1 |\lambda f(t)|\,dt = |\lambda| \int_0^1 |f(t)|\,dt$$

und

$$\int_0^1 |f(t)+g(t)|\,dt \le \int_0^t (|f(t)|+|g(t)|)\,dt \le \int_0^1 |f(t)|\,dt + \int_0^t |g(t)|\,dt.$$

$\|\cdot\|_p$*:* Punkte (1) und (2) von Definition 6.2.1 sind wieder leicht zu beweisen. Die Dreiecksungleichung muss ähnlich wie in der endlichen Dimension bewiesen werden, wir führen dies hier nicht durch. Das wichtigste Werkzeug dazu, die Höldersche Ungleichung, führen wir hier ohne Beweis an. □

Satz 6.2.8 (Höldersche Ungleichung)**.** *Seien f, g integrierbare Funktionen von einem (endlichen oder unendlichen) Intervall I in die reellen oder komplexen Zahlen. Seien $p, q > 0$ so, dass $p^{-1} + q^{-1} = 1$. Dann ist*

$$\int_I |f(t)g(t)|\,dt \le \left(\int_I |f(t)|^p\,dt\right)^{1/p} \left(\int_I |g(t)|^q\,dt\right)^{1/q}.$$

6.2.2. *Innere Produkte.*

Definition 6.2.9. Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Abbildung

$$a : \begin{cases} V \times V & \to \mathbb{K}, \\ (x,y) & \mapsto a(x,y) \end{cases}$$

heißt

Bilinearform, wenn gilt:	Sesquilinearform, wenn gilt:
$a(z, \lambda x + \mu y) = \lambda a(z,x) + \mu a(z,y)$	$a(z, \lambda x + \mu y) = \lambda a(z,x) + \mu a(z,y)$
$a(\lambda x + \mu y, z) = \lambda a(x,z) + \mu a(y,z)$	$a(\lambda x + \mu y, z) = \overline{\lambda} a(x,z) + \overline{\mu} a(y,z)$

Eine Bilinearform bzw. Sesquilinearform a heißt

symmetrisch, wenn gilt:	hermitesch, wenn gilt:
$a(y,x) = a(x,y)$,	$a(y,x) = \overline{a(x,y)}$.

Eine Bilinearform bzw. Sesquilinearform a heißt
positiv semidefinit, wenn für alle $x \in V$ gilt: $a(x,x) \in [0,\infty)$,
positiv definit, wenn für alle $x \in V \setminus \{0\}$ gilt $a(x,x) \in (0,\infty)$.

Bemerkung 6.2.10. Viele Autoren setzen in der Definition der Sesquilinearform die Konjugation bei der Summe in der zweiten Variablen:

$$\begin{aligned} a(\lambda x + \mu y, z) &= \lambda a(x,z) + \mu a(y,z) \\ a(z, \lambda x + \mu y) &= \overline{\lambda} a(z,x) + \overline{\mu} a(z,y). \end{aligned}$$

Bemerkung 6.2.11.

1) Sei $A \in \mathbb{R}^{n\times n}$. Dann ist
$$a : \begin{cases} \mathbb{R}^n \times \mathbb{R}^n & \to \mathbb{R} \\ (x,y) & \mapsto x^T A y \end{cases}$$
eine Bilinearform. Diese Bilinearform ist genau dann symmetrisch, wenn gilt $A^T = A$.
2) Sei $A \in \mathbb{C}^{n\times n}$. Dann ist
$$a : \begin{cases} \mathbb{C}^n \times \mathbb{C}^n & \to \mathbb{C} \\ (x,y) & \mapsto x^* A y \end{cases}$$
eine Bilinearform. Diese Bilinearform ist genau dann hermitesch, wenn gilt $A^* = A$.

Definition 6.2.12.

1) Sei V ein Vektorraum über $\mathbb{R}$. Ein inneres Produkt auf V ist eine symmetrische, positiv definite Bilinearform $\langle\cdot,\cdot\rangle$ von $V \times V$ nach $\mathbb{R}$.
2) Sei V ein Vektorraum über $\mathbb{C}$. Ein inneres Produkt auf V ist eine hermitesche, positiv definite Sesquilinearform $\langle\cdot,\cdot\rangle$ von $V \times V$ nach $\mathbb{C}$.

Satz 6.2.13 (Cauchy-Schwarz'sche Ungleichung)**.** *Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und $\langle\cdot,\cdot\rangle$ ein inneres Produkt auf V. Dann gilt:*

1) *Für alle $x, y \in V$ ist $|\langle x,y\rangle| \le \sqrt{\langle x,x\rangle\,\langle y,y\rangle}$.*
2) *Gleichheit in der obigen Ungleichung gilt genau dann, wenn x und y linear abhängig sind.*

Beweis. Wir bringen den Beweis für $\mathbb{K} = \mathbb{C}$. Für reelle Vektorräume geht er genauso. Seien zunächst x, y linear abhängig, also oBdA $y = \lambda x$ für ein $\lambda \in \mathbb{C}$. Es gilt
$$\begin{aligned} |\langle x,y\rangle| &= |\langle x,\lambda x\rangle| = |\lambda|\langle x,x\rangle\,, \\ (\langle x,x\rangle)^{1/2}\,(\langle y,y\rangle)^{1/2} &= (\langle x,x\rangle)^{1/2}\,(\langle \lambda x,\lambda x\rangle)^{1/2} = (\langle x,x\rangle)^{1/2}\left(|\lambda|^2\langle x,x\rangle\right)^{1/2} \\ &= |\lambda|\langle x,x\rangle. \end{aligned}$$
Seien nun x, y linear unabhängig. Vorläufig fixieren wir ein $\alpha \in \mathbb{C}$. Es ist $x+\alpha y \neq 0$. Wegen der positiven Definitheit und der Sesquilinearität des inneren Produktes gilt
$$\begin{aligned} 0 &< \langle x+\alpha y, x+\alpha y\rangle = \langle x,x\rangle + \overline{\alpha}\langle y,x\rangle + \alpha\langle x,y\rangle + \overline{\alpha}\alpha\langle y,y\rangle \\ &= \langle x,x\rangle + \overline{\alpha\langle x,y\rangle} + \alpha\langle x,y\rangle + |\alpha|^2\langle y,y\rangle. \end{aligned}$$
Wegen der linearen Unabhängigkeit ist $y \neq 0$ und daher $\langle y,y\rangle > 0$. Wir setzen
$$\alpha := -\,\frac{\overline{\langle x,y\rangle}}{\langle y,y\rangle}\,.$$
Dann ist
$$\alpha\langle x,y\rangle = -\,\frac{|\langle x,y\rangle|^2}{\langle y,y\rangle} \quad \text{und} \quad |\alpha|^2\langle y,y\rangle = \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle}\,.$$
Daher ist
$$0 < \langle x,x\rangle - \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle} - \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle} + \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle} = \langle x,x\rangle - \frac{|\langle x,y\rangle|^2}{\langle y,y\rangle}\,.$$
Also gilt
$$|\langle x,y\rangle|^2 < \langle x,x\rangle\,\langle y,y\rangle\,.$$
□

Definition 6.2.14. Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ mit einem inneren Produkt $\langle \cdot, \cdot \rangle$. Wir definieren die zugehörige Norm durch

$$\|x\| := \sqrt{\langle x, x \rangle}\,.$$

(Dabei wird die positive Quadratwurzel genommen.)
Diese Norm wird auch oft als die von $\langle \cdot, \cdot \rangle$ induzierte Norm bezeichnet.

Satz 6.2.15. *Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ mit einem inneren Produkt $\langle \cdot, \cdot \rangle$, und sei $\| \cdot \|$ definiert wie in Definition 6.2.14. Dann ist $\| \cdot \|$ eine Norm.*

Beweis. Wir führen den Beweis wieder für komplexe Vektorräume, der reelle Fall geht ebenso. Wir müssen die drei Punkte von Definition 6.2.1 nachweisen.
Punkt (1): Wegen der Bilinearität bzw. Sesquilinearität ist immer

$$\langle 0, 0 \rangle = \langle 0 \cdot 0, 0 \cdot 0 \rangle = 0 \langle 0, 0 \rangle = 0.$$

Also ist $\|0\| = 0$. Wegen der positiven Definitheit des inneren Produktes gilt $\langle x, x \rangle > 0$ für alle $x \in V \setminus \{0\}$. Damit gibt es auch eine positive Quadratwurzel von $\langle x, x \rangle$, und $\|x\| > 0$.
Punkt (2): Es ist für alle $\lambda \in \mathbb{C}$, $x \in V$:

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{|\lambda|^2 \langle x, x \rangle} = |\lambda| \, \|x\|\,.$$

Punkt (3) Seien $x, y \in V$. In der folgenden Rechnung verwenden wir die Cauchy-Schwarz'sche Ungleichung Satz 6.2.13.

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\
&= \langle x, x \rangle + \langle x, y \rangle + \overline{\langle x, y \rangle} + \langle y, y \rangle \\
&= \langle x, x \rangle + 2\Re(\langle x, y \rangle) + \langle y, y \rangle \\
&\leq \langle x, x \rangle + 2\sqrt{\langle x, x \rangle \langle y, y \rangle} + \langle y, y \rangle \\
&= (\sqrt{\langle x, x \rangle} + \sqrt{\langle y, y \rangle})^2 = (\|x\| + \|y\|)^2\,.
\end{aligned}$$

□

Beispiel 6.2.16. Die euklidische Norm $\| \cdot \|_2$ in $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ gehört zum inneren Produkt $\langle x, y \rangle = x^* y$.

Beispiel 6.2.17. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Auf dem Raum $\mathcal{C}([0, 1], \mathbb{K})$ der stetigen Funktionen von $[0, 1]$ nach $\mathbb{K}$ definieren wir das innere Produkt

$$\langle f, g \rangle := \int_0^1 \overline{f(t)} g(t)\, dt\,.$$

Die dazugehörige Norm ist

$$\|f\|_{L^2} = \left(\int_0^1 |f(t)|^2\, dt \right)^{1/2}.$$

Lemma 6.2.18 (Parallelogrammgleichung)**.** *Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ mit einem inneren Produkt $\langle \cdot, \cdot \rangle$ und zugehöriger Norm $\| \cdot \|$. Dann gilt für alle $x, y \in V$:*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2\,.$$

In Wirklichkeit ist die Parallelogrammgleichung auch eine hinreichende Bedingung, dass eine Norm von einem inneren Produkt stammt.

Beweis.

$$\begin{aligned} & \|x+y\|^2 + \|x-y\|^2 = \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\ = \;& \langle x,x\rangle + \langle y,x\rangle + \langle x,y\rangle + \langle y,y\rangle + \langle x,x\rangle - \langle y,x\rangle - \langle x,y\rangle + \langle y,y\rangle \\ = \;& 2\langle x,x\rangle + 2\langle y,y\rangle = 2\|x\|^2 + 2\|y\|^2 . \end{aligned}$$

□

Beispiel 6.2.19. Die folgenden Normen gehören zu keinen inneren Produkten, weil sie die Parallelogrammgleichung (Lemma 6.2.18) nicht erfüllen:

Die Normen $\|\cdot\|_p$ in $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ für $p \neq 2$,
die Normen $\|\cdot\|_{L^\infty}$ und $\|\cdot\|_{L^p}$ für $p \neq 2$ auf $\mathcal{C}([0,1],\mathbb{R})$ oder $\mathcal{C}([0,1],\mathbb{C})$.

6.2.3. *Norm und Grenzwert.*

Die Norm bringt nicht nur den Längenbegriff in die Geometrie der Vektorräume, sie ist vor allem der Schlüssel zur Analysis (Grenzwertrechnung) in Vektorräumen. Wir gehen davon aus, dass Sie wissen, was der Grenzwert für reelle Zahlen bedeutet:

Definition 6.2.20. Sei $(\alpha_n)_{n=1}^\infty$ eine Folge in $\mathbb{R}$ und sei $\beta \in \mathbb{R}$. Wir sagen, α_n konvergiert gegen β oder

$$\lim_{n\to\infty} \alpha_n = \beta$$

genau dann, wenn

$$(\forall \epsilon > 0)\,(\exists n_0 \in \mathbb{N})\,(\forall n > n_0)\,|\alpha_n - \beta| < \epsilon\,.$$

Anders ausgedrückt, wenn ein Abstand ϵ vorgegeben ist, lässt sich immer ein n_0 finden, sodass ab $n = n_0$ alle α_n von β einen kleineren Abstand als ϵ haben.

Im normierten Vektorraum wird der Abstand zwischen zwei Vektoren x und y durch die Norm $\|y - x\|$ gegeben. Entsprechend liest sich die Bedingung der Konvergenz:

Definition 6.2.21. Sei $(V, \|\cdot\|)$ ein normierter Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$. Sei (x_n) eine Folge in V und $y \in V$. Wir sagen, x_n konvergiert gegen y (bezüglich der Norm $\|\cdot\|$), oder

$$\lim_{n\to\infty} x_n = y$$

genau dann, wenn gilt:

$$\lim_{n\to\infty} \|x_n - y\| = 0.$$

Bemerkung 6.2.22. Anders ausgedrückt: Eine Folge von Vektoren x_n konvergiert genau dann gegen y, wenn gilt

$$(\forall \epsilon > 0)\,(\exists n_0 \in \mathbb{N})\,(\forall n > n_0)\,\|x_n - y\| < \epsilon\,.$$

Das folgende Beispiel zeigt, dass es auch auf die Norm ankommt, ob eine Folge konvergiert:

Beispiel 6.2.23. Im Raum $\mathcal{C}([0,1],\mathbb{R})$ der stetigen Funktionen von $[0,1]$ nach $\mathbb{R}$ betrachten wir die beiden Normen $\|\cdot\|_{L^\infty}$ und $\|\cdot\|_{L^1}$ wie in Beispiel 6.2.7. Wir betrachten die Funktionen

$$f_n(t) = \begin{cases} 2nt & \text{falls } t \in [0, 1/(2n)), \\ 2(1-nt) & \text{falls } t \in [1/(2n), 1/n), \\ 0 & \text{falls } t \in [1/n, 1]. \end{cases}$$

Es ist für alle $n \in \mathbb{N}$

$$\|f_n\|_{L^\infty} = \max_{t\in[0,1]} |f_n(t)| = 1,$$

daher konvergiert f_n nicht gegen Null bezüglich $\|\cdot\|_{L^\infty}$. Andererseits ist

$$\|f_n\|_{L^1} = \int_0^1 f_n(t)\,dt = \frac{1}{2n} \to 0,$$

also konvergiert f_n gegen Null bezüglich $\|\cdot\|_{L^1}$.

Überlegen Sie selbst dass $\|f_n - f_{2n}\|_{L^\infty} = 1$, sodass die Folge (f_n) bezüglich $\|\cdot\|_{L^\infty}$ überhaupt keinen Grenzwert besitzen kann. Überlegen Sie andererseits, dass für alle $t \in [0,1]$ gilt $f_n(t) \to 0$.

Definition 6.2.24. Sei V ein Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$ mit zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$.

1) Die Norm $\|\cdot\|_1$ heißt stärker als die Norm $\|\cdot\|_2$, wenn für alle Folgen (x_n) in V gilt:
$$\lim_{n\to\infty} \|x_n\|_1 = 0 \;\Rightarrow\; \lim_{n\to\infty} \|x_n\|_2 = 0.$$
2) Die beiden Normen heißen äquivalent, wenn jede stärker als die andere ist, d.h., für alle Folgen (x_n) in V gilt:
$$\lim_{n\to\infty} \|x_n\|_1 = 0 \;\Leftrightarrow\; \lim_{n\to\infty} \|x_n\|_2 = 0.$$

"Wenn eine Vektorfolge bezüglich der stärkeren Norm konvergiert, dann konvergiert sie erst recht bezüglich der schwächeren Norm."

Satz 6.2.25. *Sei V ein Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$ mit zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$. Dann gilt:*
Die Norm $\|\cdot\|_1$ ist genau dann stärker als die Norm $\|\cdot\|_2$, wenn gilt:

$$(\exists M > 0)\,(\forall x \in V)\,\|x\|_2 \le M\|x\|_1\,.$$

Die beiden Normen sind genau dann äquivalent, wenn gilt:

$$(\exists m, M > 0)\,(\forall x \in V)\,m\|x\|_1 \le \|x\|_2 \le M\|x\|_1\,.$$

Beweis. Sei zunächst $\|x\|_2 \le M\|x\|_1$ für alle $x \in V$. Konvergiert die Folge (x_n) gegen Null bezüglich $\|\cdot\|_1$, so folgt

$$\|x_n\|_2 \le M\|x_n\|_1 \to 0,$$

also konvergiert (x_n) auch bezüglich $\|\cdot\|_2$ gegen Null.
Angenommen, es gibt nun kein $M > 0$, sodass für alle $x \in V$ gilt $\|x\|_2 \le M\|x\|_1$. Insbesondere gibt es also für jedes $M = n^2$ einen Vektor $x_n \in V$ mit

$$\|x_n\|_2 > n^2\|x_n\|_1\,.$$

Wir setzen nun

$$y_n = \frac{1}{n\|x_n\|_1}\,x_n.$$

Dann ist

$$\|y_n\|_1 = \frac{1}{n}, \quad \|y_n\|_2 = \frac{\|x_n\|_2}{n\|x_n\|_1} > n\,.$$

Also geht $\|y_n\|_1$ gegen Null, aber $\|y_n\|_2$ geht gegen unendlich. Daher ist dann $\|\cdot\|_1$ nicht stärker als $\|\cdot\|_2$. □

Beispiel 6.2.26. In $\mathcal{C}([0,1],\mathbb{R})$ ist $\|\cdot\|_{L^\infty}$ stärker als $\|\cdot\|_{L^1}$.

Beweis. Sei $f \in \mathcal{C}([0,1],\mathbb{R})$. Es ist

$$\begin{aligned} \|f\|_{L^\infty} &= \max_{s\in[0,1]} |f(s)|\,, \\ \|f\|_{L^1} &= \int_0^1 |f(t)|\,dt \le \int_0^1 \|f\|_{L^\infty}\,dt = \|f\|_{L^\infty}\,. \end{aligned}$$

□

In endlichdimensionalen Vektorräumen ist der Konvergenzbegriff unabhängig von der Wahl der Norm:

Satz 6.2.27 (Normenäquivalenzsatz)**.** *Sei V ein endlichdimensionaler Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$. Dann sind alle Normen auf V zueinander äquivalent.*

Wir geben hier keinen Beweis. Ein wichtiges Werkzeug ist der Begriff der Kompaktheit aus der Analysis. Es ist aber leicht zu zeigen, dass in $\mathbb{R}^n$ die Normen $\|\cdot\|_p$ aus Beispiel 6.2.5 für alle $p \in [1, \infty)$ zur Norm $\|\cdot\|_\infty$ äquivalent sind. Versuchen Sie es!

6.3. **Orthogonale Projektion.**

Wir beginnen nun die Theorie der Räume mit innerem Produkt. Viele Beweise laufen in Vektorräumen über $\mathbb{R}$ genauso wie in Vektorräumen über $\mathbb{C}$. Wir geben in diesen Fällen nur den Beweis für den komplexen Fall. Den Beweis für den reellen Fall erhält man, indem man die komplexe Konjugation weglässt. Für jede reelle Zahl α ist ja $\overline{\alpha} = \alpha$.

6.3.1. *Orthogonalität und orthogonales Komplement.*

Definition 6.3.1. Sei V ein Vektorraum über $\mathbb{C}$ oder $\mathbb{R}$ mit innerem Produkt $\langle\cdot,\cdot\rangle$. Seien $a, b \in V$, $U, W \subset V$ nichtleere Teilmengen. Wir sagen:

1) a steht orthogonal auf b (Schreibweise: $a \perp b$), wenn $\langle a, b\rangle = 0$.
2) a steht orthogonal auf U, wenn $(\forall u \in U)\langle a, u\rangle = 0$.
3) U steht orthogonal auf W, wenn $(\forall u \in U)\,(\forall w \in W)\,\langle u, w\rangle = 0$.
4) Das orthogonale Komplement von U ist
   $U^\perp = \{x \in V \mid (\forall u \in U)\,\langle x, u\rangle = 0\}$.

Diese Definition ist wörtlich dieselbe wie Definition 6.1.15, die wir aber nur für das euklidische innere Produkt in $\mathbb{R}^n$ formuliert hatten.
Die Bezeichnung orthogonales Komplement $U^\perp$ ist allgemein üblich für den Fall, dass U ein Unterraum von V ist. Aber es ist auch manchmal praktisch, für andere Teilmengen den Begriff des orthogonalen Komplementes bereit zu haben.

Bemerkung 6.3.2. Orthogonalität von Vektoren ist eine symmetrische Relation, d.h.

$$a \perp b \Leftrightarrow b \perp a.$$

Beweis.

$$\langle a, b\rangle = 0 \Leftrightarrow \langle b, a\rangle = \overline{\langle a, b\rangle} = 0.$$

□

Satz 6.3.3. *Sei V ein Vektorraum über $\mathbb{C}$ oder $\mathbb{R}$ mit innerem Produkt $\langle\cdot,\cdot\rangle$ und dazugehöriger Norm $\|\cdot\|$. Seien $a, b \in V$. Dann gilt:*

$$a \perp b \;\Rightarrow\; \|a + b\|^2 = \|a\|^2 + \|b\|^2 .$$

In reellen Vektorräumen gilt auch die Umkehrung.

Beweis.

$$\begin{aligned} & \|a + b\|^2 = \langle a + b, a + b\rangle = \langle a, a\rangle + \langle a, b\rangle + \langle b, a\rangle + \langle b, b\rangle \\ = \;& \|a\|^2 + \langle a, b\rangle + \overline{\langle a, b\rangle} + \|b\|^2 = \|a\|^2 + \|b\|^2 + 2\Re(\langle a, b\rangle) . \end{aligned}$$

□

Satz 6.3.4. *Sei V ein Vektorraum mit innerem Produkt $\langle\cdot,\cdot\rangle$ über $\mathbb{R}$ oder $\mathbb{C}$. Sei $U \subset V$ eine nichtleere Teilmenge. Dann gilt:*

1) *$U^\perp$ ist ein Unterraum von V.*
2) *Ist $W \subset U$, so ist $U^\perp \subset W^\perp$.*

3) *Für das Komplement des Komplementes einer Menge gilt* $U \subset (U^\perp)^\perp$.
4) *Das Komplement der linearen Hülle von* U *ist* $(\mathcal{L}(U))^\perp = U^\perp$.
5) $\{0\}^\perp = V$.
6) $U \cap U^\perp \subset \{0\}$.
7) $V^\perp = \{0\}$.

Beweis. Punkt (1): Da $\langle 0, u\rangle = 0$ für alle $u \in U$, ist jedenfalls $0 \in U^\perp$, und damit ist $U^\perp \neq \emptyset$.
Seien nun $x, y \in U^\perp$, $\lambda, \mu \in \mathbb{C}$. Es ist für alle $u \in U$

$$\langle \lambda x + \mu y, u\rangle = \overline{\lambda}\langle x, u\rangle + \overline{\mu}\langle y, u\rangle = \overline{\lambda}\,0 + \overline{\mu}\,0 = 0.$$

Also ist $\lambda x + \mu y \in U^\perp$.

Punkt (2): Sei $W \subset U$ und $x \in U^\perp$. Dann steht x orthogonal auf alle Vektoren $u \in U$, und da $W \subset U$, steht x insbesondere orthogonal auf alle Vektoren aus W. Folglich ist $x \in W^\perp$.

Punkt (3): Sei $u \in U$. Dann gilt (nach Definition des Komplementes) für alle $x \in U^\perp$

$$\langle u, x\rangle = \overline{\langle x, u\rangle} = 0.$$

Daher steht u orthogonal auf ganz $U^\perp$, d.h., $u \in (U^\perp)^\perp$.

Punkt (4): Da $U \subset \mathcal{L}(U)$, gilt wegen Punkt (2) jedenfalls $(\mathcal{L}(U))^\perp \subset U^\perp$. Andererseits gilt wegen Punkt (3) $U \subset (U^\perp)^\perp$, und $(U^\perp)^\perp$ ist wegen Punkt (1) ein Unterraum von V. Daher ist auch $\mathcal{L}(U) \subset (U^\perp)^\perp$. Anders ausgedrückt, jedes Element aus $U^\perp$ steht orthogonal auf jedes Element aus $\mathcal{L}(U)$, d.h., $U^\perp \subset (\mathcal{L}(U))^\perp$.

Punkt (5): Trivial.

Punkt (6): Sei $u \in U \cap U^\perp$. Dann ist also $\langle u, u\rangle = 0$, das geht aber wegen der Definition eines inneren Produktes nur für $u = 0$.

Punkt (7): Wegen Punkt (6) enthält $V^\perp = V \cap V^\perp$ höchstens den Nullvektor. Da aber $V^\perp$ ein Unterraum von V ist, enthält er jedenfalls den Nullvektor. □

6.3.2. *Orthogonale Projektion.*

Definition 6.3.5. Sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Sei $U \subset V$ eine nichtleere Teilmenge. Die Menge U heißt konvex, wenn gilt:

$$(\forall \lambda \in [0,1])\,(\forall x, y \in U)\,\lambda x + (1-\lambda)y \in U.$$

Anschaulich gesprochen: Eine Menge U ist konvex, wenn die Verbindungsstrecke von je 2 Punkten aus U zur Gänze in U verläuft.

Satz 6.3.6. *Sei* V *ein Vektorraum über* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ *mit einem inneren Produkt* $\langle \cdot, \cdot\rangle$ *und der dazugehörigen Norm* $\|\cdot\|$. *Sei* U *eine konvexe Teilmenge von* V. *Sei* $x \in V$ *und* $u \in U$. *Dann sind die folgenden Aussagen äquivalent:*

1) *Für alle* $y \in U \setminus \{u\}$ *gilt* $\|u - x\| \leq \|y - x\|$.
2) *Für alle* $y \in U \setminus \{u\}$ *gilt* $\Re(\langle y - u, x - u\rangle) \leq 0$.
3) *Für alle* $y \in U \setminus \{u\}$ *gilt* $\|u - x\| < \|y - x\|$.

Ist U *ein Unterraum*[23] *von* V, *so sind diese Aussagen auch äquivalent zu:*

4) *Für alle* $y \in U \setminus \{u\}$ *gilt* $y - u \perp x - u$.

[23]Unterräume sind natürlich immer konvex.

Beweis. (1) ⇒ (2): Wir nehmen an, dass (1) gilt. Sei $y \in U$, $y \neq u$. Sei $\epsilon \in (0,1)$ beliebig. Wegen der Konvexität von U ist dann auch $z := u + \epsilon(y-u) \in U$. Nun gilt wegen (1):

$$\begin{aligned} 0 &\leq \|z-x\|^2 - \|u-x\|^2 = \|u+\epsilon(y-u)-x\|^2 - \|u-x\|^2 \\ &= \langle (u-x)+\epsilon(y-u), (u-x)+\epsilon(y-u)\rangle - \|u-x\|^2 \\ &= \|u-x\|^2 + \epsilon^2\|y-u\|^2 + \epsilon\langle y-u, u-x\rangle + \epsilon\langle u-x, y-u\rangle - \|u-x\|^2 \\ &= \epsilon^2\|y-u\|^2 + 2\epsilon\Re(\langle y-u, u-x\rangle). \end{aligned}$$

Wir dividieren diese Ungleichung durch ϵ und erhalten für alle $\epsilon \in (0,1)$

$$-2\Re(\langle y-u, u-x\rangle) \leq \epsilon\|y-u\|^2.$$

Lassen wir nun ϵ gegen Null gehen, erhalten wir Punkt (2).

(2) ⇒ (3): Angenommen, Punkt(2) gilt. Sei $y \in U$, $y \neq u$. Dann ist

$$\begin{aligned} \|y-x\|^2 &= \|(u-x)-(u-y)\|^2 \\ &= \|u-x\|^2 - 2\Re(\langle u-x, u-y\rangle) + \|u-y\|^2 \\ &\geq \|u-x\|^2 + \|u-y\|^2 > \|u-x\|^2. \end{aligned}$$

(3) ⇒ (1): Trivial.

(2) ⇒ (4): Wir nehmen nun an, dass U ein Unterraum ist. Ist $y \in U$ so ist wegen Punkt (2) $\Re(\langle x-u, y-u\rangle) \leq 0$. Weil U ein Unterraum ist, ist aber auch $z = u-(y-u) \in U$, und es ist

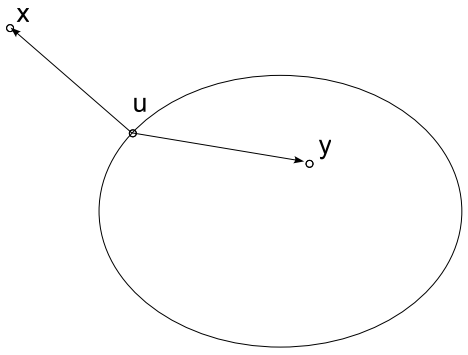
$$0 \geq \Re(\langle x-u, z-u)\rangle) = -\Re(\langle x-u, y-u\rangle).$$

Daher ist $\Re(\langle x-u, y-u\rangle) = 0$. Ist V ein Vektorraum über $\mathbb{C}$, so ist auch $z = u+i(y-u) \in U$, und damit ist

$$0 = \Re(\langle x-u, z-u\rangle) = \Re(\langle x-u, i(y-u)\rangle) = -i\Im(\langle x-u, y-u\rangle).$$

(4) ⇒ (2): Trivial. □

Anschaulich gesprochen: Der Punkt $u \in U$ liegt genau dann von allen Punkten aus U am nächsten an x, wenn für jedes $y \in U$ der Winkel zwischen $x-u$ und $y-u$ ein stumpfer Winkel ist.



Korollar 6.3.7. *Sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit einem inneren Produkt $\langle\cdot,\cdot\rangle$ und der dazugehörigen Norm $\|\cdot\|$. Sei U eine konvexe Teilmenge von V. Sei $x \in V$. Dann gibt es höchstens einen Punkt $u \in U$, sodass für alle $y \in U$ gilt:*

$$\|u-x\| \leq \|y-x\|.$$

Insbesondere gilt: Ist U ein Unterraum von V, so gibt es höchstens ein $u \in U$ so, dass $x - u \in U^\perp$.

Beweis. Seien u_1, u_2 zwei solche Punkte. Wegen Satz 6.3.6(3) folgt dann für alle $y \in U$:

$$\|u_1 - x\| < \|y - x\| \quad \text{und} \quad \|u_2 - x\| < \|y - x\|.$$

Setzt man $y = u_1$ bzw. $y = u_2$, so erhält man

$$\|u_1 - x\| < \|u_2 - x\| \quad \text{und} \quad \|u_2 - x\| < \|u_1 - x\|,$$

das ist aber ein Widerspruch.
Ist U ein Unterraum, so gilt $u \in U$, $x - u \in U^\perp$ genau dann, wenn $u \in U$ der nächste Punkt an x ist. Und wenn ein solcher existiert, ist er (wie eben bewiesen) eindeutig bestimmt. □

Definition 6.3.8. Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Sei U ein Unterraum von V, sodass $U + U^\perp = V$. Wir definieren eine Abbildung $\pi_U : V \to V$ so dass für jedes $x \in V$ das Bild $\pi_U(x) \in U$ und $x - \pi_U(x) \in U^\perp$.
Dann heißt die Abbildung π_U die orthogonale Projektion von V auf U.

Rechtfertigung der Definition: Die Bedingung $U + U^\perp = V$ sagt genau: Zu jedem $x \in V$ existiert mindestens ein $u \in U$ so, dass $x - u \in U^\perp$. Wegen Korollar 6.3.7 ist dieses u eindeutig bestimmt.

Satz 6.3.9. *Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Sei U ein Unterraum von V so, dass $U + U^\perp = V$. Sei π_U die orthogonale Projektion von V auf U. Dann gilt*

1) *Sei $x \in V$. Dann ist $x \in U$ genau dann wenn $\pi_U(x) = x$.*
2) *Sei $x \in V$. Dann ist $x \in U^\perp$ genau dann wenn $\pi_U(x) = 0$.*
3) *Die Abbildung π_U ist idempotent, d.h. $\pi_U \circ \pi_U = \pi_U$.*
4) *Die Abbildung $\mathrm{id}_V - \pi_U$ ist die orthogonale Projektion von V auf $U^\perp$.*
5) *Die Abbildung π_U ist linear.*
6) *Für alle $x, y \in V$ gilt*
$$\langle x, \pi_U(y)\rangle = \langle \pi_U(x), y\rangle = \langle \pi_U(x), \pi_U(y)\rangle.$$
7) *Für alle $x \in V$ gilt: $\|\pi_U(x)\| \le \|x\|$. Gleichheit gilt genau dann, wenn $x \in U$.*
8) *Der Vektorraum V ist die* direkte *Summe $U + U^\perp$.*
9) $U = (U^\perp)^\perp$.

Beweis. Punkt (1): Ist $x \in U$, so ist $x - x = 0 \in U^\perp$, und daher ist $\pi_U(x) = x$. Umgekehrt gilt für alle $x \in V$, dass $\pi_U(x) \in U$. Ist also $x = \pi_U(x)$, so ist $x \in U$.

Punkt (2): Ist $x \in U^\perp$, so ist $x = 0 + x$ mit $0 \in U$ und $x \in U^\perp$. Also ist $\pi_U(x) = 0$. Andererseits ist $x - \pi_U(x)$ immer in $U^\perp$. Ist also $\pi_U(x) = 0$, so ist $x = x - \pi_U(x) \in U^\perp$.

Punkt (3): Für alle $x \in V$ ist $\pi_U(x) \in U$. Daher gilt nach Punkt (1) $[\pi_U \circ \pi_U](x) = \pi_U(\pi_U(x)) = \pi_U(x)$.

Punkt (4): Ist $x \in V$, so ist $x - \pi_U(x) \in U^\perp$ und $\pi_U(x) \in U \subset (U^\perp)^\perp$. Daher ist $x - \pi_U(x) = \pi_{U^\perp}(x)$.

Punkt (5): Seien $x, y \in V$, $\lambda, \mu \in \mathbb{K}$. Zur Abkürzung schreiben wir:

$$\pi_U(x) = u_x \in U,\ \pi_U(y) = u_y \in U,\ x - \pi_U(x) = v_x \in U^\perp,\ y - \pi_U(y) = v_y \in U^\perp.$$

Es ist dann

$$\begin{aligned}\lambda x + \mu y &= \lambda(u_x + v_x) + \mu(u_y + v_y) \\ &= (\lambda u_x + \mu u_y) + (\lambda v_x + \mu v_y)\,.\end{aligned}$$

Weil U und $U^\perp$ Unterräume von V sind, gilt

$$\lambda u_x + \mu u_y \in U \quad \lambda v_x + \mu v_y \in U^\perp\,.$$

Damit ist aber

$$\pi_U(\lambda x + \mu y) = \lambda u_x + \mu u_y = \lambda \pi_U(x) + \mu \pi_U(y)\,.$$

Punkt (6): Wir verwenden, dass $x - \pi_U(x)$ und $y - \pi_U(y)$ orthogonal auf U stehen. Es ist

$$\begin{aligned}&\langle x, \pi_U(y)\rangle = \langle x, \pi_U(y)\rangle + \langle \pi_U(x) - x, \pi_U(y)\rangle \\ = \;& \langle \pi_U(x), \pi_U(y)\rangle = \langle \pi_U(x), \pi_U(y)\rangle + \langle \pi_U(x), y - \pi_U(y)\rangle \\ = \;& \langle \pi_U(x), y\rangle\,.\end{aligned}$$

Punkt (7): Weil $\pi_U(x) \perp (x - \pi_U(x))$, gilt nach Pythagoras

$$\|x\|^2 = \|\pi_U(x)\|^2 + \|x - \pi_U(x)\|^2 \geq \|\pi_U(x)\|^2\,.$$

Gleichheit gilt offensichtlich genau dann, wenn $x - \pi_U(x) = 0$, das heißt wegen Punkt (1) dasselbe wie $x \in U$.

Punkt (8): Nach Voraussetzung ist $V = U + U^\perp$. Wegen Satz 6.3.4(6) ist $U \cap U^\perp = \{0\}$.

Punkt (9): Bekanntlich ist $U \subset (U^\perp)^\perp$. Sei $x \in (U^\perp)^\perp$. Es ist $x - \pi_U(x) \in U^\perp$ und orthogonal zu $\pi_U(x)$, daher ist

$$0 = \langle x, x - \pi_U(x)\rangle = \langle x - \pi_U(x), x - \pi_U(x)\rangle = \|x - \pi_U(x)\|^2\,.$$

Also ist $x = \pi_U(x)$, und nach Punkt (1) ist also $x \in U$. Daher ist $(U^\perp)^\perp \subset U$. □

6.3.3. *Orthogonale Projektion auf endlichdimensionale Unterräume.*

Definition 6.3.10. Sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit einem inneren Produkt $\langle\cdot,\cdot\rangle$ und der dazugehörigen Norm $\|\cdot\|$. Sei $(b_1, \cdots, b_n)$ ein n-Tupel von Vektoren aus V. Wir definieren die Gramsche Matrix

$$\mathrm{Gram}(b_1, \cdots, b_n) := \begin{pmatrix} \langle b_1, b_1\rangle & \cdots & \langle b_1, b_n\rangle \\ \vdots & & \vdots \\ \langle b_n, b_1\rangle & \cdots & \langle b_n, b_n\rangle \end{pmatrix}.$$

Bemerkung 6.3.11. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. In $\mathbb{K}^n$ betrachten wir wie üblich das innere Produkt $\langle x, y\rangle = x^* y$. Sei $(b_1, \cdots, b_m)$ ein m-Tupel von Vektoren aus $\mathbb{K}^n$. Wir bilden eine Matrix $B \in \mathbb{K}^{n\times m}$, deren Spalten die jeweiligen b_i sind. Dann ist

$$\mathrm{Gram}(b_1, \cdots, b_m) = B^* B.$$

Satz 6.3.12. *Sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit einem inneren Produkt $\langle\cdot,\cdot\rangle$ und der dazugehörigen Norm $\|\cdot\|$. Sei $(b_1, \cdots, b_n)$ ein n-Tupel von Vektoren aus V, und $G = \mathrm{Gram}(b_1, \cdots, b_n)$ die zugehörige Gramsche Matrix. Dann gilt:*

1) *G ist hermitesch.*

2) *Für alle* $x = (\xi_1, \cdots, \xi_n)^T \in \mathbb{K}^n$ *ist*

$$\begin{aligned} Gx &= \begin{pmatrix} \langle b_1, \sum_{j=1}^n \xi_j b_j \rangle \\ \vdots \\ \langle b_n, \sum_{j=1}^n \xi_j b_j \rangle \end{pmatrix}, \\ x^*Gx &= \| \sum_{j=1}^n \xi_j b_j \|^2 . \end{aligned}$$

3) *G ist positiv semidefinit (d.h. für alle $x \in \mathbb{K}^n$ ist $x^*Gx \geq 0$).*
4) *G ist genau dann positiv definit, wenn das System $(b_1, \cdots, b_n)$ linear unabhängig ist. (Positiv definit heißt, für alle $x \in \mathbb{K}^n \setminus \{0\}$ ist $x^*Gx > 0$.)*
5) *Sei $x = (\xi_1, \cdots, \xi_n)^T \in \mathbb{K}^n$. Dann gilt:*

$$x \in \ker(G) \Leftrightarrow \sum_{i=1}^n \xi_i b_i = 0 .$$

6) *G ist genau dann regulär, wenn das System $(b_1, \cdots, b_n)$ linear unabhängig ist.*

Beweis. Punkt (1): Das folgt aus der Tatsache, dass $\langle b_j, b_i \rangle = \overline{\langle b_i, b_j \rangle}$.
Punkt (2): Es ist

$$\begin{aligned} Gx &= \begin{pmatrix} \sum_{j=1}^n \langle b_1, b_j \rangle \xi_j \\ \vdots \\ \sum_{j=1}^n \langle b_n, b_j \rangle \xi_j \end{pmatrix} = \begin{pmatrix} \langle b_1, \sum_{j=1}^n \xi_j b_j \rangle \\ \vdots \\ \langle b_n, \sum_{j=1}^n \xi_j b_j \rangle \end{pmatrix}, \\ x^*Gx &= \sum_{i=1}^n \overline{\xi}_i \langle b_i, \sum_{j=1}^n \xi_j b_j \rangle = \langle \sum_{i=1}^n \xi_i b_i, \sum_{j=1}^n \xi_j b_j \rangle = \| \sum_{i=1}^n \xi_i b_i \|^2 . \end{aligned}$$

Punkt (3): Direkte Folge von Punkt (2).
Punkt (4): Sei das n-Tupel $(b_1, \cdots, b_n)$ linear unabhängig. Dann gilt für alle Vektoren $x = (\xi_1, \cdots, \xi_n)^T \neq 0$ dass $\sum_{i=1}^n \xi_i b_i \neq 0$. Folglich ist auch

$$x^*Gx = \| \sum_{i=1}^n \xi_i b_i \|^2 > 0.$$

Also ist G positiv definit.
Sei umgekehrt G positiv definit und $\sum_{i=1}^n \xi_i b_i = 0$ für einen Vektor $x = (\xi_1, \cdots, \xi_n)^T$. Dann ist also

$$x^*Gx = \| \sum_{i=1}^n \xi_i b_i \|^2 = \|0\|^2 = 0,$$

und wegen der positiven Definitheit von G ist $x = 0$, also sind alle $\xi_i = 0$. Folglich ist $(b_1, \cdots, b_n)$ linear unabhängig.
Punkt (5): Sei $x \in \ker(G)$. Es ist dann

$$\| \sum_{i=1}^n \xi_i b_i \|^2 = x^*Gx = x^*0 = 0 ,$$

und daher ist $\sum_{i=1}^n \xi_i b_i = 0$.
Sei nun umgekehrt $\sum_{i=1}^n \xi_i b_i = 0$. Wegen Punkt (2) ist dann auch $Gx = 0$, also $x \in \ker(G)$.
Punkt (6): Es ist G genau dann regulär, wenn $\ker(G) = \{0\}$. Wegen Punkt (5) ist dies äquivalent dazu, dass die einzige Linearkombination $\sum_{i=1}^n \xi_i b_i$, welche Null ergibt, mit den Koeffizienten $\xi_1 = \cdots = \xi_n = 0$ gebildet wird, d.h. $(b_1, \cdots, b_n)$ sind linear unabhängig. □

Satz 6.3.13. *Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Sei U ein endlichdimensionaler Unterraum von V mit einer Basis $(b_1, \cdots, b_n)$. Sei $x \in V$.*

1) *Dann gibt es genau ein $u \in U$ so, dass $x - u \in U^\perp$.*
2) *Ist $G = \mathrm{Gram}(b_1, \cdots, b_n)$ die Gramsche Matrix des Systems $(b_1, \cdots, b_n)$, so schreibt sich*

$$u = \sum_{j=1}^{n} \eta_j b_j,$$

wobei der Vektor $(\eta_1, \cdots, \eta_n)^T$ das folgende Gleichungssystem löst:

$$G \cdot \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} = \begin{pmatrix} \langle b_1, x\rangle \\ \vdots \\ \langle b_n, x\rangle \end{pmatrix}. \tag{6.3.1}$$

Beweis. Wir wissen schon aus Korollar 6.3.7, dass es höchstens eine Zerlegung $x = u + w$ mit $u \in U$, $w \in U^\perp$ geben kann. Wir zeigen nun die Existenz von u:
Weil $(b_1, \cdots, b_n)$ ein linear unabhängiges System ist, ist die zugehörige Gramsche Matrix regulär. Daher ist das Gleichungssystem (6.3.1) eindeutig lösbar. Wir setzen nun $u = \sum_{j=1}^n \eta_j b_j$. Da alle b_j in U enthalten sind, ist $u \in U$. Wir zeigen nun: $x - u \in U^\perp$:
Wegen Satz 6.3.12 ist

$$\begin{pmatrix} \langle b_1, \sum_{j=1}^n \eta_j b_j\rangle \\ \vdots \\ \langle b_n, \sum_{j=1}^n \eta_j b_j\rangle \end{pmatrix} = G \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}.$$

Daher ergibt sich aus dem Gleichungssystem (6.3.1) für jedes $i = 1, \cdots, n$:

$$\langle b_i, x - u\rangle = \langle b_i, x\rangle - \langle b_i, \sum_{j=1}^{n} \eta_j b_j\rangle = 0\,.$$

Da $x-u$ auf alle Vektoren der Basis $(b_1, \cdots, b_n)$ orthogonal steht, ist $x-u \in U^\perp$. □

Bemerkung 6.3.14. In Satz 6.3.13 kann V ein unendlichdimensionaler Vektorraum sein. Ist aber U unendlichdimensional, so muss es nicht unbedingt einen nächsten Punkt geben. Mit Mitteln der Analysis (Vollständigkeit) kann man auch unendlichdimensionale Unterräume charakterisieren, in denen es zu jedem $x \in V$ einen nächsten Punkt gibt.

Satz 6.3.15. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. In $\mathbb{K}^n$ betrachten wir das übliche innere Produkt $\langle x, y\rangle = x^* y$. (Im reellen Fall setzen wir $x^* = x^T$.) Sei $f : \mathbb{K}^n \to \mathbb{K}^n$ eine lineare Abbildung und $P \in \mathbb{K}^{n\times n}$ die dazugehörige Matrix. Dann sind äquivalent:*

1) *Es gibt einen Unterraum $U \subset \mathbb{K}^n$ so, dass $f = \pi_U$.*
2) *P ist idempotent (d.h. $P^2 = P$) und hermitesch (d.h. $P^* = P$).*

In diesem Fall ist U der Spaltenraum von P und $U^\perp$ der Kern von P.

Beweis. (1) ⇒ (2): Sei also $f = \pi_U$ die orthogonale Projektion auf einen Unterraum U. Nach Satz 6.3.9 (3) gilt: $\pi_U \circ \pi_U = \pi_U$. Daher ist $P = P^2$. Seien nun e_i, e_j zwei Einheitsvektoren im $\mathbb{K}^n$, und $P = (\alpha_{i,j})_{i,j=1\cdots n}$. Es gilt (wegen Satz 6.3.9 (6)

$$\begin{aligned} \alpha_{i,j} &= e_i^* P e_j = \langle e_i, \pi_U(e_j)\rangle = \langle \pi_U(e_i), e_j\rangle \\ &= \overline{\langle e_j, \pi_U(e_i)\rangle} = \overline{e_j^* P e_i} = \overline{\alpha_{j,i}}\,. \end{aligned}$$

Es ist dann $U = \mathrm{rg}(\pi_U) = \mathrm{rg}(P)$ und $U^\perp = \ker(\pi_U) = \ker(P)$.

(2) ⇒ (1): Sei nun P idempotent und hermitesch. Sei U der Spaltenraum von P. Für jedes $x = Pv \in U$ ist dann wegen der Idempotenz $Px = PPv = Pv = x$. Sei nun $y \in \mathbb{K}^n$. Es gilt für alle $x \in U$ (unter Benützung der Idempotenz und Hermitizität):

$$x^*(y - Py) = (Px)^*(y - Py) = x^*P^*(y - Py) = x^*Py - x^*P^2y = 0\,.$$

Also ist $y - Py \in U^\perp$. Nun ist also $y = Py + (y - Py)$ mit $Py \in U$ und $y - Py \in U^\perp$. Daher ist $Py = \pi_U(y)$. □

Wir werden später von orthogonalen Matrizen reden. Vorsicht, die Matrix einer orthogonalen Projektion ist normalerweise *keine* orthogonale Matrix! (Einzige Ausnahme: Einheitsmatrix.)

6.3.4. *Beispiele:*

Beispiel 6.3.16. Wir betrachten den $\mathbb{R}^3$ mit der euklidischen Norm. Seien

$$b_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}\,, \quad b_2 = \begin{pmatrix}0\\1\\1\end{pmatrix}\,, \quad x = \begin{pmatrix}1\\2\\6\end{pmatrix}\,.$$

Sei U der von b_1, b_2 aufgespannte Unterraum des $\mathbb{R}^3$. Wir berechnen die orthogonale Projektion (bezüglich des üblichen inneren Produktes) von x auf U.

Lösung: Wir fassen die beiden Basisvektoren von U zu einer Matrix B zusammen und berechnen die Gramsche Matrix:

$$\begin{aligned} B &= \begin{pmatrix}1&0\\0&1\\1&1\end{pmatrix}\,, \\ \mathrm{Gram}(b_1, b_2) &= B^*B = \begin{pmatrix}1&0&1\\0&1&1\end{pmatrix}\begin{pmatrix}1&0\\0&1\\1&1\end{pmatrix} = \begin{pmatrix}2&1\\1&2\end{pmatrix}\,. \end{aligned}$$

Die rechte Seite für unser Gleichungssystem ist der Vektor

$$\begin{pmatrix}\langle b_1, x\rangle\\ \langle b_2, x\rangle\end{pmatrix} = \begin{pmatrix}7\\8\end{pmatrix}\,.$$

Wir lösen das Gleichungssystem

$$\begin{pmatrix}2&1\\1&2\end{pmatrix}\begin{pmatrix}\eta_1\\\eta_2\end{pmatrix} = \begin{pmatrix}7\\8\end{pmatrix}$$

mit der Cramerschen Regel:

$$\begin{aligned} \eta_1 &= \frac{\begin{vmatrix}7&1\\8&2\end{vmatrix}}{\begin{vmatrix}2&1\\1&2\end{vmatrix}} = \frac{6}{3} = 2\,, \\ \eta_2 &= \frac{\begin{vmatrix}2&7\\1&8\end{vmatrix}}{\begin{vmatrix}2&1\\1&2\end{vmatrix}} = \frac{9}{3} = 3\,. \end{aligned}$$

Daher ist

$$u = \pi_U(x) = 2\begin{pmatrix}1\\0\\1\end{pmatrix} + 3\begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}2\\3\\5\end{pmatrix}\,.$$

□

Überzeugen Sie sich zur Probe selbst, dass $x - u$ auf b_1 und b_2, und somit auf ganz U orthogonal steht.

Beispiel 6.3.17. Gegeben sei die Funktion $p(t) = t^3$. Finden Sie ein Polynom $q(x)$ vom Grad höchstens 2, sodass

$$\int_{-1}^{1} |p(t) - q(t)|^2 \, dt$$

möglichst klein wird.

Lösung: Auf dem Raum $\mathcal{C}([-1,1], \mathbb{R})$ der stetigen Funktionen von $[1,1]$ in $\mathbb{R}$ betrachten wir das innere Produkt

$$\langle u, v \rangle := \int_{-1}^{1} u(t)v(t) \, dt$$

und die dazugehörige Norm

$$\|u\|_{L^2} := \left(\int_{-1}^{1} |u(t)|^2 \, dt \right)^{1/2}.$$

Die Funktionen

$$f_0(t) = 1\,, \quad f_1(t) = t\,, \quad f_2(t) = t^2$$

spannen den Raum U der Polynome vom Grad kleiner als 3 auf. Wir suchen die orthogonale Projektion von p auf U.

Zur Aufstellung des Gleichungssystems brauchen wir zunächst folgende innere Produkte (mit $i, j \in \{0, 1, 2\}$):

$$\begin{aligned}
\langle f_i, f_j \rangle &= \int_{-1}^{1} t^{i+j} \, dt = \frac{1}{i+j+1} \left[1^{i+j+1} - (-1)^{i+j+1} \right] \\
&= \begin{cases} \frac{2}{i+j+1} & \text{wenn } \; i+j \text{ gerade,} \\ 0 & \text{wenn } \; i+j \text{ ungerade.} \end{cases} \\
\langle f_i, p \rangle &= \begin{cases} \frac{2}{i+4} & \text{wenn } \; i \text{ ungerade,} \\ 0 & \text{wenn } \; i \text{ gerade.} \end{cases}
\end{aligned}$$

Damit ist das Gleichungssystem

$$\begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix} \begin{pmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{2}{5} \\ 0 \end{pmatrix}.$$

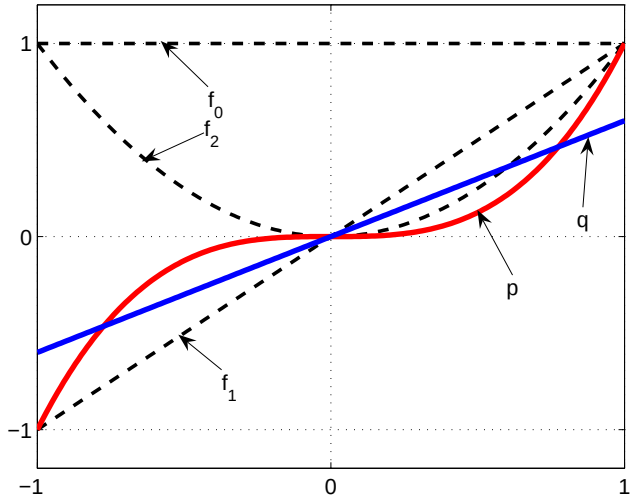
Die Lösung ist

$$\eta_0 = \eta_2 = 0\,, \; \eta_1 = \frac{3}{5}\,.$$

Das gesuchte Polynom ist

$$q(t) = \frac{3}{5}\, t.$$

Die folgende Grafik zeigt die drei Polynome f_0, f_1, f_2 (strichliert), die Funktion p und das Polynom p, welches q am besten annähert:



□

Wenn eine Funktion p durch Funktionen aus einer vorgegebenen Menge U möglichst gut angenähert werden soll, sprechen wir von einem Approximationsproblem. Ist das Kriterium für die Güte der Approximation, dass $\|p-u\|_{L^2}$ möglichst klein sein soll, sprechen wir von Approximation nach der Methode der kleinsten Quadrate (least squares). Steht für eine Least-Squares-Approximation ein Vektorraum U von Funktionen zur Verfügung, so führt das auf die Suche nach der orthogonalen Projektion von p auf U.

6.4. **Orthonormalsysteme.**

6.4.1. *Konstruktion von Orthonormalsystemen.*

Definition 6.4.1. Sei V ein Vektorraum über $\mathbb{K}=\mathbb{R}$ oder $\mathbb{K}=\mathbb{C}$ mit einem inneren Produkt $\langle\cdot,\cdot\rangle$. Sei $(a_1,\cdots,a_n)$ ein n-Tupel von Vektoren aus V. Dann heißt $(a_1,\cdots,a_n)$ ein Orthonormalsystem, wenn für alle $i,j\in\{1,\cdots,n\}$ gilt:

$$\langle a_i,a_j\rangle=\delta_{i,j}=\begin{cases}1 & \text{falls } i=j,\\ 0 & \text{falls } i\neq j.\end{cases}$$

Ist zugleich $(a_1,\cdots,a_n)$ eine Basis von V, so heißt $(a_1,\cdots,a_n)$ eine Orthonormalbasis von V.

Im Ortho*normal*system steht also jeder Vektor orthogonal auf jeden anderen Vektor, und jeder Vektor hat *Norm* $\|a_i\|=1$. Steht in einem Vektorsystem jeder Vektor orthogonal auf jeden anderen, und keiner ist der Nullvektor, so spricht man von einem Ortho*gonal*system.

Beispiel 6.4.2. Einige Beispiele von Orthonormalsystemen in $\mathbb{C}^2$ bzw. $\mathbb{C}^3$:

Die Einheitsbasen,

$$\left(\begin{pmatrix}\frac{1+i}{2}\\ \frac{1-i}{2}\end{pmatrix},\begin{pmatrix}\frac{1+i}{2}\\ \frac{-1+i}{2}\end{pmatrix}\right),$$

$$\left(\begin{pmatrix}\cos(\phi)\\ -\sin(\phi)\end{pmatrix},\begin{pmatrix}\sin(\phi)\\ \cos(\phi)\end{pmatrix}\right)\text{ mit }\phi\in[0,2\pi),$$

$$\left(\begin{pmatrix}2/3\\ 1/3\\ 2/3\end{pmatrix},\begin{pmatrix}-2/3\\ 2/3\\ 1/3\end{pmatrix},\begin{pmatrix}1/3\\ 2/3\\ -2/3\end{pmatrix}\right).$$

Satz 6.4.3. *Sei V ein Vektorraum über $\mathbb{K}\in\{\mathbb{R},\mathbb{C}\}$ mit innerem Produkt. Sei $(a_1,\cdots,a_n)$ ein Orthonormalsystem in V. Dann ist $(a_1,\cdots,a_n)$ linear unabhängig.*

Beweis. Sei $\sum_{i=1}^{n} \lambda_i a_i = 0$. Dann ist

$$0 = \langle a_j, \sum_{i=1}^{n} \lambda_i a_i \rangle = \sum_{i=1}^{n} \lambda_i \langle a_j, a_i \rangle = \sum_{i \neq j} 0\lambda_i + 1\lambda_j.$$

Da dies für alle j gilt, sind also alle $\lambda_j = 0$. □

Satz 6.4.4 (Orthonormalisierungsverfahren von Gram - Schmidt)**.** *Sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit innerem Produkt. Sei $(b_1, \cdots, b_n)$ ein linear unabhängiges Vektorsystem in V. Dann gibt es ein Orthonormalsystem $(a_1, \cdots, a_n)$ so, dass für alle $k = 1, \cdots, n$ gilt*

$$\mathcal{L}(\{a_1, \cdots, a_k\}) = \mathcal{L}(\{b_1, \cdots, b_k\}).$$

Die Vektoren a_i lassen sich rekursiv mit folgendem Algorithmus berechnen:

$$\begin{aligned}
\tilde{a}_1 &= b_1, \quad \textit{Nur für konsistente Schreibweise im Beweis.} \\
a_1 &= \frac{1}{\|\tilde{a}_1\|} \tilde{a}_1, \\
\tilde{a}_{k+1} &= b_{k+1} - \sum_{i=1}^{k} \langle a_i, b_{k+1} \rangle a_i, \\
a_{k+1} &= \frac{1}{\|\tilde{a}_{k+1}\|} \tilde{a}_{k+1}.
\end{aligned}$$

Beweis. Wir zeigen durch vollständige Induktion nach k:

1) $\tilde{a}_k \neq 0$,
2) $(a_1, \cdots, a_k)$ sind ein Orthonormalsystem,
3) $\mathcal{L}(\{a_1, \cdots, a_k\}) = \mathcal{L}(\{b_1, \cdots, b_k\})$.

Die Induktionsverankerung für $k = 1$ ist trivial. Sei nun der Satz gültig bis k. Dann ist

$$u := \sum_{i=1}^{k} \langle a_i, b_{k+1} \rangle a_i \in \mathcal{L}(\{a_1, \cdots, a_k\}) = \mathcal{L}(\{b_1, \cdots, b_k\}).$$

Andererseits gilt wegen der linearen Unabhängigkeit der $(b_1, \cdots, b_{k+1})$:

$$b_{k+1} \notin \mathcal{L}(\{b_1, \cdots, b_k\}).$$

Also ist $b_{k+1} \neq u$ und damit $\tilde{a}_{k+1} = b_{k+1} - u \neq 0$.
Da $u \in \mathcal{L}(\{b_1, \cdots, b_k\})$, ist $\tilde{a}_{k+1} \in \mathcal{L}(\{b_1, \cdots, b_{k+1}\})$. Ebenso zeigt man: $b_{k+1} \in \mathcal{L}(\{a_1, \cdots, a_k, \tilde{a}_{k+1}\}$. Es folgt, dass die beiden linearen Hüllen gleich sind.
Für jedes $j \leq k$ gilt

$$\langle a_j, \tilde{a}_{k+1} \rangle = \langle a_j, b_{k+1} \rangle - \sum_{i=1}^{k} \langle a_i, b_{k+1} \rangle \langle a_j, a_i \rangle = \langle a_j, b_{k+1} \rangle - \langle a_j, b_{k+1} \rangle \cdot 1 = 0.$$

Daher steht $\tilde{a}_{k+1}$ orthogonal zu allen a_j mit $j \leq k$. Letztlich normieren wir den Vektor $\tilde{a}_{k+1}$ und erhalten einen Vektor a_{k+1} mit Norm 1. An den linearen Hüllen und der Orthogonalität ändert sich dabei natürlich nichts. □

Korollar 6.4.5. *Sei V ein endlichdimensionaler Vektorraum mit innerem Produkt. Dann lässt sich jedes Orthonormalsystem in V zu einer Orthonormalbasis erweitern. Insbesondere besitzt also V eine Orthonormalbasis.*

Beweis. Wegen Satz 6.4.3 ist jedes Orthonormalsystem linear unabhängig und lässt sich also nach dem Basisergänzungssatz zu einer (nicht unbedingt orthonormalen) Basis von V erweitern. Auf diese Basis wendet man dann das Gram-Schmidt-Verfahren an. □

Beispiel 6.4.6.

$$b_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},\ b_2 = \begin{pmatrix} -1 \\ 5 \\ -1 \\ 5 \end{pmatrix},\ b_3 = \begin{pmatrix} 5 \\ 0 \\ -1 \\ -6 \end{pmatrix}.$$

Lösung: Normierung von b_1: $\|b_1\| = \sqrt{1+1+1+1} = 2$, also setze

$$a_1 = \frac{1}{2}\, b_1 = \begin{pmatrix} 0,5 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix}.$$

Ansatz: $\tilde{a}_2 = b_2 - \langle a_1, b_2\rangle\, a_1$:
$\langle a_1, b_2\rangle = -0,5 + 2,5 - 0,5 + 2,5 = 4$, also setze

$$\tilde{a}_2 = \begin{pmatrix} -1 \\ 5 \\ -1 \\ 5 \end{pmatrix} - 4 \begin{pmatrix} 0,5 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \\ -3 \\ 3 \end{pmatrix}.$$

Es ist $\|\tilde{a}_2\| = 6$, daher

$$a_2 = \frac{1}{6} \begin{pmatrix} -3 \\ 3 \\ -3 \\ 3 \end{pmatrix} = \begin{pmatrix} -0,5 \\ 0,5 \\ -0,5 \\ 0,5 \end{pmatrix}.$$

Ansatz: $\tilde{a}_3 = b_3 - \langle a_1, b_3\rangle\, a_1 - \langle a_2, b_3\rangle\, a_2$.
$\langle a_1, b_3\rangle = -1$, $\langle a_2, b_3\rangle = -5$, also

$$\tilde{a}_3 = \begin{pmatrix} 5 \\ 0 \\ -1 \\ -6 \end{pmatrix} + 1 \begin{pmatrix} 0,5 \\ 0,5 \\ 0,5 \\ 0,5 \end{pmatrix} + 5 \begin{pmatrix} -0,5 \\ 0,5 \\ -0,5 \\ 0,5 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ -3 \\ -3 \end{pmatrix}.$$

Normierung ergibt

$$a_3 = \begin{pmatrix} 0,5 \\ 0,5 \\ -0,5 \\ -0,5 \end{pmatrix}$$

□

6.4.2. *Orthogonale Projektion mittels Orthonormalbasen.*

Satz 6.4.7. *Sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit einem inneren Produkt. Sei U ein endlichdimensionaler Unterraum von V mit einer Orthonormalbasis $(a_1, \cdots, a_n)$. Sei $x \in V$ und $\gamma_1, \cdots, \gamma_n \in \mathbb{K}$. Dann gilt:*

1) *Die orthogonale Projektion von x auf U errechnet sich durch*

$$\pi_U(x) = \sum_{i=1}^{n} \langle a_i, x\rangle\, a_i.$$

2)

$$\|\sum_{i=1}^{n} \gamma_i a_i\|^2 = \sum_{i=1}^{n} |\gamma_i|^2.$$

3)

$$\sum_{i=1}^{n} |\langle a_i, x\rangle|^2 \leq \|x\|^2 \quad \textit{(“Besselsche Ungleichung”).}$$

Gleichheit gilt genau dann, wenn $x \in U$ (”Parsevalsche Gleichung”).

Beweis. Die Beweise dieses Satzes beruhen alle auf dem Prinzip, dass Summen gebildet werden, in deren Summanden die inneren Produkte $\langle a_j, a_i\rangle$ stehen. Weil $(a_1, \cdots, a_n)$ ein Orthonormalsystem sind, fallen nun alle Summanden mit $i \neq j$ weg, weil dann $\langle a_j, a_i\rangle = 0$.
Punkt (1): Offensichtlich ist $u := \sum_{i=1}^{n} \langle a_i, x\rangle a_i \in U$. Wir zeigen, dass $x - u$ orthogonal auf alle a_j und folglich auf ganz U steht: Es ist

$$\begin{aligned} \langle a_j, x-u\rangle &= \langle a_j, x\rangle - \langle a_j, \sum_{i=1}^{n} \langle a_i, x\rangle a_i\rangle = \langle a_j, x\rangle - \sum_{i=1}^{n} \langle a_i, x\rangle \, \langle a_j, a_i\rangle \\ &= \langle a_j, x\rangle - \langle a_j, x\rangle \cdot 1 = 0. \end{aligned}$$

Punkt (2):

$$\|\sum_{i=1}^{n} \gamma_i a_i\|^2 = \langle \sum_{j=1}^{n} \gamma_j a_j, \sum_{i=1}^{n} \gamma_i a_i\rangle = \sum_{j=1}^{n}\sum_{i=1}^{n} \overline{\gamma}_j \gamma_i \langle a_j, a_i\rangle = \sum_{j=1}^{n} \overline{\gamma}_j \gamma_j = \sum_{j=1}^{n} |\gamma_j|^2.$$

Punkt (3): Nach Satz 6.3.9 (7) gilt $\|\pi_U(x)\| \leq \|x\|$ mit Gleichheit genau dann, wenn $x \in U$. Nach den eben bewiesenen Punkten (1,2) ist $\|\pi_U(x)\| = \sum_{i=1}^{n} |\langle a_i, x\rangle|^2$. □

Beispiel 6.4.8. Auf dem Raum $\mathcal{C}([0,1];\mathbb{C})$ der stetigen Funktionen von $[0,1]$ nach $\mathbb{C}$ betrachten wir das innere Produkt

$$\langle u, v\rangle = \int_0^1 \overline{u}(t) v(t)\, dt.$$

Für $k \in \mathbb{Z}$ betrachten wir die Funktionen $a_k(t) = e^{2k\pi i t}$. (Hier bedeutet i die imaginäre Zahl.) Dann ist für alle $n \in \mathbb{N}$ das System $(a_1, \cdots, a_n)$ ein Orthonormalsystem.

Beweis. Es ist

$$\langle a_k, a_j\rangle = \int_0^1 e^{-2k\pi i t} e^{2j\pi i t}\, dt = \int_0^1 e^{2(j-k)\pi i t}\, dt = \begin{cases} 1 & \text{wenn } j-k=0, \\ 0 & \text{wenn } j-k \neq 0. \end{cases}$$

□

Beispiel 6.4.9. Wir betrachten $\mathcal{C}([0,1];\mathbb{C})$ und darauf das L^2-innere Produkt wie in Beispiel 6.4.8. Wieder sei $a_k(t) = e^{2k\pi i t}$. Sei $v(t) = t$. Für $n \in \mathbb{N}$ sei U_n die lineare Hülle von $(a_{-n}, \cdots, a_n)$. Wir berechnen die orthogonale Projektion von v auf U_n.

Lösung: Nach Satz 6.4.7 berechnet man die orthogonale Projektion von x durch die Formel

$$\pi_{U_n}(v) = \sum_{k=1}^{n} \langle a_k, v\rangle a_k.$$

Für $k = 0$ erhalten wir

$$\langle a_0, v\rangle = \int_0^1 e^{0t} t\, dt = \frac{1}{2}.$$

Für alle $k \neq 0$ ergibt partielle Integration

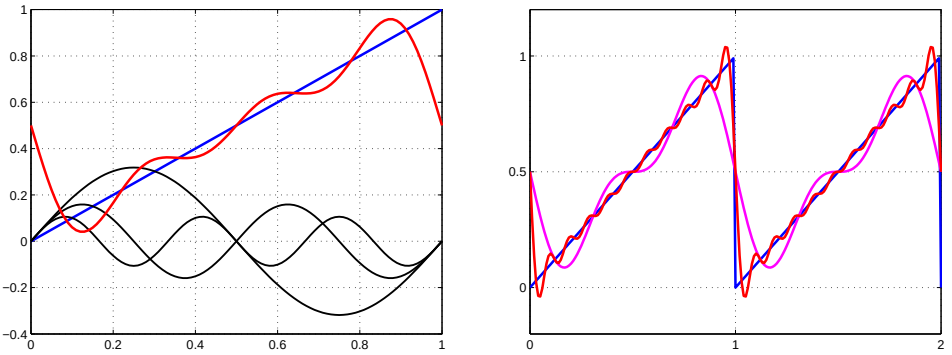
$$\langle a_k, v\rangle = \int_0^1 e^{-2k\pi i t} t\, dt = -\frac{1}{2k\pi i}\, e^{-2k\pi i t}\big|_0^1 + \int_0^1 \frac{1}{2k\pi i}\, e^{-2k\pi i t}\, dt = -\frac{1}{2k\pi i}.$$

Daher ist

$$\begin{aligned}[\pi_{U_n}(v)](t) &= \frac{1}{2}\,a_0(t) - \sum_{k=1}^{n}\left(\frac{1}{2k\pi i}\,a_k(t) - \frac{1}{2k\pi i}\,a_{-k}(t)\right)\\ &= \frac{1}{2} - \sum_{k=1}^{n}\left(\frac{1}{2k\pi i}\,e^{2k\pi it} - \frac{1}{2k\pi i}\,e^{-2k\pi it}\right)\\ &= \frac{1}{2} - \sum_{k=1}^{n}\frac{1}{k\pi}\,\sin(2k\pi t)\,.\end{aligned}$$

□

Die linke Grafik zeigt die orthogonale Projektion der Funktion $v(t) = t$ auf $\mathcal{L}(\{e^{2k\pi it} \mid k = -3, \cdots, 3\})$. Wir sehen die Funktion v, ihre orthogonale Projektion, und die Beiträge der drei Sinusschwingungen zur orthogonalen Projektion. In der rechten Grafik wurden alle Funktion periodisch (mit Periode 1) fortgesetzt. Wir sehen die orthogonalen Projektionen auf $\mathcal{L}(\{e^{2k\pi it} \mid k = -n, \cdots, n\})$ für $n = 2$ und $n = 10$. Außer an der Unstetigkeitsstelle, die bei der periodischen Fortsetzung von v entsteht, wird v durch die orthogonale Projektion gut angenähert.



Die Zerlegung einer periodischen Funktion in Sinusschwingungen ist Ausgangspunkt der Fourieranalysis, einer Denkweise, die nicht nur in der theoretischen Mathematik, sondern auch in Regelttechnik, Elektronik, Datenanalyse und Bildverarbeitung ungeheure Bedeutung besitzt. Fourieranalysis wäre ohne die Möglichkeit, Funktionen als Vektoren zu betrachten, wesentlich schwieriger zu formulieren und zu verstehen.

Satz 6.4.10. *Sei V ein endlichdimensionaler Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$ mit innerem Produkt $\langle\cdot,\cdot\rangle$. Sei $A = (a_1, \cdots, a_n)$ eine Orthonormalbasis von V und sei $x \in V$. Der Koordinatenvektor von x bezüglich der Basis A ist*

$$\phi_A^{-1}(x) = \begin{pmatrix}\langle a_1, x\rangle\\ \vdots\\ \langle a_n, x\rangle\end{pmatrix}\,.$$

Beweis. Natürlich ist x die orthogonale Projektion von x auf ganz V. Nach Satz 6.4.7 ist also

$$x = \pi_V(x) = \sum_{k=1}^{n}\langle a_k, x\rangle a_k\,.$$

Das heißt aber nichts anderes, als dass die Koordinaten von x bezüglich A die Produkte $\langle a_1, x\rangle$, …, $\langle a_n, x\rangle$ sind. □

6.5. **Dualraum, duale Abbildung, adjungierte Matrix.**

6.5.1. *Dualraum und duale Abbildung.*

Zur Entwicklung der Theorie des Dualraumes und der dualen Abbildung braucht man kein inneres Produkt. Wir werden also in diesem Unterabschnitt Vektorräume über beliebigen Körpern betrachten, und nicht von inneren Produkten und Normen reden.

Definition 6.5.1. Sei V ein endlich dimensionaler Vektorraum über einem beliebigen Körper $\mathbb{K}$.

1) Eine lineare Abbildung von V nach $\mathbb{K}$ heißt Linearform oder lineares Funktional.
2) Der Raum $\mathrm{Hom}(V, \mathbb{K})$ der linearen Funktionale auf V heißt der Dualraum V^* von V.
3) Auf V^* betrachten wir (wie üblich) die Operationen

$$\begin{aligned}[f+g](v) &:= f(v)+g(v)\,,\\ [\lambda f](v) &:= \lambda f(v)\,,\end{aligned}$$

für alle $f, g \in V^*$, $\lambda \in \mathbb{K}$ und $v \in V$.

Bemerkung 6.5.2.

1) Hier tritt $\mathbb{K}$ in seiner Rolle als Vektorraum auf: $\mathbb{K} = \mathbb{K}^1$. Die linearen Funktionale sind also Homomorphismen von V nach $\mathbb{K}^1$.
2) Wir wissen bereits, dass für alle Vektorräume V, W über demselben Körper der Homomorphismenraum $\mathrm{Hom}(V, W)$ ein Vektorraum ist.
3) Definition 6.5.1 hängt zunächst nicht von der Voraussetzung ab, dass V endliche Dimension hat. Jedoch bewährt sie sich in dieser Form nicht für unendlichdimensionale Vektorräume. In der Funktionalanalysis, also der Analysis auf unendlich dimensionalen Vektorräumen, definiert man den Dualraum als die Menge der *stetigen* linearen Funktionale. In endlich dimensionalen normierten Vektorräumen über $\mathbb{R}$ oder $\mathbb{C}$ macht das keinen Unterschied: Dort sind alle linearen Abbildungen stetig.

Satz 6.5.3. *Sei V ein endlich dimensionaler Vektorraum über einem beliebigen Körper $\mathbb{K}$ mit einer Basis $B = (b_1, \cdots, b_n)$. Für alle $k = 1, \cdots, n$ definieren wir eine lineare Abbildung $f_k : V \to \mathbb{K}$ so, dass*

$$f_k(b_j) := \begin{cases} 1 & \text{falls } j = k,\\ 0 & \text{falls } j \neq k.\end{cases}$$

(Laut Satz 4.1.12 ist dadurch f_k eindeutig bestimmt.)

1) *Dann ist $(f_1, \cdots, f_n)$ eine Basis von V^*, und jedes $g \in V^*$ lässt sich schreiben $g = \sum_{k=1}^n g(b_k) f_k$.*
2) *V^* ist isomorph zu V.*

Beweis. Wir zeigen zunächst, dass jedes $g \in V^*$ die Form $g = h := \sum_{k=1}^n g(b_k) f_k$ hat. Nach Satz 4.1.12 ist eine lineare Abbildung eindeutig durch ihre Werte auf einer Basis bestimmt, d.h., wir müssen nur zeigen, dass für alle b_j gilt: $g(b_j) = h(b_j)$. Nun ist

$$h(b_j) = \sum_{k=1}^{n} g(b_k) f_k(b_j) = g(b_j)$$

weil $f_k(b_j) = \delta_{k,j}$ (Kronecker-δ), sodass aus der ganzen Summe nur der Term für $k = j$ übrig bleibt. Also ist $g = h$, und insbesondere ist $(f_1, \cdots, f_n)$ ein Erzeugendensystem von V^*.

Wir zeigen nun, dass $(f_1, \cdots, f_n)$ linear unabhängig ist. Sei $\sum_{k=1}^n \lambda_k f_k = 0$. Dann ist für alle $j = 1, \cdots, n$:

$$0 = [\sum_{k=1}^n \lambda_k f_k](b_j) = \sum_{k=1}^n \lambda_k f_k(b_j) = \sum_{k=1}^n \lambda_k \delta_{j,k} = \lambda_j.$$

Also sind alle $\lambda_j = 0$, und die lineare Unabhängigkeit ist gezeigt.
Nun hat V^* also eine Basis aus n Funktionalen. Daher ist $\dim(V^*) = n = \dim(V)$. Vektorräume mit derselben Dimension sind aber nach Satz 4.4.3 isomorph. □

Definition 6.5.4. Sei V ein endlich dimensionaler Vektorraum über einem Körper $\mathbb{K}$, sei $(b_1, \cdots, b_n)$ eine Basis von V, und sei für $k = 1, \cdots, n$ das Funktional f_k durch die Bedingung $f_k(b_j) = \delta_{k,j}$ festgelegt (wie in Satz 6.5.3).
Dann heißt $(f_1, \cdots, f_n)$ die zu $(b_1, \cdots, b_n)$ duale Basis von V^*.

Bemerkung 6.5.5. Sei $\mathbb{K}$ ein Körper und $(e_1, \cdots, e_n)$ die Einheitsbasis in $\mathbb{K}^n$. Jedes lineare Funktional $f : \mathbb{K}^n \to \mathbb{K}$ lässt sich bekanntlich durch eine $\mathbb{K}^{1\times n}$-Matrix F darstellen: $f(x) = F \cdot x$. Wir können also $(\mathbb{K}^n)^*$ mit $\mathbb{K}^{1\times n}$ identifizieren. Es ist dann $(e_1^T, \cdots, e_n^T)$ die zur Einheitsbasis duale Basis des $\mathbb{K}^{1\times n}$.

6.5.2. *Dualität und Vektorräume mit innerem Produkt.*
Wir könnten die Theorie der Dualräume für allgemeine Vektorräume noch ein gutes Stück weiterentwickeln. Aber wir konzentrieren uns aus Zeitgründen nun auf den Sonderfall von Vektorräumen mit innerem Produkt. Inneres Produkt und Dualität spielen sehr gut zusammen.

Satz 6.5.6. *Sei V ein endlichdimensionaler Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit innerem Produkt $\langle \cdot, \cdot \rangle$. Sei V^* sein Dualraum. Dann gilt:*

1) *Für jeden Vektor $u \in V$ ist die Abbildung*

$$f_u : \begin{cases} V & \to \mathbb{K}, \\ x & \mapsto \langle u, x \rangle, \end{cases}$$

ein lineares Funktional aus V^*.

2) *Für jedes $g \in V^*$ gibt es einen eindeutigen Vektor $u \in V$ sodass $g = f_u$, d.h., für alle $x \in V$ gilt*

$$g(x) = \langle u, x \rangle.$$

3) *Die Abbildung*

$$\begin{cases} V & \to V^*, \\ u & \mapsto f_u \end{cases}$$

ist antilinear, d.h., für alle $u, v \in V$, $\lambda, \mu \in \mathbb{K}$ gilt

$$f_{\lambda u + \mu v} = \overline{\lambda} f_u + \overline{\mu} f_v.$$

4) *Die Abbildung $u \mapsto f_u$ ist bijektiv.*

Achtung, manche Autoren setzen für das innere Produkt die Linearität in der ersten Komponente und die Antilinearität in der zweiten Komponente voraus. Der Satz ist dann entsprechend abzuändern.

Beweis. Punkt (1): Das innere Produkt ist linear in der zweiten Variablen:

$$\langle u, \lambda x + \mu y\rangle = \lambda\langle u, x\rangle + \mu\langle u, y\rangle.$$

Punkt (2): Nach Satz 6.4.4 besitzt V eine Orthonormalbasis $(a_1, \cdots, a_n)$. Sei $g \in V^*$. Wir definieren

$$u := \sum_{k=1}^{n} \overline{g(a_k)} a_k.$$

Nach Satz 6.4.10 gilt für jedes $x \in V$:

$$x = \sum_{k=1}^{n} \langle a_k, x\rangle a_k.$$

Dann ist

$$g(x) = g(\sum_{k=1}^{n} \langle a_k, x\rangle a_k) = \sum_{k=1}^{n} \langle a_k, x\rangle g(a_k) = \langle \sum_{k=1}^{n} \overline{g}(a_k) a_k, x\rangle = \langle u, x\rangle.$$

Um die Eindeutigkeit zu zeigen, nehmen wir an $f_u = f_v$. Dann ist also für alle $x \in V$

$$\langle u, x\rangle = f_u(x) = f_v(x) = \langle v, x\rangle.$$

Betrachten wir insbesondere $x = u - v$. Dann ist

$$\|u - v\|^2 = \langle u - v, u - v\rangle = \langle u, u - v\rangle - \langle v, u - v\rangle = 0.$$

Also ist $u = v$.

Punkt (3):

$$f_{\lambda u + \mu v}(x) = \langle \lambda u + \mu v, x\rangle = \overline{\lambda}\langle u, x\rangle + \overline{\mu}\langle v, x\rangle = \overline{\lambda} f_u(x) + \overline{\mu} f_v(x).$$

Punkt (4): Direkte Folge der Punkte (1,2). □

Bemerkung 6.5.7. Auf Grund von Satz 6.5.6 können wir jeden endlich dimensionalen Vektorraum mit innerem Produkt mit seinem Dualraum identifizieren: Jedes $u \in V$ wird zugleich identifiziert mit der linearen Abbildung $x \mapsto \langle u, x\rangle$.

Ein analoger Satz gilt auch für *vollständige* unendlich dimensionale Vektorräume mit innerem Produkt. Auf den Begriff der Vollständigkeit aus der Analysis können wir aber in dieser Vorlesung noch nicht eingehen.

6.5.3. *Adjungierte Abbildung.*

Satz 6.5.8. *Seien V, W endlich dimensionale Vektorräume über demselben Körper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit inneren Produkten $\langle\cdot,\cdot\rangle_V$, $\langle\cdot,\cdot\rangle_W$. Sei $g \in \mathrm{Hom}(V, W)$.*

1) *Für jedes $w \in W$ ist die Abbildung*

$$\begin{cases} V & \to \mathbb{K}, \\ v & \mapsto \langle w, g(v)\rangle_W, \end{cases}$$

ein lineares Funktional aus V^*.

2) *Für jedes $w \in W$ gibt es einen eindeutig bestimmten Vektor $g^*(w) \in V$ sodass für alle $v \in V$ gilt*

$$\langle g^*(w), v\rangle_V = \langle w, g(v)\rangle_W.$$

3) *Die Abbildung*

$$g^* : \begin{cases} W & \to V, \\ w & \mapsto g^*(w) \end{cases}$$

ist linear.

4) *Identifiziert man jedes $v \in V$ bzw. $w \in W$ mit den linearen Funktionalen*

$$f_v : \begin{cases} V & \to \mathbb{K}, \\ x & \mapsto \langle v, x\rangle_V, \end{cases} \quad \text{bzw.} \quad f_w : \begin{cases} W & \to \mathbb{K}, \\ y & \mapsto \langle w, y\rangle_W, \end{cases}$$

dann entspricht $g^*(w)$ dem Funktional $f_w \circ g \in V^*$.

Beweis. Punkt (1): Es ist

$$\langle w, g(\lambda u + \mu v)\rangle_W = \langle w, \lambda g(u) + \mu g(v)\rangle_W = \lambda\langle w, g(u)\rangle_W + \mu\langle w, g(v)\rangle_W.$$

Punkt (2): Da $v \mapsto \langle w, g(v)\rangle_W$ ein lineares Funktional auf V ist, gibt es nach Satz 6.5.6 einen eindeutig bestimmten Vektor $y \in V$ sodass für alle $v \in V$ gilt: $\langle w, g(v)\rangle_W = \langle y, v\rangle_V$. Wir bezeichnen diesen Vektor y mit $g^*(w)$.

Punkt (3): Für alle $v \in V$, $w_1, w_2 \in W$, $\lambda_1, \lambda_2 \in \mathbb{K}$ gilt:

$$\begin{aligned} & \langle g^*(\lambda_1 w_1 + \lambda_2 w_2), v\rangle_V = \langle \lambda_1 w_1 + \lambda_2 w_2, g(v)\rangle_W = \overline{\lambda_1}\langle w_1, g(v)\rangle_W + \overline{\lambda_2}\langle w_2, g(v)\rangle_W \\ = \quad & \overline{\lambda_1}\langle g^*(w_1), v\rangle_V + \overline{\lambda_1}\langle g^*(w_2), v\rangle_V = \langle \lambda_1 g^*(w_1) + \lambda_2 g^*(w_2), v\rangle_V. \end{aligned}$$

Punkt (4): Es ist

$$\langle g^*(w), v\rangle_V = \langle w, g(v)\rangle_W = f_w(g(v)) = [f_w \circ g](v).$$

□

Definition 6.5.9. Seien V, W endlich dimensionale Vektorräume über demselben Körper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit inneren Produkten $\langle\cdot,\cdot\rangle_V$, $\langle\cdot,\cdot\rangle_W$. Sei $g \in \mathrm{Hom}(V, W)$ und für alle $w \in W$ sei (wie in Satz 6.5.8) $g^*(w) \in V$ definiert durch

$$(\forall v \in V) \quad \langle g^*(w), v\rangle_V = \langle w, g(v)\rangle_W.$$

Dann heißt die Abbildung

$$g^* : \begin{cases} W & \to V, \\ w & \mapsto g^*(w) \end{cases}$$

die adjungierte Abbildung zu g.

Bemerkung 6.5.10. Achtung: Geht g von V nach W, so geht g^* von W nach V.

Auch in den Dualräumen von Räumen ohne inneres Produkt kann man eine analoge Abbildung wie g^* konstruieren, die sogenannte duale Abbildung.

Satz 6.5.11. *Seien U, V, W endlich dimensionale Vektorräume über demselben Körper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit inneren Produkten $\langle\cdot,\cdot\rangle_U$, $\langle\cdot,\cdot\rangle_V$, $\langle\cdot,\cdot\rangle_W$. Sei $f \in \mathrm{Hom}(U, V)$, $g \in \mathrm{Hom}(V, W)$. Dann gilt:*

1) $[g \circ f]^* = f^* \circ g^*$.
2) $(f^*)^* = f$.
3) $\ker(f) = [\mathrm{rg}(f^*)]^\perp$.
4) $\mathrm{rg}(f) = [\ker(f^*)]^\perp$.
5) *f ist genau dann injektiv, wenn f^* surjektiv ist.*
6) *f ist genau dann surjektiv, wenn f^* injektiv ist.*

Beweis. Punkt (1): Für alle $u \in U$, $w \in W$ gilt:

$$\langle [g \circ f]^*(w), u\rangle_U = \langle w, g(f(u))\rangle_W = \langle g^*(w), f(u)\rangle_V = \langle f^*(g^*(w)), u\rangle_U.$$

Punkt (2): Für alle $u \in U$, $v \in V$ ist

$$\langle (f^*)^*(u), v\rangle_V = \langle u, f^*(v)\rangle_U = \overline{\langle f^*(v), u\rangle_U} = \overline{\langle v, f(u)\rangle_V} = \langle f(u), v\rangle_V.$$

Punkt (3): Sei $u \in \ker(f)$. Dann ist für alle $v \in V$

$$\langle f^*(v), u\rangle_U = \langle v, f(u)\rangle_V = 0,$$

daher ist $u \in [\mathrm{rg}(f^*)]^\perp$.
Sei umgekehrt $u \in [\mathrm{rg}(f^*)]^\perp$. Wir setzen $y = f(u)$. Dann ist

$$\|f(u)\|_V^2 = \langle y, f(u)\rangle_V = \langle f^*(y), u\rangle_U = 0.$$

Also ist $u \in \ker(f)$.
Punkt (4): Punkt (3), angewendet auf f^*, wobei wir benützen, dass $(f^*)^* = f$.
Punkt (5): Es ist wegen Satz 4.1.10 und Punkt (3)

$$f \text{ injektiv} \Leftrightarrow \ker(f) = \{0\} \Leftrightarrow [\ker(f)]^\perp = U \Leftrightarrow \mathrm{rg}(f^*) = U \Leftrightarrow f^* \text{ surjektiv}.$$

Punkt (6): ist Punkt (5), angewendet auf f^*. □

Erinnern Sie sich, dass man die adjungierte Matrix A^* zu einer Matrix A erhält, indem man zuerst A transponiert, und anschließend von den Koeffizienten jeweils die konjugiert komplexe Zahl nimmt.

Satz 6.5.12. *Wir betrachten über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ die Vektorräume $\mathbb{K}^m, \mathbb{K}^n$ und in beiden Räumen das innere Produkt $\langle x, y\rangle = x^* y$. Sei f eine lineare Abbildung*

$$f : \begin{cases} \mathbb{K}^n & \to \mathbb{K}^m, \\ x & \mapsto Ax, \end{cases}$$

mit ihrer Matrix $A \in \mathbb{K}^{m\times n}$.
Dann gehört zur adjungierten Abbildung f^* die adjungierte Matrix A^*.

Beweis. Sei $w \in \mathbb{K}^m$. Es gilt für alle $v \in \mathbb{K}^n$

$$\langle f^*(w), v\rangle = \langle w, f(v)\rangle = w^*(Av) = (w^*A)v = (A^*w)^*v = \langle A^*w, v\rangle.$$

Also ist $f^*(w) = A^*w$. □

An dieser Stelle unterbrechen wir unsere Diskussion von Räumen mit inneren Produkten. Wir benötigen zuerst das wichtige und schlagkräftige Instrumentarium der Eigenwerttheorie, um später damit lineare Abbildungen in Räumen mit innerem Produkt besser zu verstehen.

7. Eigenwerte

7.1. **Polynome.**

7.1.1. *Polynom und Polynomfunktion.*

Definition 7.1.1. Sei $\mathbb{K}$ ein Körper. Eine Polynomfunktion f über $\mathbb{K}$ ist eine Funktion

$$f : \begin{cases} \mathbb{K} & \to \mathbb{K}, \\ x & \mapsto \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n. \end{cases}$$

Dabei sind $\alpha_0, \cdots, \alpha_n$ Elemente aus dem Körper $\mathbb{K}$ und $n \in \mathbb{N} \cup \{0\}$.

Satz 7.1.2. *Seien $f(x) = \sum_{i=0}^{m} \alpha_i x^i$ und $g(x) = \sum_{j=0}^{n} \beta_j x^j$ zwei Polynomfunktionen über einem Körper $\mathbb{K}$. Wir definieren Addition und Multiplikation von Funktionen wie üblich: $[f+g](x) := f(x) + g(x)$, $[fg](x) := f(x)g(x)$. Dann sind auch $f+g$ und fg Polynomfunktionen, und es gilt:*

$$\begin{aligned} [f+g](x) &= \sum_{k=0}^{\max(m,n)} (\alpha_k + \beta_k) x^k \,, \\ [fg](x) &= \sum_{k=0}^{m+n} \left(\sum_{j=0}^{k} \alpha_{k-j} \beta_j \right) x^k \,. \end{aligned}$$

In der Formel für die Summe setzen wir $\alpha_i = 0$ falls $i > m$ und $\beta_j = 0$ falls $j > n$.

Beweis. Um die Summenformel zu zeigen, setzen wir $N = \max(m,n)$, und $\alpha_i = 0$ für $i > m$, $\beta_j = 0$ für $j > n$, und können daher die Polynomfunktionen schreiben

$$f(x) = \sum_{k=0}^{N} \alpha_k x^k \quad g(x) = \sum_{k=0}^{N} \beta_k x^k \,.$$

Nun ist (unter Ausnützung der Kommutativität und Assoziativität der Körperaddition, und eines Distributivgesetzes)

$$f(x) + g(x) = \sum_{k=0}^{N} \alpha_k x^k + \sum_{k=0}^{N} \beta_k x^k = \sum_{k=0}^{N} (\alpha_k + \beta_k) x^k \,.$$

Für die Multiplikation erhalten wir

$$\begin{aligned} f(x)g(x) &= \left(\sum_{i=0}^{m} \alpha_i x^i \right) \left(\sum_{j=0}^{n} \beta_j x^j \right) \\ &= \sum_{j=0}^{n} \sum_{i=0}^{m} \alpha_i \beta_j x^{i+j} && \text{(Klammern ausmultipliziert)} \\ &= \sum_{j=0}^{n} \sum_{k=j}^{m+j} \alpha_{k-j} \beta_j x^k && \text{(Verschiebung des Summationsindex: } k := i+j\text{)} \\ &= \sum_{k=0}^{m+n} \sum_{j=0}^{k} \alpha_{k-j} \beta_j x^k && \text{(Vertauschung der Summationsreihenfolge)} \\ &= \sum_{k=0}^{m+n} \left(\sum_{j=0}^{k} \alpha_{k-j} \beta_j \right) x^k \,. \end{aligned}$$

□

Beispiel 7.1.3. Auf $\mathbb{K} = \{0, 1\}$ definieren wir die Rechenoperationen

$+$	0	1
0	0	1
1	1	0

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

Man zeigt leicht, dass $\mathbb{K}$ ein Körper ist. Wir berechnen die Polynomfunktion $f(x) = x^2 + x$ für $x \in \mathbb{K}$.

Lösung:

$$\begin{aligned} f(0) &= 0^2 + 0 = 0 + 0 = 0, \\ f(1) &= 1^2 + 1 = 1 + 1 = 0. \end{aligned}$$

□

Dieses Beispiel zeigt, dass über diesem Körper die beiden Polynomfunktionen $f(x) = x^2 + x$ und $g(x) = 0$ dieselbe Funktion beschreiben. Trotzdem gibt es Gründe, zwischen den Ausdrücken $x^2 + x$ und 0 zu unterscheiden. Das ist der Grund, warum wir ein Polynom nicht einfach als Polynomfunktion definieren, sondern mit einem formaleren Zugang. Natürlich definieren wir die Rechenvorschriften für Polynome genau so, dass sie zu den Rechenvorschriften für Polynomfunktionen passen.

Definition 7.1.4. Sei $\mathbb{K}$ ein Körper.

(1) Ein Polynom über $\mathbb{K}$ ist eine Folge $(\alpha_i)_{i=0}^{\infty}$ von Elementen $\alpha_i \in \mathbb{K}$, bei der nur endlich viele Glieder von 0 verschieden sind.
(2) Auf der Menge der Polynome über $\mathbb{K}$ führen wir die folgenden Rechenoperationen ein:

$$\begin{aligned} (\alpha_i)_{i=0}^{\infty} + (\beta_i)_{i=0}^{\infty} &= (\alpha_i + \beta_i)_{i=0}^{\infty}, \\ (\alpha_i)_{i=0}^{\infty} \cdot (\beta_i)_{i=0}^{\infty} &= (\sum_{j=0}^{k} \alpha_{k-j}\beta_j)_{k=0}^{\infty}. \end{aligned}$$

(3) Das Nullpolynom ist $0 = (0)_{i=0}^{\infty}$.
(4) Ist $p = (\alpha_i)_{i=0}^{\infty}$ nicht das Nullpolynom, so definieren wir den Grad (engl.: degree) von p durch

$$\deg(p) := \max\{i \in \mathbb{N} \mid \alpha_i \neq 0\}.$$

Der Grad des Nullpolynoms wird mit $-\infty$ angesetzt.
(5) Die Polynome von Grad 0 und das Nullpolynom heißen konstante Polynome.
(6) Die Polynome von Grad ≤ 1 heißen lineare Polynome.

Da die Schreibweise durch Folgen unhandlich und intuitiv wenig ansprechend ist, wählen wir ein beliebiges Symbol (z.B. x) und schreiben

$$(\alpha_i)_{i=0}^{\infty} = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n.$$

Wir nennen dann das Symbol x die Unbestimmte. Die Menge der Polynome über $\mathbb{K}$, mit der Unbestimmten x geschrieben, bezeichnen wir mit $\mathbb{K}[x]$.

Satz 7.1.5. *Sei $\mathbb{K}$ ein Körper und $\mathbb{K}[x]$ die Menge der Polynome über $\mathbb{K}$.*

1) *$(\mathbb{K}[x], +, \cdot)$ ist ein kommutativer Ring mit Eins. Das Nullelement ist das Nullpolynom, das Einselement ist das konstante Polynom 1. (vgl. Definition 2.3.1).*
2) *Mit der Polynomaddition und dem Produkt*

$$\lambda \cdot \sum_{i=0}^{n} \alpha_i x^i := \sum_{i=0}^{n} (\lambda\alpha_i) x^i$$

bildet $\mathbb{K}[x]$ einen Vektorraum über $\mathbb{K}$.

Beweis. Nachrechnen. □

Satz 7.1.6. *Seien $f \neq 0$, $g \neq 0$ zwei Polynome über einem Körper $\mathbb{K}$. Es gilt*

$$\begin{aligned} \deg(f+g) &= \max(\deg(f), \deg(g)) \quad \textit{falls } \deg(f) \neq \deg(g)\,, \\ \deg(f+g) &\leq \deg(f) \quad \textit{falls } \deg(f) = \deg(g)\,, \\ \deg(fg) &= \deg(f) + \deg(g). \end{aligned}$$

Beweis. Sei $f = (\alpha_i)_{i=0}^{\infty}$, $g = (\beta_i)_{i=0}^{\infty}$. Sei $m = \deg(f)$ und $n = \deg(g)$.
Ist $k > \max(m,n)$, so ist $\alpha_k = \beta_k = 0$ und folglich $\alpha_k + \beta_k = 0$. Daher kann $\deg(f+g)$ nicht größer als $\max(m,n)$ sein. Sei nun $m \neq n$, o.B.d.A. nehmen wir an $m < n$. Dann ist $\alpha_n + \beta_n = \beta_n \neq 0$, folglich ist $\deg(f+g) \geq n = \max(m,n)$.
Ist $k > m+n$ und $j \leq k$, so ist entweder $j > n$ oder $k - j > m$ (oder beides), folglich muss gelten $\beta_j = 0$ oder $\alpha_{k-j} = 0$. Daher ist

$$\sum_{j=0}^{k} \alpha_{k-j}\beta_j = 0.$$

Es kann also der Grad von fg nicht größer als $m+n$ sein. Andererseits ist $\alpha_m \neq 0$ und $\beta_n \neq 0$. Daher ist

$$\sum_{j=0}^{m+n} \alpha_{k-j}\beta_j = \alpha_m\beta_n \neq 0.$$

Also ist $\deg(fg) \geq m+n$. □

Satz 7.1.7. *Sei $\mathbb{K}$ ein Körper. Der Polynomring $\mathbb{K}[x]$ ist nullteilerfrei, d.h., sind $f, g \in \mathbb{K}[x] \setminus \{0\}$, so ist $fg \neq 0$.*

Beweis. Sind $f, g \neq 0$, so sind $\deg(f) \geq 0$, $\deg(g) \geq 0$. Dann ist nach Satz 7.1.6 auch $\deg(fg) \geq 0$, also ist $fg \neq 0$. □

7.1.2. *Division mit Rest.*

Satz 7.1.8 (Polynomdivision mit Rest)**.** *Sei $\mathbb{K}$ ein Körper, f, g zwei Polynome über $\mathbb{K}$, wobei $g \neq 0$. Dann gibt es eindeutig bestimmte Polynome p, r mit den Eigenschaften*

$$\begin{aligned} f &= pg + r\,, \\ \deg(r) &< \deg(g)\,. \end{aligned}$$

Beweis. Eindeutigkeit: Sei $f = p_1 g + r_1 = p_2 g + r_2$. Dann ist also $g(p_1 - p_2) = r_2 - r_1$. Ist $p_1 \neq p_2$, so ist $\deg(p_1 - p_2) \geq 0$ und folglich $\deg(g(p_1 - p_2)) \geq 0 + \deg(g)$. Andererseits ist $\deg(r_2 - r_1) < \deg(g)$, und es entsteht ein Widerspruch. Also ist $p_1 = p_2$ und es folgt $r_1 = r_2$.
Existenz: Wir führen eine vollständige Induktion[24] nach $\deg(f)$:
Verankerung: Ist $\deg(f) < \deg(g)$, so ist $f = 0 \cdot g + f$. Wähle also $p = 0$, $r = f$.
Induktionsschritt: Sei nun $k \geq \deg(g)$ und jedes Polynom f mit $\deg(f) < k$ in der Form $f = pg + r$ darstellbar. Sei nun $f \in \mathbb{K}[x]$ mit $\deg(f) = k$. Wir zeigen, dass sich auch f zerlegen lässt. Es ist also $f = \sum_{i=0}^{k} \alpha_i x^i$, $g = \sum_{j=0}^{m} \beta_j x^j$, wobei $m = \deg(g) \leq k$, $\beta_m \neq 0$. Es ist dann

$$f_1 := f - \frac{\alpha_k}{\beta_m} x^{k-m} g$$

[24]Das ist eine sogenannte starke Induktion. Angenommen eine Aussage A gilt für eine ganze Zahl m (Verankerung). Wenn man zeigen kann, dass die Aussage auch für $k+1$ gilt, falls sie für alle ganzen Zahlen $m, m+1, \cdots, k$ gilt (Induktionsschritt), dann gilt die Aussage für alle ganzen Zahlen $\geq m$.

ein Polynom von Grad kleiner als k. Denn weil $\deg(f) = k$, $\deg(x^{k-m}g) = (k-m)+m = k$, ist $\deg(f_1) \leq k$. Aber der k-te Koeffizient von f_1 ist $\alpha_k - \frac{\alpha_k}{\beta_m}\beta_m = 0$. Nach Induktionsvoraussetzung kann man also zerlegen: $f_1 = p_1 g + r$ mit $\deg(r) < \deg(g)$. Nun ist

$$f = f_1 + \frac{\alpha_k}{\beta_m} x^{k-m} g = \left(p_1 + \frac{\alpha_k}{\beta_m} x^{k-m} \right) g + r.$$

□

Definition 7.1.9. Seien f, g Polynome über $\mathbb{K}$ mit $g \neq 0$, sei $f = pg + r$ mit $\deg(r) < \deg(g)$. Wir nennen das Polynom r den Rest bei der Division von f durch g.

Bemerkung 7.1.10. Der Existenzbeweis von Satz 7.1.8 liefert zugleich den Algorithmus für die Auffindung von p und r.

Beispiel 7.1.11. Dividiere mit Rest:

$$\underbrace{6x^4 - 7x^3 + 7x^2 + 5x - 5}_{f} : \underbrace{2x^2 - 3x + 2}_{g} .$$

Lösung:

$6x^4$	$-7x^3$	$+7x^2$	$+5x$	-5	$\frac{6x^4}{2x^2} = 3x^2$
$6x^4$	$-9x^3$	$+6x^2$			$3x^2 g$
	$2x^3$	$+x^2$	$+5x$	-5	$f_1 = f - 3x^2 g$
	$2x^3$	$-3x^2$	$+2x$		xg
		$4x^2$	$+3x$	-5	$f_2 = f_1 - xg$
		$4x^2$	$-6x$	$+4$	$2g$
			$9x$	-9	$r = f_2 - 2g$

Also ist

$$f = (3x^2 + x + 2)g + (9x - 9) .$$

□

Einfach, aber von fundamentaler Bedeutung ist der folgende Satz. Er stellt die Verbindung zwischen der Theorie der Nullstellen und der Lehre von der Teilbarkeit der Polynome her:

Satz 7.1.12. *Sei $f = \sum_{i=0}^{n} \alpha_i x^i$ ein Polynom über einem Körper $\mathbb{K}$. Die dazugehörige Polynomfunktion sei*

$$\tilde{f} : \begin{cases} \mathbb{K} & \to \mathbb{K}, \\ t & \mapsto \sum_{i=0}^{n} \alpha_i t^i . \end{cases}$$

Sei $\gamma \in \mathbb{K}$ und $f = p(x-\gamma)+r$ mit einem Polynom p (zugehörige Polynomfunktion $\tilde{p}$) und einem konstanten Polynom $r \in \mathbb{K}$.
Dann ist $r = \tilde{f}(\gamma)$.

Beweis. Wir setzen für x den Wert γ ein:

$$\tilde{f}(\gamma) = \tilde{p}(\gamma)\,(\gamma - \gamma) + r = r.$$

□

Algorithmus 7.1.13 (Horner-Schema)**.** Sei $\mathbb{K}$ ein Körper, sei $f = \sum_{i=0}^{n} \alpha_i x^i \in \mathbb{K}[x]$, sei $\gamma \in \mathbb{K}$. Gesucht ist ein Polynom p und eine Konstante r so, dass $f = p(x-\gamma)+r$. (Insbesondere ist $r = \tilde{f}(\gamma)$, wobei $\tilde{f}$ die zu f gehörende Polynomfunktion ist.)

$$\begin{aligned} \xi_n &= \alpha_n , \\ \xi_k &= \gamma \xi_{k+1} + \alpha_k \quad \text{für } k = n-1, \cdots, 0 . \end{aligned}$$

Dann ist

$$p = \sum_{i=0}^{n-1} \xi_{i+1} x^i\,, \quad r = \xi_0\,.$$

Beweis.

$$\begin{aligned}
& (x-\gamma)\sum_{i=0}^{n-1} \xi_{i+1}x^i + \xi_0 \\
= & \sum_{i=0}^{n-1} \xi_{i+1}x^{i+1} - \sum_{i=0}^{n-1} \gamma\xi_{i+1}x^i + \xi_0 \\
= & \sum_{i=1}^{n} \xi_i x^i + \xi_0 x^0 - \sum_{i=0}^{n-1} \gamma\xi_{i+1}x^i \\
= & \xi_n x^n + \sum_{i=0}^{n-1} (\xi_i - \gamma\xi_{i+1})x^i \\
= & \alpha_n x^n + \sum_{i=0}^{n-1} \alpha_i x^i.
\end{aligned}$$

□

Sie können sich auch überlegen, dass das Horner-Schema nichts anderes ist als eine verkürzte Schreibweise für die Division mit Rest: $f : (x-\gamma)$.

Beispiel 7.1.14. Sei $f = 4x^5 + 2x^4 - 6x^3 + 0x^2 - 3x + 2$, $\gamma = 2$. Schreiben Sie $f = p(x-2) + r$. Was ist $\tilde{f}(2)$?

Lösung:

i	5	4	3	2	1	0
α_i	4	2	-6	0	-3	2
ξ_i	4	10	14	28	53	108

Es ist also $p = 4x^4 + 10x^3 + 14x^2 + 28x + 53$, $r = \tilde{f}(2) = 108$. □

Viel schneller als die einzelnen Potenzen auszurechnen: $\tilde{f}(2) = 4 \cdot 2^5 + 2 \cdot 2^4 - 6 \cdot 2^3 - 3 \cdot 2 + 2$!

7.1.3. *Nullstellen.*

Definition 7.1.15. Sei $\mathbb{K}$ ein Körper und $f \in \mathbb{K}[x]$ ein Polynom mit zugehöriger Polynomfunktion $\tilde{f}$. Ein $\gamma \in \mathbb{K}$ heißt Nullstelle von f, wenn gilt $\tilde{f}(\gamma) = 0$.

Satz 7.1.16. *Sei $\mathbb{K}$ ein Körper und $f \in \mathbb{K}[x]$ ein Polynom. Sei $\gamma \in \mathbb{K}$ und $f = p(x-\gamma) + r$. Dann ist γ ist genau dann eine Nullstelle von f, wenn $r = 0$, d.h., f ist durch $x - \gamma$ ohne Rest teilbar.*

Beweis. Direkte Folge von Satz 7.1.12. □

Definition 7.1.17. Sei $\mathbb{K}$ ein Körper, $f \neq 0 \in \mathbb{K}[x]$, und $\gamma \in \mathbb{K}$ eine Nullstelle von f. Die Vielfachheit der Nullstelle γ in f ist die größte Potenz m so, dass $(x-\gamma)^m$ ein Teiler von f ist.

Anders ausgedrückt: Ist m die Vielfachheit von γ in f, so ist $f = (x-\gamma)^m p$, wobei γ keine Nullstelle von p ist.

Beispiel 7.1.18. Überprüfen Sie, dass -3 eine Nullstelle von $f = x^4 + 7x^3 + 9x^2 - 27x - 54$ ist, und bestimmen Sie die Vielfachheit.

Lösung: Wie oft läßt sich $x+3$ aus f ohne Rest abdividieren? Wir dividieren natürlich mit dem Horner-Schema:

1	7	9	-27	- 54	f
1	4	-3	-18	0	Rest 0, $f = f_1(x+3)$, $f_1 = x^3 + 4x^2 - 3x - 18$,
1	1	-6	0		Rest 0, $f_1 = f_2(x+3)$, $f_2 = x^2 + x - 6$,
1	-2	0			Rest 0, $f_2 = f_3(x+3)$, $f_3 = x - 2$,
1	-5				Rest $\tilde{f}_3(-3) = -5 \neq 0$.

Also ist $f = (x+3)^3 f_3$ mit $\tilde{f}_3(-3) = -5 \neq 0$. Die Vielfachheit von -3 als Nullstelle von f ist 3. □

Die folgende nützliche Regel zur Bestimmung von Vielfachheiten verwendet den Begriff der Ableitung aus der Differentialrechnung, und dieser beruht auf der Grenzwertrechnung. Grenzwertrechnung funktioniert nicht über jedem Körper, aber der Begriff der Ableitung läßt sich für Polynome rein algebraisch definieren.

Bemerkung 7.1.19. Sei $\mathbb{K}$ ein Körper. Für ein Polynom $f = \sum_{i=0}^{n} \alpha_i x^i$ definieren wir die erste Ableitung $f' := \sum_{i=1}^{n} i\alpha_i x^{i-1}$, und die höheren Ableitungen f'', $f^{(3)}$, $f^{(4)}$ ...wie üblich durch wiederholte Anwendung der ersten Ableitung. $f^{(0)}$ ist einfach f.
Sei nun $f \neq 0 \in \mathbb{K}[x]$ und $\gamma \in \mathbb{K}$ eine Nullstelle von f. Dann gilt:
Die Vielfachheit von γ in f ist m genau dann, wenn γ eine Nullstelle der Polynome f, f', ..., $f^{(m-1)}$ ist, aber keine Nullstelle von $f^{(m)}$.

Beweis. Wir geben nur eine Skizze, da wir die Theorie der Ableitung von Polynomen nicht im Detail entwickeln wollen. Der erste Schritt des Beweises wäre, dass wie über $\mathbb{R}$ und $\mathbb{C}$ für die Ableitung die folgenden Regeln gelten:

$$\begin{aligned} [fg]' &= f'g + g'f \quad \text{(Produktregel)}, \\ [(x-\gamma)^m]' &= m(x-\gamma)^{m-1}. \end{aligned}$$

Sei nun γ eine Nullstelle von f mit der Vielfachheit m. Wir zeigen durch vollständige Induktion: Für $k = 0, \cdots, m-1$ ist γ eine Nullstelle der Vielfachheit $m-k$ von $f^{(k)}$.
Induktionsverankerung: Nach Voraussetzung ist γ eine Nullstelle der Vielfachheit m von $f^{(0)} = f$.
Induktionsschritt: Sei $k < m$ und γ eine Nullstelle der Vielfachheit $m-k$ von $f^{(k)}$, das heißt $f^{(k)} = (x-\gamma)^{m-k}p$, wobei γ keine Nullstelle von p ist. Dann ist

$$\begin{aligned} f^{(k+1)} &= (m-k)(x-\gamma)^{m-k-1}p + (x-\gamma)^{m-k}p' \\ &= (x-\gamma)^{m-k-1}\left((m-k)p + (x-\gamma)p'\right) = (x-\gamma)^{m-k-1}q. \end{aligned}$$

Einsetzen in die Polynomfunktionen ergibt

$$\tilde{q}(\gamma) = (m-k)\tilde{p}(\gamma) + (\gamma-\gamma)\tilde{p}'(\gamma) = (m-k)\tilde{p}(\gamma) \neq 0\,.$$

Daher ist γ keine Nullstelle von q. Ist nun $k+1 < m$, so folgt dass γ eine Nullstelle der Vielfachheit $m-k-1$ von $f^{(k+1)}$ ist. Damit ist die Induktion fertig. Ist aber $k+1 = m$, so ist $m-k-1 = 0$ und $f^{(m)} = q$. Daher ist γ keine Nullstelle von $f^{(m)}$. □

Satz 7.1.20. *Sei $\mathbb{K}$ ein Körper. Sei $\gamma \in \mathbb{K}$ und $f = (x-\gamma)^m p \neq 0$ ein Polynom über $\mathbb{K}$. Sei $\lambda \neq \gamma$ und $n \in \mathbb{N}$.*
Dann ist λ genau dann eine Nullstelle der Vielfachheit n von f, wenn λ eine Nullstelle der Vielfachheit n von p ist.

Beweis. Wenn $(x-\lambda)^n$ das Polynom p teilt, so ist also $p = q(x-\lambda)^n$ und folglich $f = q(x-\gamma)^m(x-\lambda)^n$. Es folgt daher: Ist λ eine Nullstelle der Vielfachheit n von p, so ist λ auch eine Nullstelle von f mit Vielfachheit von mindestens n.
Wir zeigen nun durch vollständige Induktion nach n: Wenn $(x-\lambda)^n$ ein Teiler von f ist, dann teilt $(x-\lambda)^n$ das Polynom p. Wenn wir das gezeigt haben, können wir schließen: Ist λ eine Nullstelle der Vielfachheit n von f, so ist λ eine Nullstelle von p mit Vielfachheit von mindestens n, und wir sind fertig.
Induktionsverankerung $n = 1$: Sei λ eine Nullstelle von f, d.h., $0 = \tilde{f}(\lambda) = (\lambda - \gamma)^m\tilde{p}(\lambda)$. Weil $\lambda - \gamma \neq 0$, ist auch $(\lambda-\gamma)^m \neq 0$, und es folgt $\tilde{p}(\lambda) = 0$. Also ist λ auch eine Nullstelle von p.
Induktionsschritt: Die Behauptung gelte für n. Sei $(x-\lambda)^{n+1}q = f$. Auf Grund der Induktionsverankerung wissen wir, dass $(x-\lambda)$ auch ein Teiler von p ist. Es ist also $p = (x-\lambda)p_1$. Nun ist

$$\begin{aligned}(x-\lambda)^{n+1}q = f &= (x-\gamma)^m p_1(x-\lambda)\,,\\ (x-\lambda)\left[(x-\lambda)^n q - (x-\gamma)^m p_1\right] &= 0\,.\end{aligned}$$

Wegen der Nullteilerfreiheit des Polynomringes (Satz 7.1.7) folgt

$$(x-\lambda)^n q = (x-\gamma)^m p_1\,.$$

Nach der Induktionsvoraussetzung ist also $(x-\lambda)^n$ ein Teiler von p_1, d.h., $p_1 = (x-\lambda)^n p_2$. Damit ist

$$p = (x-\lambda)p_1 = (x-\lambda)^{n+1}p_2\,,$$

und die Induktion ist fertig. □

Satz 7.1.21. *Sei $\mathbb{K}$ ein Körper und $f \neq 0 \in \mathbb{K}[x]$. Seien $\lambda_1, \cdots, \lambda_k$ die Nullstellen von f mit den Vielfachheiten $m_1, \cdots, m_k$.*

1) *Dann gibt es ein Polynom $p \in \mathbb{K}[x]$, welches keine Nullstellen besitzt, so dass*
$$f = (x-\lambda_1)^{m_1}(x-\lambda_2)^{m_2}\cdots(x-\lambda_k)^{m_k}p\,.$$
2) *Das Polynom f hat höchstens so viele Nullstellen (gezählt mit Vielfachheit) wie sein Grad ist: $\sum_{j=1}^k m_j \leq \deg(f)$.*

Beweis. Teil (1): Vollständige Induktion nach k:
Induktionsverankerung $k = 1$: Weil λ_1 eine m_1-fache Nullstelle von f ist, kann f zerlegt werden: $f = (x-\lambda_1)^{m_1}p$. Da jede Nullstelle von p auch eine weitere Nullstelle von f wäre, hat p keine Nullstellen.
Induktionsschritt: Der Satz sei richtig für ein Polynom mit k Nullstellen. Sei nun f ein Polynom mit $k+1$ Nullstellen. Weil λ_{k+1} eine m_{k+1}-fache Nullstelle von f ist, gilt $f = (x-\lambda_{k+1})^{m_{k+1}}f_1$, dabei sind alle weiteren Nullstellen von f auch Nullstellen von f_1 mit derselben Vielfachheit. Daher hat f_1 genau k Nullstellen, und wir können die Induktionsvoraussetzung anwenden: $f_1 = p\,\Pi_{j=1}^k(x-\lambda_j)^{m_j}$. Dabei hat p keine Nullstellen. Schließlich erhalten wir

$$f = (x-\lambda_{k+1})^{m_{k+1}}f_1 = p\,\Pi_{j=1}^{k+1}(x-\lambda_j)^{m_j}$$

und die Induktion ist fertig.
Teil (2): In Hinblick auf Teil (1) ist

$$\deg(f) = \deg(p\,\Pi_{j=1}^{k+1}(x-\lambda_j)^{m_j}) = \deg(p) + \sum_{j=1}^k m_j \geq \sum_{j=1}^k m_j\,.$$

□

Korollar 7.1.22. *Sei $\mathbb{K}$ ein Körper mit unendlich vielen Elementen, und seien f, g zwei verschiedene Polynome über $\mathbb{K}$. Dann sind auch die Polynomfunktionen $\tilde{f}$, $\tilde{g}$ verschieden.*

Beweis. Sei $\tilde{f} = \tilde{g}$, d.h., für alle $t \in \mathbb{K}$ ist $\tilde{f}(t) - \tilde{g}(t) = 0$. Es ist also jedes Element aus $\mathbb{K}$ eine Nullstelle des Polynoms $f - g$. Wenn $f - g$ nicht das Nullpolynom wäre, könnte es aber nur endlich viele Nullstellen geben. Also ist $f = g$. □

Überlegen Sie selbst, dass es über einem Körper mit endlich vielen Elementen immer Polynomfunktionen gibt, die durch mehrere verschiedene Polynome dargestellt werden können.

Definition 7.1.23. Sei $\mathbb{K}$ ein Körper. $\mathbb{K}$ heißt algebraisch abgeschlossen, wenn jedes nicht konstante Polynom über $\mathbb{K}$ mindestens eine Nullstelle besitzt.

Bemerkung 7.1.24. $\mathbb{K}$ ist genau dann algebraisch abgeschlossen, wenn jedes Polynom $f \neq 0$ über $\mathbb{K}$ in das Produkt von Linearfaktoren und einer Konstanten zerfällt, also wenn es $\lambda_1, \cdots, \lambda_k, \alpha \in \mathbb{K}$ und $m_1, \cdots, m_k \in \mathbb{N}$ gibt, so dass

$$f = \alpha \, \Pi_{j=1}^{k} (x - \lambda_j)^{m_j}.$$

Anders ausgedrückt: Die Anzahl der Nullstellen von f, gezählt mit Vielfachheiten, ergibt genau den Grad von f.

Beweis. Nach Satz 7.1.21 lässt sich f zerlegen

$$f = p \, \Pi_{j=1}^{k} (x - \lambda_j)^{m_j} .$$

Dabei hat p keine Nullstellen.
"⇒": Sei $\mathbb{K}$ algebraisch abgeschlossen, und hat p keine Nullstellen, so muss p ein konstantes Polynom sein. Es ist dann auch $\deg(p) = 0$ und folglich $\deg(f) = \sum_{j=1}^{k} m_j$.
"⇐": Sei $\mathbb{K}$ nicht algebraisch abgeschlossen. Dann gibt es ein nicht konstantes Polynom f ohne Nullstellen. Es ist dann in der obigen Zerlegung $f = p$ und p nicht konstant. □

Für den Beweis des folgenden tiefen Satzes von fundamentaler Bedeutung fehlen uns hier die Grundlagen:

Satz 7.1.25 (Hauptsatz der Algebra)**.** *Der Körper $\mathbb{C}$ der komplexen Zahlen ist algebraisch abgeschlossen.*

Überlegen Sie selbst, dass der Körper $\mathbb{R}$ der reellen Zahlen nicht algebraisch abgeschlossen ist.

7.2. **Eigenvektoren und Eigenwerte.**

7.2.1. *Beispiele.*

Beispiel 7.2.1. Wir betrachten die folgende Matrix $A \in \mathbb{R}^{2\times 2}$ und die Vektoren $x, y \in \mathbb{R}^2$:

$$A = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad x = \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \quad y = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}.$$

Wir berechnen Ax und Ay. Welche Folgerungen können wir über die lineare Abbildung f schließen, die durch A gegeben wird?

Lösung: Einfache Rechnung zeigt: $Ax = -x$, $Ay = y$. Überdies ist $B = (x, y)$ eine Orthonormalbasis des $\mathbb{R}^2$. Die Matrix von f bezüglich der Basis B erhalten wir nach der Regel: Die Spalten sind die Bilder der Basisvektoren (natürlich auch in neuen Koordinaten geschrieben):

$$M_B^B(f) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Also gibt es ein Koordinatensystem in $\mathbb{R}^2$, sodass die Abbildung f das Vorzeichen der ersten Koordinate umkehrt, und die zweite Koordinate unverändert lässt. Die Abbildung f ist also eine Spiegelung an der Geraden

$$\begin{pmatrix}0\\0\end{pmatrix} + t\begin{pmatrix}\frac{\sqrt{3}}{2}\\ \frac{1}{2}\end{pmatrix}.$$

□

Beispiel 7.2.2. Wir betrachten die lineare Abbildung mit der folgenden Matrix $A \in \mathbb{R}^{3\times 3}$ und die Vektoren $x, y, z \in \mathbb{R}^3$.

$$A = \begin{pmatrix}4 & 2 & -6\\ 8 & 19 & -42\\ 4 & 8 & -18\end{pmatrix}, \quad x = \begin{pmatrix}1\\2\\1\end{pmatrix}, \quad y = \begin{pmatrix}-2\\1\\0\end{pmatrix}, \quad z = \begin{pmatrix}1\\4\\2\end{pmatrix}.$$

Wieder berechnen wir Ax, Ay und Az.

Lösung: Wir erhalten: $Ax = 2x$, $Ay = 3y$, $Az = 0$. Verwenden wir $B = (x, y, z)$ als Basis für ein neues Koordinatansystem, so ist

$$M_B^B(f) = \begin{pmatrix}2 & 0 & 0\\ 0 & 3 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

In diesem Koordinatensystem lässt sich f also leicht beschreiben: Die erste Koordinate wird verdoppelt, die zweite Koordinate auf das Dreifache gestreckt, und die dritte Koordinate auf 0 gesetzt. Beachten Sie aber, dass diese Basis kein Orthonormalsystem ist. □

In beiden Beispielen gewinnen wir Einsicht in die Struktur der linearen Abbildung, indem wir das Koordinatensystem wechseln. Unsere Basis bestand immer aus Vektoren der Form $Ax = \lambda x$ mit $\lambda \in \mathbb{K}$.

Beispiel 7.2.3. Sei $A \in \mathbb{R}^{3\times 3}$ die Matrix einer Drehung um die Achse $0 + tv$. Dann lässt A die Punkte auf der Achse fest, also $Av = v$.

7.2.2. *Eigenwerte, Eigenvektoren, Eigenraum.*

Definition 7.2.4. Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix.
Ein Vektor $v \in \mathbb{K}^n \setminus \{0\}$ heißt Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{K}$, wenn gilt $Av = \lambda v$. Insbesondere heißt also $\lambda \in \mathbb{K}$ ein Eigenwert von A, wenn es dazu einen Eigenvektor $v \in \mathbb{K}^n \setminus \{0\}$ gibt.

Satz 7.2.5. *Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. Sei $\lambda \in \mathbb{K}$, und sei $W = \{v \in \mathbb{K}^n \mid Av = \lambda v\}$.*

1) *$W \neq \{0\}$ genau dann wenn λ ein Eigenwert von A ist.*
2) *$W = \ker(\lambda E_n - A)$.*
3) *W ist ein Unterraum von $\mathbb{K}^n$.*

Beweis. Teil (1): Für jedes $\lambda \in \mathbb{K}$ gilt $A0 = \lambda 0$. Es ist λ genau dann ein Eigenwert, wenn es zusätzlich ein $v \neq 0$ gibt mit $Av = \lambda v$.
Teil (2): Es gilt

$$Av = \lambda v \;\Leftrightarrow\; 0 = \lambda v - Av \;\Leftrightarrow\; 0 = (\lambda E_n - A)v\,.$$

Teil (3): Der Kern einer Matrix ist immer ein Unterraum. □

Definition 7.2.6. Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. Sei $\lambda \in \mathbb{K}$ ein Eigenwert von A. Der Raum $\ker(\lambda E_n - A)$ heißt Eigenraum von A zum Eigenwert λ. Die Dimension von $\ker(\lambda E_n - A)$ heißt geometrische Vielfachheit von λ als Eigenwert von A.

Schreibweise 7.2.7. Statt $\lambda E_n - A$ schreibt man auch einfach $\lambda - A$.

Bemerkung 7.2.8. Der Eigenraum von A zu $\lambda = 0$ ist einfach der Kern von A.

7.2.3. *Berechnung von Eigenwerten.*

Satz 7.2.9. *Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. Sei $\lambda \in \mathbb{K}$. Dann gilt:*
λ ist genau dann ein Eigenwert von A wenn $\det(\lambda - A) = 0$ ist.

Beweis. Wegen Satz 7.2.5 ist λ genau dann ein Eigenwert, wenn $\ker(\lambda - A) \neq \{0\}$ ist, und das gilt genau dann, wenn $(\lambda - A)$ singulär ist. Eine Matrix ist wiederum genau dann singulär, wenn ihre Determinante null ist. □

Damit ergibt sich folgendes Verfahren zur Berechnung von Eigenwerten:

Algorithmus 7.2.10. Die Eigenräume und Eigenwerte von $A \in \mathbb{K}^{n\times n}$ kann man folgendermaßen berechnen:

1) Setze λ als Unbekannte an und berechne $\det(\lambda - A)$. Es ergibt sich ein Polynom vom Grad n[25].
2) Berechne (bei Bedarf mit einem numerischen Gleichungslöser) die Nullstellen dieses Polynoms. Sie sind die Eigenwerte von A.
3) Zu jedem Eigenwert λ ist $\ker(\lambda - A)$ der zugehörige Eigenraum von A.

Für spezielle Matrizen (z.B. hermitesche) gibt es ganz andere Zugänge zur Berechnung der Eigenwerte. Sie hören davon mehr in einer Lehrveranstaltung über numerische Mathematik.

Beispiel 7.2.11. Berechnen Sie alle Eigenwerte und Eigenräume der Matrix A als Matrix über $\mathbb{R}$ und als Matrix über $\mathbb{C}$.

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

Lösung: Schritt 1: Berechnung von $\det(\lambda - A)$ mittels Entwicklung nach der ersten Zeile.

$$\begin{aligned}
& \det\Big(\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} - \begin{pmatrix} -1 & 1 & 2 \\ -1 & -5 & 2 \\ 2 & -2 & -4 \end{pmatrix} \Big) \\
= & \begin{vmatrix} \lambda+1 & -1 & -2 \\ 1 & \lambda+5 & -2 \\ -2 & 2 & \lambda+4 \end{vmatrix} \\
= & (\lambda+1) \begin{vmatrix} \lambda+5 & -2 \\ 2 & \lambda+4 \end{vmatrix} - (-1) \begin{vmatrix} 1 & -2 \\ -2 & \lambda+4 \end{vmatrix} + (-2) \begin{vmatrix} 1 & \lambda+5 \\ -2 & 2 \end{vmatrix} \\
= & (\lambda+1)[(\lambda+5)(\lambda+4)+4] + [\lambda+4-4] - 2[2+2(\lambda+5)] \\
= & \lambda^3 + 10\lambda^2 + 30\lambda\,.
\end{aligned}$$

Schritt 2: Nullstellen suchen: Offensichtlich ist $\lambda_1 = 0$ eine Nullstelle, damit haben wir einen Eigenwert. Die anderen Eigenwerte, falls welche existieren, sind Nullstellen des quadratischen Polynoms $\lambda^2 + 10\lambda + 30$. Diese gibt es nur im Körper der komplexen Zahlen: $\lambda_{2,3} = -5 \pm i\sqrt{5}$. Es hat also A über $\mathbb{R}$ nur einen Eigenwert, nämlich 0, über $\mathbb{C}$ besitzt A zusätzlich 2 konjugiert komplexe Eigenwerte.

[25]Das werden wir in Kürze beweisen.

Schritt 3: Berechnung der Eigenräume. Wir arbeiten mit Pivotschritten.
$W_1 := \ker(0 - A)$ über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$:

$$\begin{pmatrix} 1* & -1 & -2 \\ 1 & 5 & -2 \\ -2 & 2 & 4 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & -2 \\ 0 & 6* & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Die dritte Spalte bleibt als nichtbasische Spalte, wir setzen $x_3 = t$ und erhalten

$$\ker(-A) = \{t \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \mid t \in \mathbb{K}\}.$$

$W_2 = \ker(-5 + i\sqrt{5} - A)$ über $\mathbb{C}$:

$$\begin{pmatrix} -4 + i\sqrt{5} & -1 & -2 \\ 1* & i\sqrt{5} & -2 \\ -2 & 2 & -1 + i\sqrt{5} \end{pmatrix} \to \begin{pmatrix} 0 & 4 + 4i\sqrt{5} & -10 + 2i\sqrt{5} \\ 1 & i\sqrt{5} & -2 \\ 0 & 2 + 2i\sqrt{5}* & -5 + i\sqrt{5} \end{pmatrix}$$

$$\to \quad \begin{pmatrix} 0 & 0 & 0 \\ 1 & i\sqrt{5} & -2 \\ 0 & 2 + 2i\sqrt{5}* & -5 + i\sqrt{5} \end{pmatrix}$$

Wir setzen $x_3 = t$ und erhalten

$$x_2 = -t\,\frac{-5 + i\sqrt{5}}{2 + 2i\sqrt{5}} = -t\,\frac{i\sqrt{5}(i\sqrt{5} + 1)}{2(i\sqrt{5} + 1)} = -t\frac{i\sqrt{5}}{2}.$$

Letztlich ist

$$x_1 = -i\sqrt{5}\,x_2 + 2t = -\frac{t}{2}.$$

Damit ist der Eigenraum zu $\lambda_2 = -5 + i\sqrt{5}$

$$\ker(5 - i\sqrt{5} - A) = \{t \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{5}}{2} \\ 1 \end{pmatrix} \mid t \in \mathbb{C}\}.$$

Für $\lambda_3 = -5 - i\sqrt{5}$ ist die Rechnung ebenso, nur mit konjugiert komplexen Zahlen, und es ist

$$\ker(-5 + i\sqrt{5} - A) = \{t \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{5}}{2} \\ 1 \end{pmatrix} \mid t \in \mathbb{C}\}.$$

Alle 3 Eigenräume sind eindimensional. Damit besitzt A über $\mathbb{R}$ einen Eigenwert mit geometrischer Vielfachheit 1. Über $\mathbb{C}$ besitzt A drei Eigenwerte mit geometrischer Vielfachheit jeweils 1. □

7.2.4. *Charakteristisches Polynom.*

Satz 7.2.12. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. Für $J \subset \{1, \cdots, n\}$ sei $\#J$ die Anzahl der Elemente von J, und sei A_J die $(n - \#J) \times (n - \#J)$-Matrix, die man erhält wenn man alle Zeilen und Spalten mit den Indices $i \in J$ aus der Matrix A streicht. Dann gilt*

$$\begin{aligned} \det(\lambda - A) &= \sum_{k=0}^{n} c_k \lambda^k \quad \textit{mit} \\ c_k &= (-1)^{n-k} \sum_{J\subset\{1,\cdots,n\},\, \#J=k} \det(A_J) \quad \textit{für } 0 \le k < n, \\ c_n &= 1. \end{aligned}$$

Insbesondere ist also $\det(\lambda - A)$ ein Polynom vom Grad n.

Beweis. Sei $A = (\alpha_{i,j})_{i,j=1,\cdots,n}$.
Für $0 \le m \le n$ setzen wir

$$Q_m := \begin{pmatrix} \lambda & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix},$$

dabei sind die ersten m Zeilen den Term λ, die anderen sind Nullzeilen. Wir zeigen durch vollständige Induktion nach m:

$$\det(Q_m - A) = \sum_{k=0}^{m} (-1)^{n-k} \lambda^k \sum_{J \subset \{1,\cdots,m\},\, \#J=k} \det(A_{i_1,\cdots,i_k}) .$$

Für $m = n$ ergibt sich dann genau der gewünschte Satz.
Induktionsverankerung $m = 0$: Hier haben wir gar keine λ_i, und es gilt (wegen der Linearität der Determinante in den einzelnen Zeilen)

$$\det(-A) = \det((-1)A) = (-1)^n \det(A) = (-1)^{n-0} \det(A_\emptyset) .$$

Induktionsschritt von m auf $m+1$: Wir verwenden die Linearität in der $(m+1)$-ten Zeile:

$$\begin{vmatrix} \lambda - \alpha_{1,1} & \cdots & -\alpha_{1,m} & -\alpha_{1,m+1} & \cdots & -\alpha_{1,n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ -\alpha_{m,1} & \cdots & \lambda - \alpha_{m,m} & -\alpha_{m,m+1} & \cdots & -\alpha_{m,n} \\ -\alpha_{m+1,1} & \cdots & -\alpha_{m+1,m} & \lambda - \alpha_{m+1,m+1} & \cdots & -\alpha_{m+1,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ -\alpha_{n,1} & \cdots & -\alpha_{n,m} & -\alpha_{n,m+1} & \cdots & -\alpha_{n,n} \end{vmatrix}$$

$$= \begin{vmatrix} \lambda - \alpha_{1,1} & \cdots & -\alpha_{1,m} & -\alpha_{1,m+1} & \cdots & -\alpha_{1,n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ -\alpha_{m,1} & \cdots & \lambda - \alpha_{m,m} & -\alpha_{m,m+1} & \cdots & -\alpha_{m,n} \\ -\alpha_{m+1,1} & \cdots & -\alpha_{m+1,m} & -\alpha_{m+1,m+1} & \cdots & -\alpha_{m+1,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ -\alpha_{n,1} & \cdots & -\alpha_{n,m} & -\alpha_{n,m+1} & \cdots & -\alpha_{n,n} \end{vmatrix}$$

$$- \lambda \begin{vmatrix} \lambda - \alpha_{1,1} & \cdots & -\alpha_{1,m} & -\alpha_{1,m+1} & \cdots & -\alpha_{1,n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ -\alpha_{m,1} & \cdots & \lambda - \alpha_{m,m} & -\alpha_{m,m+1} & \cdots & -\alpha_{m,n} \\ 0 & \cdots & 0 & -1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ -\alpha_{n,1} & \cdots & -\alpha_{n,m} & -\alpha_{n,m+1} & \cdots & -\alpha_{n,n} \end{vmatrix}$$

$$= \det(Q_m - A) - \lambda \det(Q_m - \tilde{A}) ,$$

dabei unterscheidet sich $\tilde{A}$ von A nur in der $(m+1)$-ten Zeile, in welcher der $(m+1)$-te (Zeilen-)Einheitsvektor steht. Entwickelt man für $J \subset \{1, \cdots, m\}$ die Determinante $\det(\tilde{A}_J)$ nach dieser Zeile, so sieht man sofort:

$$\det(\tilde{A}_J) = \det(A_{J \cup \{m+1\}}) .$$

Wir wenden die Induktionsvoraussetzung auf beide Matrizen an:

$$\begin{aligned}
&\det(Q_{m+1}-A) = \det(Q_m - A) - \lambda \det(Q_m - \tilde{A}) \\
= &\sum_{k=0}^{m}(-1)^{n-k}\lambda^k \sum_{J\subset\{1,\cdots,m\},\,\#J=k} \det(A_J) \\
&-\lambda\sum_{k=0}^{m}(-1)^{n-k}\lambda^k \sum_{J\subset\{1,\cdots,m\},\,\#J=k} \det(\tilde{A}_J) \\
= &\sum_{k=0}^{m}(-1)^{n-k}\lambda^k \sum_{J\subset\{1,\cdots,m\},\,\#J=k} \det(A_J) \\
&+\sum_{k=0}^{m}(-1)^{n-k-1}\lambda^{k+1} \sum_{J\subset\{1,\cdots,m\},\,\#J=k} \det(A_{J\cup\{m+1\}})\,.
\end{aligned}$$

In der ersten Summe werden alle Teilmengen $\{i_1,\cdots,i_k\}$ von $\{1,\cdots,m+1\}$ aufgezählt, die $m+1$ nicht enthalten. In der zweiten Summe werden alle Teilmengen $\{i_1,\cdots,i_k,m+1\}$ aufgezählt, also jene, die $m+1$ enthalten. Damit werden insgesamt alle Teilmengen $J\subset\{1,\cdots,m+1\}$ aufgezählt, und die Induktion ist fertig. □

Beispiel 7.2.13. Wir berechnen $\det(\lambda - A)$ für

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

Lösung: Die Determinante von A bekommen wir mittels Regel von Sarrus:

$$\begin{aligned}
c_0 &= (-1)^3 \begin{vmatrix} 2 & 1 & 3 \\ 4 & 4 & -1 \\ 0 & -1 & 5 \end{vmatrix} = -(40-12-2-20) = -6\,, \\
c_1 &= (-1)^2 \left(\begin{vmatrix} 4 & -1 \\ -1 & 5 \end{vmatrix} + \begin{vmatrix} 2 & 3 \\ 0 & 5 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ 4 & 4 \end{vmatrix} \right) = 19+10+4 = 33\,, \\
c_2 &= (-1)^1(2+4+5) = -11\,.
\end{aligned}$$

Daher ist

$$\det(\lambda - A) = \lambda^3 - 11\lambda^2 + 33\lambda - 6\,.$$

□

Definition 7.2.14. Sei $\mathbb{K}$ ein Körper und $A\in\mathbb{K}^{n\times n}$ eine quadratische Matrix.

1) Dann heißt das Polynom $\det(\lambda - A)$ das charakteristische Polynom von A.
2) Ist γ ein Eigenwert von A, dann ist die algebraische Vielfachheit von γ als Eigenwert von A die Vielfachheit von γ als Nullstelle des charakteristischen Polynoms.

Definition 7.2.15. Sei $\mathbb{K}$ ein Körper und $A=(\alpha_{i,j})_{i,j=1\cdots n}\in\mathbb{K}^{n\times n}$ eine quadratische Matrix. Die Spur (engl.: trace) von A ist die Summe der Diagonalelemente:

$$\operatorname{trace}(A) = \sum_{i=1}^{n}\alpha_{i,i}\,.$$

Satz 7.2.16. *Sei $\mathbb{K}$ ein Körper und $A\in\mathbb{K}^{n\times n}$ eine quadratische Matrix.*

1) *Dann hat das charakteristische Polynom von A die Gestalt:*

$$\lambda^n - \operatorname{trace}(A)\lambda^{n-1} + c_{n-2}\lambda^{n-2} - \cdots \pm c_1\lambda + (-1)^n\det(A)\,.$$

2) *Hat das charakteristische Polynom (mit Vielfachheiten gezählt) n Eigenwerte $\gamma_1, \cdots, \gamma_n$, dann ist*

$$\mathrm{trace}(A) = \sum_{i=1}^{n} \gamma_i \,, \quad \det(A) = \Pi_{i=1}^{n} \gamma_i \,.$$

(Eigenwerte mit algebraischer Vielfachheit ≥ 2 werden in der Liste mehrfach aufgezählt.)

Beweis. Teil (1) ist die direkte Folge von Satz 7.2.12.
Teil (2): Es sei also das charakteristische Polynom von A

$$\det(\lambda - A) = (\lambda - \gamma_1)(\lambda - \gamma_2) \cdots (\lambda - \gamma_n) \,.$$

Ausmultiplizieren und sortieren nach Potenzen von λ gibt für

$$\begin{aligned} -\,\mathrm{trace}(A) &= c_{n-1} = -\gamma_1 - \gamma_2 - \cdots - \gamma_n \,, \\ (-1)^n \det(A) &= c_0 = (-\gamma_1)(-\gamma_2) \cdots (-\gamma_n) \,. \end{aligned}$$

□

Beispiel 7.2.17. Bestimmen Sie die Eigenwerte, Eigenräume, und die algebraischen und geometrischen Vielfachheiten der Eigenwerte der Matrix A:

$$A = \begin{pmatrix} -1 & -1 & -3 & 0 \\ 0 & -1 & -6 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 6 & 1 \end{pmatrix}$$

Lösung: Das charakteristische Polynom ermittelt man in diesem Fall am bequemsten durch den Laplaceschen Entwicklungssatz nach der vierten Spalte, dann hat man eine Dreiecksmatrix:

$$\begin{vmatrix} \lambda+1 & 1 & 3 & 0 \\ 0 & \lambda+1 & 6 & 0 \\ 0 & 0 & \lambda-1 & 0 \\ 0 & -2 & -6 & \lambda-1 \end{vmatrix} = (\lambda-1) \begin{vmatrix} \lambda+1 & 1 & 3 \\ 0 & \lambda+1 & 6 \\ 0 & 0 & \lambda-1 \end{vmatrix}$$

$$= (\lambda-1)^2(\lambda+1)^2 \,.$$

Daher besitzt A die Eigenwerte ± 1, jeweils mit der algebraischen Vielfachheit 2.
Wir berechnen jetzt den Eigenraum für $\lambda = 1$, also den Kern von $(1 - A)$, durch Pivotschritte:

$$\begin{pmatrix} 2 & 1 & 3 & 0 \\ 0 & 2 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2* & -6 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2* & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2* & -6 & 0 \end{pmatrix}$$

Die dritte und vierte Spalte bleiben nichtbasisch. Wir setzen $x_3 = s$, $x_4 = t$ und erhalten

$$\ker(1 - A) = \{ s \begin{pmatrix} 0 \\ -3 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \mid s, t \in \mathbb{R} \} \,.$$

Der Eigenraum ist zweidimensional, der Eigenwert 1 hat auch die geometrische Vielfachheit 2.

Letztlich berechnen wir den Eigenraum zu $\lambda = -1$, also den Kern von $(-1 - A)$:

$$\begin{pmatrix} 0 & 1* & 3 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & -2 & -6 & -2 \end{pmatrix} \to \begin{pmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & -2* & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \to \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2* & 0 \\ 0 & 0 & 0 & -2* \end{pmatrix}$$

Nichtbasisch bleibt nur die erste Spalte, der Kern hat die Dimension 1:

$$\ker(-1 - A) = \{t \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \mid t \in \mathbb{R}\}\,.$$

Der Eigenwert -1 hat die algebraische Vielfachheit 2 und die geometrische Vielfachheit 1. □

7.3. **Diagonalisierung.**

7.3.1. *Lineare Unabhängigkeit von Eigenvektoren.*

Das Matrizenprodukt ist nicht kommutativ, aber es gilt immer:

Lemma 7.3.1. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix, und $\lambda, \mu \in \mathbb{K}$. Dann gilt:*

$$(\lambda - A)(\mu - A) = (\mu - A)(\lambda - A)\,.$$

Beweis. Wir müssen nur die Klammern ausmultiplizieren:

$$\begin{aligned} (\lambda - A)(\mu - A) &= \lambda\mu E - (\lambda E)A - A(\mu E) + A^2 \\ = \quad \mu\lambda E - A(\lambda E) - (\mu E)A + A^2 &= (\mu - A)(\lambda - A)\,. \end{aligned}$$

□

Satz 7.3.2. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix.*

1) *Sind $\lambda \neq \mu \in \mathbb{K}$, $r, s \in \mathbb{N} \setminus \{0\}$, und ist $v \in \mathbb{K}^n$ so, dass*
$$(\lambda - A)^r v = (\mu - A)^s v = 0\,.$$
Dann ist $v = 0$.
2) *Sind $\lambda_1, \cdots, \lambda_k \in \mathbb{K}$ paarweise verschieden, $r_1, \cdots, r_k \in \mathbb{N} \setminus \{0\}$, sind $v_1, \cdots, v_k \in \mathbb{K}^n \setminus \{0\}$ so, dass für alle $i = 1, \cdots k$ gilt*
$$(\lambda_i - A)^{r_i} v_i = 0\,.$$
Dann ist das System $(v_1, \cdots, v_k)$ linear unabhängig.
3) *Sind $\lambda_1, \cdots, \lambda_k \in \mathbb{K}$ paarweise verschieden, $r_1, \cdots, r_k \in \mathbb{N} \setminus \{0\}$. Dann ist der Summenraum*
$$\ker(\lambda_1 - A)^{r_1} + \cdots + \ker(\lambda_k - A)^{r_k}$$
eine direkte Summe.

Beweis. Teil (1): Vollständige Induktion nach $r + s$.
Induktionsverankerung: $r + s = 2$, d.h., $r = s = 1$. Sei $(\lambda - A)v = (\mu - A)v = 0$. Dann ist $(\lambda - \mu)v = (\lambda - A)v - (\mu - A)v = 0$. Weil $\lambda - \mu \neq 0$, folgt $v = 0$.
Induktionsschritt: Sei der Satz richtig, falls $r + s \leq m$, und sei nun $(\lambda - A)^r v = (\mu - A)^s v = 0$ mit $r + s = m + 1$. OBdA sei $r > 1$. Wir setzen $w = (\lambda - A)v$. Es ist dann

$$\begin{aligned} (\lambda - A)^{r-1} w &= (\lambda - A)^r v = 0\,, \\ (\mu - A)^s w &= (\mu - A)^s (\lambda - A)v = (\lambda - A)(\mu - A)^s v = 0\,. \end{aligned}$$

Da $(r-1)+s=m$ folgt aus der Induktionsannahme, dass $w=(\lambda-A)v=0$. Nun ist $(\lambda-A)v=(\mu-A)^s v=0$, und weil $1+s\leq m$, gilt wieder nach der Induktionsannahme $v=0$.
Teil (2): Vollständige Induktion nach k. Für $k=1$ ist die Aussage trivial, weil $v_1\neq 0$.
Induktionsschritt: Sei der Satz richtig für k, und seien für $i=1,\cdots,k+1$ die Vektoren $v_i\in\ker((\lambda_i-A)^{r_i})\setminus\{0\}$. Für $i=1,\cdots,k$ setzen wir $w_i=(\lambda_{k+1}-A)^{r_{k+1}}v_i$. Wäre eines der $w_i=0$, so wäre

$$(\lambda_{k+1}-A)^{r_{k+1}}v_i=w_i=0=(\lambda_i-A)^{r_i}v_i\,,$$

und wegen Teil (1) wäre dann $v_i=0$ im Widerspruch zur Annahme. Daher sind alle $w_i\neq 0$. Es ist dann für $i=1\cdots k$

$$(\lambda_i-A)^{r_i}w_i=(\lambda_i-A)^{r_i}(\lambda_{k+1}-A)^{r_{k+1}}v_i=(\lambda_{k+1}-A)^{r_{k+1}}(\lambda_i-A)^{r_i}v_i=0\,,$$

also ist $w_i\in\ker((\lambda_i-A)^{r_i})$.
Sei nun

$$\sum_{i=1}^{k+1}\alpha_i v_i=0\,.$$

Multiplikation mit $(\lambda_{r+1}-A)^{r_{k+1}}$ ergibt

$$\sum_{i=1}^{k}\alpha_i w_i+0=0\,,$$

und weil $(w_1,\cdots,w_k)$ wegen der Induktionsannahme linear unabhängig sind, folgt $\alpha_1=\cdots=\alpha_k=0$. Weil $v_{k+1}\neq 0$, muss nun auch $\alpha_{k+1}=0$ sein, und das System $(v_1,\cdots,v_{k+1})$ ist linear unabhängig.
Teil (3): Wegen Teil (2) und Satz 3.5.7 (1) ist der Summenraum $\ker((\lambda_1-A)^{r_1})+\cdots+\ker((\lambda_k-A)^{r_k})$ eine direkte Summe. □

Korollar 7.3.3. *Sei $\mathbb{K}$ ein Körper und $A\in\mathbb{K}^{n\times n}$ eine quadratische Matrix. Dann gilt:*

1) *Jedes System $(v_1,\cdots,v_k)$ von Eigenvektoren von A zu paarweise verschiedenen Eigenwerten $\lambda_1,\cdots,\lambda_k$ ist linear unabhängig.*
2) *Seien $\lambda_1,\cdots,\lambda_k$ die Eigenwerte von A mit den* geometrischen *Vielfachheiten $m_1,\cdots,m_k$. Der Raum $\mathbb{K}^n$ besitzt genau dann eine Basis aus Eigenvektoren von A, wenn $m_1+\cdots+m_k=n$.*
3) *Wenn es n verschiedene Eigenwerte $\lambda_1,\cdots,\lambda_n$ von A gibt, und $v_1,\cdots,v_n$ Eigenvektoren zu $\lambda_1,\cdots,\lambda_n$ sind, dann ist $(v_1,\cdots,v_n)$ eine Basis von $\mathbb{K}^n$.*

Beweis. Teil (1): Satz 7.3.2 (2) mit $r_1=\cdots=r_k=1$.
Teil (2): Seien $W_i=\ker((\lambda_i-A))$ die Eigenräume von A. Wegen Satz 7.3.2 (3) ist die Summe $W_1+\cdots+W_k$ direkt. Wegen Satz 3.5.7 (4) ist die Dimension des Summenraums gleich $m_1+\cdots+m_k$. Insbesondere ist $W_1+\cdots+W_k=\mathbb{K}^n$ genau dann wenn $m_1+\cdots+m_k=n$.
Besitzt $\mathbb{K}^n$ eine Basis aus Eigenvektoren, so ist offensichtlich $\mathbb{K}^n=W_1+\cdots+W_k$. Sei umgekehrt $\mathbb{K}^n=W_1+\cdots+W_k$. Für jeden Eigenraum W_i wählen wir eine Basis $(v_{i,1},\cdots,v_{i,m_i})$, diese besteht natürlich aus Eigenvektoren zum Eigenwert λ_i. Wegen Satz 3.5.7 (3) ist die Vereinigung der Basen, also $(v_{1,1},\cdots,v_{k,m_k})$, eine Basis des Summenraums, folglich von $\mathbb{K}^n$.
Teil (3): Wegen Teil (1) sind $(v_1,\cdots,v_n)$ linear unabhängig, und da es n Vektoren sind, bilden sie eine Basis des $\mathbb{K}^n$. □

Versuchen Sie selbst, Korollar 7.3.3 direkt zu beweisen, ohne auf Satz 7.3.2 zurückzugreifen. Das ist etwas einfacher, weil keine Potenzen r_i ins Spiel kommen, und es würde auch für diesen Abschnitt

völlig ausreichen. Später werden wir aber brauchen können, dass wir mit Kernen von $(\lambda - A)^r$ statt nur von $(\lambda - A)$ arbeiten können.

7.3.2. *Diagonalisierung einer Matrix.*

Wir wiederholen vor diesem Unterabschnitt kurz den Abschnitt 4.4 über Koordinaten und Koordinatentransformationen. Wir werden nämlich anschließend näher darauf eingehen, dass bestimmte lineare Abbildungen des $f : \mathbb{K}^n \to \mathbb{K}^n$ bei geeigneter Wahl der Basis durch Diagonalmatrizen beschrieben werden.

Wenn in einem endlich dimensionalen Vektorraum V eine Basis $B = (b_1, \cdots, b_n)$ gegeben ist, definieren wir den Isomorphismus

$$\Phi_B : \begin{cases} \mathbb{K}^n & \to V\,, \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} & \mapsto \sum_{i=1}^n \xi_i b_i\,. \end{cases}$$

Für $x \in V$ ist dann $\Phi_B^{-1}(x)$ der Koordinatenvektor bezüglich der Basis B. Sind $B = (b_1, \cdots, b_n)$ und $C = (c_1, \cdots, c_n)$ zwei Basen in V, so berechnet man den Koordinatenvektor von x bezüglich B (also $\Phi_B^{-1}(x)$) aus dem Koordinatenvektor von x bezüglich C (also $\Phi_C^{-1}(x)$) durch die sogenannte Übergangsmatrix T_B^C:

$$\Phi_C^{-1}(x) = T_C^B \Phi_B^{-1}(x)\,.$$

Die Spalten von T_B^C erhält man, indem man die Basisvektoren $c_1, \cdots, c_n$ in B-Koordinaten schreibt, also $\Phi_B^{-1}(c_1), \cdots, \Phi_B^{-1}(c_n)$. Die Übergangsmatrix von B-Koordinaten auf C-Koordinaten ist

$$T_C^B = (T_B^C)^{-1}\,.$$

In diesem Abschnitt werden wir uns nur mit Endomorphismen beschäftigen, also linearen Abbildungen eines Vektorraumes V in sich selbst. Ist nun $B = (b_1, \cdots, b_n)$ wiederum eine Basis eines endlich dimensionalen Vektorraumes, und $f : V \to V$ ein Endomorphismus, so wird f durch eine Matrix $M_B^B(f)$ beschrieben: Die B-Koordinaten von $f(x)$ ergeben sich aus den B-Koordinaten von x durch

$$\Phi_B^{-1}(f(x)) = M_B^B \Phi_B^{-1}(x)\,.$$

Ist $V = \mathbb{K}^n$ und $E = (e_1, \cdots, e_n)$ die Einheitsbasis von $\mathbb{K}^n$, ist $A \in \mathbb{K}^{n\times n}$ und f die lineare Abbildung $f(x) = Ax$, so ist die Matrix von f bezüglich der Einheitsbasis einfach $M_E^E(f) = A$.

Sind B und C zwei Basen von V, und kennt man die Matrix $M_C^C(f)$ eines Endomorphismus bezüglich der Basis C, so erhält man die B-Koordinaten von $f(x)$ aus den B-Koordinaten von x, indem man zuerst auf C-Koordinaten umrechnet, dann $M_C^C(f)$ anwendet, und am Ende wieder auf B-Koordinaten zurückrechnet, wie es das folgende Diagramm ausdrückt:

$$\begin{array}{ccccccc} V & \xrightarrow{\mathrm{id}_V} & V & \xrightarrow{f} & V & \xrightarrow{\mathrm{id}_V} & V \\ \uparrow{\scriptstyle \Phi_B} & & \uparrow{\scriptstyle \Phi_C} & & \uparrow{\scriptstyle \Phi_C} & & \uparrow{\scriptstyle \Phi_B} \\ \mathbb{K}^n & \xrightarrow{T_C^B} & \mathbb{K}^n & \xrightarrow{M_C^C(f)} & \mathbb{K}^n & \xrightarrow{T_B^C} & \mathbb{K}^n \end{array}$$

Für die Rechnung heißt das:

$$M_B^B(f) = T_B^C M_C^C(f) T_C^B = (T_C^B)^{-1} M_C^C(f) T_C^B\,.$$

Quadratische Matrizen X und Y, die mit einer regulären Matrix T die Gleichung

$$Y = T^{-1} X T$$

erfüllen, heißen ähnlich. Zwei Matrizen sind genau dann die Matrizen eines Endomorphismus bezüglich verschiedener Koordinatensysteme, wenn sie ähnlich sind.

Definition 7.3.4. Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. A heißt genau dann diagonalisierbar, wenn es eine reguläre Matrix T und eine Diagonalmatrix D gibt, sodass $T^{-1}AT = D$ (d.h., wenn A ähnlich zu einer Diagonalmatrix ist).

Satz 7.3.5. *Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$. Sei T eine reguläre Matrix mit den Spalten $v_1, \cdots, v_n$, und sei D eine Diagonalmatrix, auf deren Diagonalen die Elemente $\lambda_1, \cdots, \lambda_n$ (nicht unbedingt paarweise verschieden) stehen. Dann gilt:*
$D = T^{-1}AT$ genau dann, wenn jedes v_i ein Eigenvektor von A zum Eigenwert λ_i ist.
Es ist also A genau dann diagonalisierbar, wenn $\mathbb{K}^n$ eine Basis aus Eigenvektoren von A hat.

Beweis. Weil T regulär ist, ist $T^{-1}AT = D$ genau dann, wenn $AT = TD$. (Wir müssen nur von links mit T multiplizieren.) Die Spalten von AT sind $Av_1, \cdots, Av_n$, die Spalten von TD sind $\lambda_1 v_1, \cdots, \lambda_n v_n$. Also ist $T^{-1}AT = D$ genau dann, wenn $Av_i = \lambda_i v_i$ für alle i gilt. □

Korollar 7.3.6. *Sei $\mathbb{K}$ ein Körper und $A \in \mathbb{K}^{n\times n}$. Sei $f : \mathbb{K}^n \to \mathbb{K}^n$ der Endomorphismus $f(x) = Ax$. Sei $B = (v_1, \cdots, v_n)$ eine Basis des $\mathbb{K}^n$ aus Eigenvektoren von A zu den Eigenwerten $\lambda_1, \cdots, \lambda_n$. Mit $E = (e_1, \cdots, e_n)$ bezeichnen wir die Einheitsbasis des $\mathbb{K}^n$. Dann gilt:*

1) *Die Spalten der Übergangsmatrix T_E^B sind die Eigenvektoren $v_1, \cdots, v_n$.*
2) *Die Matrix des Endomorphismus f bezüglich der Basis B ist*

$$M_B^B(f) = (T_E^B)^{-1} A T_E^B = D\,.$$

Beweis. Die Spalten von T_E^B sind die Basisvektoren von B, in Einheitskoordinaten geschrieben, also $v_1, \cdots, v_n$. Weil diese Eigenvektoren zu den Eigenwerten $\lambda_1, \cdots, \lambda_n$ sind, gilt wegen Satz 7.3.5: $(T_E^B)^{-1} A T_E^B = D$. □

Beispiel 7.3.7. Diagonalisieren Sie die Matrix

$$A = \begin{pmatrix} -1 & 2 & 4 \\ 4 & -3 & -8 \\ -2 & 2 & 5 \end{pmatrix}.$$

Lösung: Berechnung des charakteristischen Polynoms und der Eigenwerte.

$$\begin{aligned} c_2 &= -\operatorname{trace}(A) = -(-1-3+5) = -1\,, \\ c_1 &= \begin{vmatrix} -1 & 2 \\ 4 & -3 \end{vmatrix} + \begin{vmatrix} -1 & 4 \\ -2 & 5 \end{vmatrix} + \begin{vmatrix} -3 & -8 \\ 2 & 5 \end{vmatrix} = -5+3+1 = -1\,, \\ c_0 &= -\det(A) = 1\,. \end{aligned}$$

Damit ist das charakteristische Polynom

$$\lambda^3 - \lambda^2 - \lambda + 1 = (\lambda-1)(\lambda^2-1) = (\lambda-1)^2(\lambda+1)\,.$$

Daher hat A die Eigenwerte -1 mit algebraischer Vielfachhheit 1 und $+1$ mit algebraischer Vielfachheit 2.
Eigenräume: Berechnung von $\ker(-1-A)$.

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

$$\begin{pmatrix} 0 & -2 & -4 \\ -4 & 2 & 8 \\ 2* & -2 & -6 \end{pmatrix} \to \begin{pmatrix} 0 & -2 & -4 \\ 0 & -2* & -4 \\ 2 & -2 & -6 \end{pmatrix} \to \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2* & -4 \\ 2 & 0 & -2 \end{pmatrix}.$$

Nichtbasisch bleibt nur die dritte Variable, die wir nun mit 1 ansetzen. Wir erhalten

$$\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Berechnung von $\ker(1-A)$.

$$(1-A) = \begin{pmatrix} 1+1 & -2 & -4 \\ -4 & 1+3 & 8 \\ 2 & -2 & 1-5 \end{pmatrix} = \begin{pmatrix} 2 & -2 & -4 \\ -4 & 4 & 8 \\ 2 & -2 & -4 \end{pmatrix}.$$

$$\begin{pmatrix} 2* & -2 & -4 \\ -4 & 4 & 8 \\ 2 & -2 & -4 \end{pmatrix} \to \begin{pmatrix} 2 & -2 & -4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Nichtbasisch bleiben die zweite und dritte Variable. Wir setzen jeweils eine davon mit 1, die andere mit 0 an, und erhalten eine Basis des Kerns:

$$(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}).$$

Diagonalisierung: Damit besitzt $\mathbb{R}^3$ eine Basis aus 3 linear unabhängigen Eigenvektoren von A:

$$B = (\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}).$$

Die Übergangsmatrix T_E^B ist dann die Matrix, deren Spalten die Basisvektoren sind:

$$T_E^B = \begin{pmatrix} 1 & 1 & 2 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Aus den Eigenwerten ergibt sich die Diagonalmatrix

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

Diese Matrix stellt den Endomorphismus $f(x) = Ax$ bezüglich der Basis B dar. Man kann sich leicht durch direktes Nachrechnen (MATLAB?) überzeugen, dass tatsächlich gilt:

$$M_B^B(f) = (T_E^B)^{-1} A T_E^B = D.$$

□

Satz 7.3.8. *Ähnliche Matrizen haben dasselbe charakteristische Polynom.*

Beweis. Weil $\det(T)\det(T^{-1}) = 1$, gilt

$$\begin{aligned} & \det(\lambda - T^{-1}AT) = \det(\lambda T^{-1}T - T^{-1}AT) = \det(T^{-1}(\lambda - A)T) \\ = \quad & \det(T^{-1})\det(\lambda - A)\det(T) = \det(\lambda - A). \end{aligned}$$

□

7.3.3. *Anwendungsbeispiele.*

Beispiel 7.3.9. Berechnen Sie die Potenzen A^m für $m \in \mathbb{N}$ und

$$A = \begin{pmatrix} 5 & -2 \\ 7 & -4 \end{pmatrix} .$$

Lösung: Wenn A diagonalisierbar sein sollte, also $D = T^{-1}AT$, so wäre $A = TDT^{-1}$ und

$$\begin{aligned} & A^m = (TDT^{-1})(TDT^{-1}) \cdots (TDT^{-1}) \\ = \quad & TD(T^{-1}T)D(T^{-1}T) \cdots (T^{-1}T)DT^{-1} = TD^mT^{-1} . \end{aligned}$$

Die Potenzen einer Diagonalmatrix sind aber leicht zu berechnen:

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^m = \begin{pmatrix} \lambda_1^m & 0 \\ 0 & \lambda_2^m \end{pmatrix} .$$

Wir versuchen also, A zu diagonalisieren.
Das charakteristische Polynom ist

$$\lambda^2 - (5-4)\lambda + \begin{vmatrix} 5 & -2 \\ 7 & -4 \end{vmatrix} = \lambda^2 - \lambda - 6 .$$

Die Eigenwerte sind $\lambda_1 = -2$ und $\lambda_2 = 3$. Weil A also zwei verschiedene Eigenwerte besitzt, hat $\mathbb{R}^2$ eine Basis aus Eigenvektoren, und A ist diagonalisierbar. Wir berechnen die Eigenräume:

$$\begin{aligned} \ker(-2 - A) &= \ker \begin{pmatrix} -7 & 2 \\ -7 & 2 \end{pmatrix} = \mathcal{L}(\begin{pmatrix} 2 \\ 7 \end{pmatrix}) , \\ \ker(3 - A) &= \ker \begin{pmatrix} -2 & 2 \\ -7 & 7 \end{pmatrix} = \mathcal{L}(\begin{pmatrix} 1 \\ 1 \end{pmatrix}) . \end{aligned}$$

Wir haben also $D = T^{-1}AT$ mit

$$D = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix} , \quad T = \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix} , \quad T^{-1} = \begin{pmatrix} -0,2 & 0,2 \\ 1,4 & -0,4 \end{pmatrix} .$$

Dann ist

$$\begin{aligned} & A^m = TD^mT^{-1} = \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} (-2)^m & 0 \\ 0 & 3^m \end{pmatrix} \begin{pmatrix} -0,2 & 0,2 \\ 1,4 & -0,4 \end{pmatrix} \\ = \quad & (-2)^m \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -0,2 & 0,2 \\ 1,4 & -0,4 \end{pmatrix} + 3^m \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -0,2 & 0,2 \\ 1,4 & -0,4 \end{pmatrix} \\ = \quad & (-2)^m \begin{pmatrix} -0,4 & 0,4 \\ -1,4 & 1,4 \end{pmatrix} + 3^m \begin{pmatrix} 1,4 & -0,4 \\ 1,4 & -0,4 \end{pmatrix} . \end{aligned}$$

□

Ist eine $n \times n$-Matrix diagonalisierbar mit Eigenwerten $\lambda_1, \cdots, \lambda_k$, so ist $A^m = \lambda_1^m C_1 + \cdots + \lambda_k^m C_k$ mit Matrizen $C_1, \cdots, C_k$, welche von m nicht abhängen.

Beispiel 7.3.10. Lösen Sie die folgende Differentialgleichung

$$\frac{d}{dt}\, x(t) = \begin{pmatrix} 5 & -2 \\ 7 & -4 \end{pmatrix} x(t) , \quad x(0) = x_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} .$$

Ausgeschrieben haben wir ein System von zwei Differentialgleichungen in zwei unbekannten Funktionen:

$$\begin{aligned} \frac{d}{dt}\, x_1(t) &= 5x_1(t) - 2x_2(t) , \quad x_1(0) = 1 , \\ \frac{d}{dt}\, x_2(t) &= 7x_1(t) - 4x_2(t) , \quad x_2(0) = 2 . \end{aligned}$$

Lösung: Die Matrix aus dem Gleichungssystem ist die Matrix A aus Beispiel 7.3.9. Wir wissen aus diesem Beispiel, dass A diagonalisierbar ist: $D = T^{-1}AT$ mit

$$D = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}, \quad T = \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} -0,2 & 0,2 \\ 1,4 & -0,4 \end{pmatrix}.$$

Wir setzen nun

$$y(t) := T^{-1}x(t), \quad \text{also } x(t) = Ty(t).$$

Nun ist

$$\frac{d}{dt}y(t) = \frac{d}{dt}T^{-1}x(t) = T^{-1}Ax(t) = T^{-1}ATy(t) = Dy(t).$$

Wir erhalten also das System

$$\frac{d}{dt}y(t) = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix} y(t), \quad y(0) = T^{-1}x(0) = \begin{pmatrix} 0,2 \\ 0,6 \end{pmatrix}.$$

Ausgeschrieben ergeben sich zwei entkoppelte (das heißt, voneinander unabhängige) Differentialgleichungen:

$$\begin{aligned} \frac{d}{dt}y_1(t) &= -2y_1(t), \quad y_1(0) = 0,2, \\ \frac{d}{dt}y_2(t) &= 3y_2(t), \quad y_2(0) = 0,6. \end{aligned}$$

Die Lösungen sind

$$y_1(t) = 0,2\,e^{-2t}, \quad y_2(t) = 0,6\,e^{3t}.$$

Dann ist

$$x(t) = Ty(t) = \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} 0,2\,e^{-2t} \\ 0,6\,e^{3t} \end{pmatrix} = e^{-2t} \begin{pmatrix} 0,4 \\ 1,4 \end{pmatrix} + e^{3t} \begin{pmatrix} 0,6 \\ 0,6 \end{pmatrix}.$$

□

Ist A diagonalisierbar mit Eigenwerten $\lambda_1, \cdots, \lambda_k$, so schreibt sich jede Lösung der Differentialgleichung $x'(t) = Ax(t)$ in der Form

$$x(t) = e^{\lambda_1 t}v_1 + \cdots + e^{\lambda_k t}v_k$$

mit Vektoren $v_i \in \ker(\lambda_i - A)$, welche von t nicht abhängen.

7.4. **Hauptraumzerlegung.**

7.4.1. *Invariante Unterräume.*

Definition 7.4.1.

1) Sei V ein Vektorraum, W ein Unterraum von V, und $f : V \to V$ ein Endomorphismus. Dann heißt W invariant bezüglich f, wenn gilt: $f(W) \subset W$.
2) Sei $\mathbb{K}$ ein Körper, W ein Unterraum von $\mathbb{K}^n$, und $A \in \mathbb{K}^{n\times n}$. Dann heißt W invariant bezüglich A, wenn gilt: $(\forall w \in W)\, Aw \in W$.

Bemerkung 7.4.2. Sei $A \in \mathbb{K}^{n\times n}$. Man sieht leicht, dass die folgenden Räume A-invariant sind:

$\{0\}$,
$\mathbb{K}^n$,
$\ker(A)$,
$\mathrm{rg}(A)$.

Außerdem gilt: Ist W invariant bezüglich A, so ist W auch invariant bezüglich jeder Potenz A^m und bezüglich $(\lambda - A)$ für jeden Skalar λ.

Bemerkung 7.4.3. Sei V ein Vektorraum, $f : V \to V$ ein Endomorphismus, und $W_1, \cdots, W_k$ f-invariante Unterräume von V. Dann sind auch Summenraum $W_1 + \cdots + W_k$ und Durchschnitt $\bigcap_{i=1}^k W_i$ f-invariante Unterräume.

Lemma 7.4.4. *Ist V ein Vektorraum, $f : V \to V$ ein Endomorphismus, und W ein f-invarianter Unterraum von V, dann ist die Einschränkung*

$$f\,|_W \colon \begin{cases} W & \to W \\ x & \mapsto f(x) \end{cases}$$

ein Endomorphismus auf W.

Beweis. Weil f den Raum W nach Voraussetzung in W abbildet, ist $f\,|_W$ als Abbildung von W nach W wohldefiniert. Die Linearität erbt die Einschränkung natürlich von f. □

Satz 7.4.5. *Sei V ein endlich dimensionaler Vektorraum, $f : V \to V$ ein Endomorphismus. Seien $W_1, \cdots, W_k$ f-invariante Unterräume von V, sodass V die direkte Summe ist: $V = W_1 + \cdots + W_k$. Für $i = 1, \cdots, k$ sei $(b_{i,1}, \cdots, b_{i,m_i})$ eine Basis von W_i, und B sei die Basis $(b_{1,1}, b_{1,2}, \cdots, b_{k,m_k})$ von V, die sich als Vereinigung der Basen der W_i ergibt. Dann hat die Matrix $M_B^B(f)$ von f bezüglich B Blockdiagonalform (wir indizieren diesmal, der besseren Übersichtlichkeit wegen, die Zeilen und Spalten doppelt, und schreiben die Indices an den Rand der Matrix):*

$$\begin{pmatrix}
 & (1) & \cdots & (m_1) & (1) & \cdots & (m_2) & \cdots \\
(1) & * & \cdots & * & 0 & \cdots & 0 & \cdots \\
 & \vdots & & \vdots & \vdots & & \vdots & \\
(m_1) & * & \cdots & * & 0 & \cdots & 0 & \cdots \\
(1) & 0 & \cdots & 0 & * & \cdots & * & \cdots \\
 & \vdots & & \vdots & \vdots & & \vdots & \\
(m_2) & 0 & \cdots & 0 & * & \cdots & * & \cdots \\
 & \vdots & & \vdots & \vdots & & \vdots & \ddots
\end{pmatrix}$$

Beweis. Die Spalten von $M_B^B(f)$ sind die Bilder der Basisvektoren $b_{1,1}, \cdots, b_{k,m_k}$, geschrieben in Koordinaten bezüglich B. Jeder Basisvektor $b_{i,j}$ liegt in W_i, und weil W_i invariant ist, ist auch $f(b_{i,j}) \in W_i$. Damit kann also $f(b_{i,j})$ als Linearkombination von $(b_{i,1}, \cdots, b_{i,m_i})$ geschrieben werden, ohne die anderen Vektoren von B heranzuziehen. □

7.4.2. *Lemma von Fitting.*

Satz 7.4.6 (Lemma von Fitting)**.** *Sei $\mathbb{K}$ ein Körper, $F \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. Dann gilt:*

1)
$$\{0\} \subset \ker(F) \subset \ker(F^2) \subset \cdots,$$
$$\mathbb{K}^n \supset \operatorname{rg}(F) \supset \operatorname{rg}(F^2) \supset \cdots.$$

2) *Es gibt ein $r \leq n$ sodass $\ker(F^r) = \ker(F^{r+1})$.*
3) *Sei $r \in \mathbb{N}$. Folgende Aussagen sind äquivalent:*
   (i) $\ker(F^r) = \ker(F^{r+1})$,
   (ii) $\operatorname{rg}(F^r) = \operatorname{rg}(F^{r+1})$,
   (iii) $(\forall s > r)\ \ker(F^r) = \ker(F^s)$,
   (iv) $(\forall s > r)\ \operatorname{rg}(F^r) = \operatorname{rg}(F^s)$,
   (v) $\ker(F^r) \cap \operatorname{rg}(F^r) = \{0\}$,
   (vi) $\mathbb{K}^n$ *ist die direkte Summe* $\ker(F^r) + \operatorname{rg}(F^r)$.

Beweis. Teil (1): Ist $x \in \ker(F^i)$, so ist $F^{i+1}x = FF^ix = F0 = 0$, also $x \in \ker(F^{i+1})$. Ist $x \in \mathrm{rg}(F^{i+1})$, so gibt es ein $y \in \mathbb{K}^n$ sodass $x = F^{i+1}y$. Dann ist aber $x = F^iFy \in \mathrm{rg}(F^i)$.

Teil (2): Wäre $\ker(F^r) \neq \ker(F^{r+1})$ für $r = 0, \cdots, n$, so wäre

$$0 < \dim(\ker(F)) < \dim(\ker(F^2)) < \cdots < \dim(\ker(F^{n+1}))\,,$$

und es würde folgen $\dim(\ker(F^{n+1})) > n$, was für einen Unterraum des $\mathbb{K}^n$ nicht möglich ist.

Teil (3):

(i) $\Leftrightarrow$ (ii): Da $\ker(F^r) \subset \ker(F^{r+1})$, gilt

$$\ker(F^r) = \ker(F^{r+1}) \;\Leftrightarrow\; \dim(\ker(F^r)) = \dim(\ker(F^{r+1}))\,.$$

Nun gilt aber:

$$\dim(\mathrm{rg}(F^r)) = n - \dim(\ker(F^r))\,,$$

sodass

$$\dim(\ker(F^r)) = \dim(\ker(F^{r+1})) \;\Leftrightarrow\; \dim(\mathrm{rg}(F^r)) = \dim(\mathrm{rg}(F^{r+1}))\,.$$

Das ist wiederum äquivalent zu $\mathrm{rg}(F^r) = \mathrm{rg}(F^{r+1})$, weil $\mathrm{rg}(F^r) \supset \mathrm{rg}(F^{r+1})$.

(iii) $\Leftrightarrow$ (iv) ebenso.

(i) $\Leftrightarrow$ (iii): (iii)$\Rightarrow$(i) ist trivial. Wir nehmen an, dass (i) gilt, nicht (iii). Sei s die kleinste Zahl für die gilt: $s > r$ und $\ker(F^s) \neq \ker(F^{s+1})$. Es ist also $\ker(F^{s-1}) = \ker(F^s)$. Sei $x \in \ker(F^{s+1}) \setminus \ker(F^s)$. Dann ist $F(x) \in \ker(F^s) \setminus \ker(F^{s-1}) = \emptyset$, und wir erhalten einen Widerspruch.

(ii) $\Leftrightarrow$ (v): Wir betrachten den Homomorphismus

$$f : \begin{cases} \mathrm{rg}(F^r) & \to \mathrm{rg}(F^r)\,, \\ x & \mapsto F^rx\,. \end{cases}$$

Klarerweise ist $\ker(f) = \ker(F^r) \cap \mathrm{rg}(F^r)$. Da f ein Endomorphismus des endlichdimensionalen Vektorraumes $\mathrm{rg}(F^r)$ ist, ist f genau dann injektiv, wenn f surjektiv ist. Also gilt:

$$\begin{aligned} &\ker(F^r) \cap \mathrm{rg}(F^r) = \{0\} \;\Leftrightarrow\; f \text{ injektiv} \;\Leftrightarrow\; f \text{ surjektiv} \\ \Leftrightarrow\quad &\mathrm{rg}(F^r) = \mathrm{rg}(F^{2r}) \;\Leftrightarrow\; \mathrm{rg}(F^r) = \mathrm{rg}(F^{r+1}). \end{aligned}$$

(v) $\Leftrightarrow$ (vi): (vi)$\Rightarrow$(v) ist offensichtlich. Gilt (v), so ist die Summe $\ker(F^r) + \mathrm{rg}(F^r)$ direkt. Daher ist die Dimension dieses Summenraumes $(n - \rho(F^r)) + \rho(F^r) = n$, und der Summenraum ist ganz $\mathbb{K}^n$. □

7.4.3. *Verallgemeinerte Eigenräume.*

Definition 7.4.7. Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{n\times n}$, $\lambda \in \mathbb{K}$ ein Eigenwert von A. Der verallgemeinerte Eigenraum (oder Hauptraum) von A zum Eigenwert λ ist

$$\ker(\lambda - A)^n.$$

Bemerkung 7.4.8. Wegen Satz 7.4.6 ist der Hauptraum von A zum Eigenwert λ zugleich der größtmögliche Raum, der sich als $\ker(\lambda - A)^r$ mit irgendeiner Potenz r schreiben läßt. Wenn gilt

$$0 \subsetneq \ker(\lambda - A) \subsetneq \ker(\lambda - A)^2 \subsetneq \cdots \subsetneq \ker(\lambda - A)^r = \ker(\lambda - A)^{r+1}\,,$$

dann ist $\ker(\lambda - A)^r$ der Hauptraum.

Lemma 7.4.9. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{n\times n}$, $\lambda_1, \cdots, \lambda_k$ (nicht notwendigerweise alle) Eigenwerte von A, $\ker((\lambda_i - A)^{r_i})$ die Haupträume dazu. Dann gilt:*
Ist $x \in \bigcap_{i=1}^{k} \operatorname{rg}(\lambda_i - A)^{r_i}$ und $(\lambda_1 - A)^{r_1} \cdots (\lambda_k - A)^{r_k} x = 0$, dann ist $x = 0$.

Beweis. Vollständige Induktion nach k. Für $k = 1$ ist nach Satz 7.4.6 der Durchschnitt $\ker(\lambda_1 - A)^{r_1} \cap \operatorname{rg}(\lambda_1 - A)^{r_1} = \{0\}$. Damit ist die Induktion verankert. Induktionsschritt von k auf $k+1$: Sei nun

$$\begin{aligned} x &\in \bigcap_{i=1}^{k+1} \operatorname{rg}(\lambda_i - A)^{r_i}\,, \\ (\lambda_1 - A)^{r_1} &\cdots (\lambda_{k+1} - A)^{r_{k+1}} x = 0\,. \end{aligned}$$

Wir setzen $y = (\lambda_{k+1} - A)^{r_{k+1}} x$. Man sieht leicht, dass

$$\begin{aligned} y &\in \bigcap_{i=1}^{k} \operatorname{rg}(\lambda_i - A)^{r_i}\,, \\ (\lambda_1 - A)^{r_1} &\cdots (\lambda_k - A)^{r_k} y = 0\,. \end{aligned}$$

Daher ist $y = 0$. Nun gilt also $x \in \operatorname{rg}(\lambda_{k+1} - A)^{r_{k+1}}$ und $(\lambda_{k+1} - A)^{r_{k+1}} x = y = 0$, und daraus folgt nach der Induktionsverankerung $x = 0$. □

Satz 7.4.10. *Sei $\mathbb{K}$ ein Körper, $A \in \mathbb{K}^{n\times n}$ und f der zugehörige Endomorphismus. Seien $\lambda_1, \cdots, \lambda_k$ paarweise verschiedene Eigenwerte von A.*

(1) *Dann ist $\mathbb{K}^n$ die direkte Summe*

$$\mathbb{K}^n = \ker(\lambda_1 - A)^n + \cdots + \ker(\lambda_k - A)^n + \bigcap_{i=1}^{k} \operatorname{rg}(\lambda_i - A)^n\,.$$

(2) *Die Einschränkung von f als Endomorphismus*

$$f : \bigcap_{i=1}^{k} \operatorname{rg}(\lambda_i - A)^n \to \bigcap_{i=1}^{k} \operatorname{rg}(\lambda_i - A)^n$$

besitzt $\lambda_1, \cdots, \lambda_k$ nicht als Eigenwerte.

Beweis. Wir zeigen zunächst durch vollständige Induktion nach k:

$$\mathbb{K}^n = \ker(\lambda_1 - A)^n + \cdots + \ker(\lambda_k - A)^n + \bigcap_{i=1}^{k} \operatorname{rg}(\lambda_i - A)^n\,.$$

Für $k = 1$ ist das Satz 7.4.6. Sei der Satz richtig bis k. Wir betrachten zunächst den invarianten Unterraum

$$W = \bigcap_{i=1}^{k} \operatorname{rg}(\lambda_i - A)^n$$

und die Einschränkung

$$f\mid_W : \begin{cases} W & \to W, \\ x & \mapsto Ax. \end{cases}$$

Wegen Satz 7.4.6 ist W die direkte Summe:

$$\begin{aligned} W &= \ker(\lambda_{k+1} - f\mid_W)^n + \operatorname{rg}(\lambda_{k+1} - f\mid_W)^n \\ &\subset \ker(\lambda_{k+1} - A)^n + [W \cap \operatorname{rg}(\lambda_{k+1} - A)^n] \\ &= \ker(\lambda_{k+1} - A)^n + \bigcap_{i=1}^{k+1} \operatorname{rg}(\lambda_i - A)^n\,. \end{aligned}$$

Daher ist

$$\mathbb{K}^n = \ker(\lambda_1 - A)^n + \cdots + \ker(\lambda_k - A)^n + \ker(\lambda_{k+1} - A)^n + \bigcap_{i=1}^{k+1} \operatorname{rg}(\lambda_i - A)^n \,.$$

Damit ist die Induktion fertig. Um zu zeigen, dass die Summe direkt ist, nehmen wir an:

$$\sum_{i=1}^{k} x_i + y = 0 \quad \text{mit } x_i \in \ker(\lambda_i - A)^n,\ y \in \bigcap_{i=1}^{k} \operatorname{rg}(\lambda_i - A)^n \,.$$

Multiplikation mit $\Pi_{i=1}^k(\lambda_i - A)^n$ ergibt

$$\sum_{i=1}^{k} 0 + \Pi_{i=1}^k(\lambda_i - A)^n y = 0 \,.$$

Nach Lemma 7.4.9 ist $y = 0$. Damit ist $\sum_{i=1}^k x_i = 0$. Die Summe der Haupträume ist aber wegen Satz 7.3.2 eine direkte Summe, also sind auch alle $x_i = 0$.

Nochmals nach Satz 7.4.6 ist für alle j der Durchschnitt

$$\ker(\lambda_j - A)^n \cap \bigcap_{i=1}^{k} \operatorname{rg}(\lambda_i - A)^n \subset \ker(\lambda_j - A)^n \cap \operatorname{rg}(\lambda_j - A)^n = \{0\} \,.$$

Jeder Eigenvektor zu einem der Eigenwerte $\lambda_1, \cdots, \lambda_k$ ist in $\ker(\lambda_j - A)$ und kann daher nicth in $\bigcap_{i=1}^k \operatorname{rg}(\lambda_i - A)^n$ liegen. □

Satz 7.4.11 (Hauptraumzerlegung)**.** *Sei $\mathbb{K}$ ein algebraisch abgeschlossener Körper und $A \in \mathbb{K}^{n\times n}$. Seien $\lambda_1, \cdots, \lambda_k$ die Eigenwerte von A. Dann ist $\mathbb{K}^n$ die direkte Summe*

$$\mathbb{K}^n = \ker(\lambda_1 - A)^n + \cdots + \ker(\lambda_k - A)^n \,.$$

Beweis. Jede Matrix hat ihr charakteristisches Polynom, und über einem algebraisch abgeschlossenen Körper besitzt dieses eine Nullstelle. Es kann also in einem endlichdimensionalen Vektorraum über $\mathbb{K}$ (außer dem Nullraum) keinen Endomorphismus ohne Eigenwert geben. In der Zerlegung aus Satz 7.4.10 muss dann aber $\bigcap_{i=1}^k \operatorname{rg}(\lambda_i - A)^n$ der Nullraum sein. □

7.4.4. *Jordansche Normalform.*
Auf den einzelnen Haupträumen kann man die Struktur von A noch weiter untersuchen und Basen suchen, bezüglich derer sich A besonders einfach darstellen lässt. Wir werden nun eine spezielle Basis des Hauptraumes konstruieren, die die Matrix auf eine besonders einfache Form reduziert, die Jordansche Normalform.

Lemma 7.4.12. *Sei $\mathbb{K}$ ein Körper und $F \in \mathbb{K}^{n\times n}$ eine singuläre Matrix. Sei $r \in \mathbb{N}$ so dass $\ker(F^{r-1}) \neq \ker(F^r) = \ker(F^{r+1})$. Dann gibt es für $t = 1, \cdots, r$ Zahlen $k_t \in \mathbb{N} \cup \{0\}$ und, falls $k_t > 0$, Vektoren $w_{t,1}, \cdots, w_{t,k_t}$, sodass das System*

$$\{F^i w_{t,j} \mid t = 1 \cdots r,\ j = 1 \cdots k_t,\ i = 0 \cdots t-1\}$$

eine Basis von $\ker(F^r)$ ist.

Beweis. Wir beginnen mit $U(0) := \emptyset$ und einer Basis $U(1) = \{u_{1,1}, \cdots, u_{1,m_1}\}$ von $\ker(F)$ und erweitern diese schrittweise zu Basen $U(s)$ von $\ker(F^s)$ für $s = 2, \cdots, r$:

$$U(s) := \{u_{t,j} \mid t = 1 \cdots s,\ j = 1 \cdots m_t\} \,.$$

Nun konstruieren wir, absteigend von $s = r$ bis $s = 1$, Vektoren $w_{s,1}, \cdots, w_{s,k_s} \in \ker(F^s)$ und dazu die Mengen

$$W(s) := \{F^{t-s} w_{t,j} \mid t = s \cdots r,\ j = 1 \cdots k_t\} \,,$$

sodass $W(s)+U(s-1)$ eine Basis von $\ker(F^s)$ ist. (Dabei dürfen auch einzelne k_s null sein, das heißt, nicht bei jedem Abstiegsschrit von $s+1$ auf s müssen neue Vektoren $w_{s,j}$ dazukommen.)

Für $s=r$ wählen wir $\{w_{r,1},\cdots,w_{r,k_r}\}:=\{u_{r,1},\cdots,u_{r,k_r}\}$.

Nehmen wir nun an, dass wir mit der Konstruktion bis $\{w_{s+1,1},\cdots,w_{s+1,k_{s+1}}\}$ vorangekommen sind. Wir zeigen zunächst, dass das System

$$\{F^{t-s}w_{t,j}\mid t=s+1\cdots r,\ j=1\cdots k_t\}\cup U(s-1)$$

ein linear unabhängiges System in $\ker(F^s)$ ist. Auf Grund der Konstruktion ist für $t>s$ der Vektor $w_{t,j}\in\ker(F^t)$, und daher ist $F^{t-s}w_{t,j}\in\ker(F^s)$. Für $t<s$ ist $u_{t,j}\in\ker(F^t)\subset\ker(F^{s-1})\subset\ker(F^s)$. Wir zeigen die lineare Unabhängigkeit. Es sei

$$\sum_{t=s+1}^{r}\sum_{j=1}^{k_t}\lambda_{t,j}F^{t-s}w_{t,j}+\sum_{t=1}^{s-1}\sum_{j=1}^{m_t}\mu_{t,j}u_{t,j}=0\,.$$

Dann ist also

$$F(\sum_{t=s+1}^{r}\sum_{j=1}^{k_t}\lambda_{t,j}F^{t-s-1}w_{t,j})\in\ker(F^{s-1})$$

und daher

$$\sum_{t=s+1}^{r}\sum_{j=1}^{k_t}\lambda_{t,j}F^{t-s-1}w_{t,j}\in\ker(F^s)\,.$$

Weil $U(s)$ eine Basis von $\ker(F^s)$ ist, gibt es also Koeffizienten $\nu_{t,j}$ sodass

$$\sum_{t=s+1}^{r}\sum_{j=1}^{k_t}\lambda_{t,j}F^{t-s-1}w_{t,j}=\sum_{t=1}^{s}\sum_{j=1}^{m_t}\nu_{t,j}u_{t,j}\,.$$

Nun ist aber

$$\{F^{t-s-1}w_{t,j}\mid t=s+1\cdots r,\ j=1\cdots k_t\}\cup U(s)$$

eine Basis von $\ker(F^{s+1})$ und daher linear unabhängig. Darum müssen alle $\lambda_{t,j}$ und alle $\nu_{t,j}$ gleich null sein. Es bleibt also

$$0+\sum_{t=1}^{s-1}\sum_{j=1}^{m_t}\mu_{t,j}u_{t,j}=0\,.$$

Aber $U(s-1)$ ist eine Basis von $\ker(F^{s-1})$ und daher linear unabhängig. Also sind auch alle $\mu_{t,j}$ gleich Null. Damit ist die lineare Unabhängigkeit von

$$\{F^{t-s}w_{t,j}\mid t=s+1\cdots r,\ j=1\cdots k_t\}\cup U(s-1)$$

bewiesen.

Falls dieses System noch nicht den ganzen Kern von F^s aufspannt, können wir das System durch geeignete Vektoren $w_{s,1},\cdots,w_{s,k_s}$ zu einer Basis von $\ker(F^s)$ ergänzen.

Wir zeigen nun durch Induktion nach $t=1,\cdots,r$, dass

$$V(s):=\{F^iw_{t,j}\mid t=0\cdots r,\ j=1\cdots k_t,\ i=0\cdots t-1,\ t-i\le s\}$$

eine Basis von $\ker(F^s)$ ist. Für $s=1$ haben wir bereits die Basis

$$W(1)+U(0)=W(1)=\{F^{t-1}w_{t,j}\mid t=1\cdots r,\ j=1\cdots k_t\}=V(1)\,.$$

Sei nun $V(s)$ eine Basis von $\ker(F^s)$, andererseits ist $U(s)$ ebenfalls eine Basis von $\ker(F^s)$. Wir können also in der folgenden Rechnung diese beiden Basen gegeneinander austauschen. Nach Konstruktion von $W(s+1)$ ist

$$W(s+1)\cup U(s)=\{F^{t-s-1}w_{t,j}\mid t=s+1\cdots r,\ j=1\cdots k_t\}\cup U(s)$$

eine Basis von $\ker(F^{s+1})$, daher ist auch

$$V(s+1) = \{F^{t-s-1}w_{t,j} \mid t = s+1 \cdots r,\ j = 1 \cdots k_t\} \cup V(s)$$

eine Basis von $\ker(F^{s+1})$.

Für $s = r$ erhalten wir also, dass

$$V(r) = \{F^i w_{t,j} \mid t = 1 \cdots r,\ i = 1 \cdots t-1,\ j = 1 \cdots k_t\}$$

eine Basis von $\ker(F^r)$ ist. □

Definition 7.4.13. Die $r \times r$-Matrix

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

heißt Jordanblock der Länge r zum Eigenwert λ.
Ein Jordanblock der Länge 1 ist einfach eine (1×1)-Matrix (λ).
Eine Blockdiagonalmatrix, deren Blöcke Jordanblöcke sind, heißt Matrix in Jordanscher Normalform.

Satz 7.4.14 (Jordansche Normalform)**.** *Sei $\mathbb{K}$ ein algebraisch abgeschlossener Körper und $A \in \mathbb{K}^{n\times n}$. Dann ist A ähnlich zu einer Matrix in Jordanscher Normalform:*

$$T^{-1}AT = \begin{pmatrix} J_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_q \end{pmatrix}.$$

Dabei sind die J_m Jordanblöcke zu (nicht unbedingt paarweise verschiedenen) Eigenwerten λ_i von A.

Beweis. Seien $\lambda_1, \cdots, \lambda_k$ die Eigenwerte der Matrix A. Wir wissen aus Satz 7.4.11 dass $\mathbb{K}^n$ die Summe der Haupträume ist:

$$\mathbb{K}^n = \ker((\lambda_1 - A)^{r_1}) + \cdots + \ker((\lambda_k - A)^{r_k}).$$

Für jeden Eigenwert λ_m setzen wir $F_m = (\lambda_m - A)$ und erhalten nach Lemma 7.4.12 Zahlen $k_1^m, \cdots, k_{r_m}^m$ und ein Vektorsystem

$$\{w_{t,j}^m \mid t = 1 \cdots r_m,\ j = 1 \cdots k_t^m\},$$

sodass

$$\{(\lambda_m - A)^i w_{t,j}^m \mid t = 1 \cdots r_m,\ j = 1 \cdots k_t^m,\ i = 0 \cdots t-1\}$$

eine Basis des Hauptraumes $\ker((\lambda_m - A)^{r_m})$ ist. Wir fixieren nun ein $w_{t,j}^m$ und betrachten das Vektorsystem $(v_1, \cdots, v_t)$ mit $v_i := (A - \lambda_m)^{t-i} w_{t,j}^m$. Für $i = 2, \cdots, t$ gilt dann $v_{i-1} = (A - \lambda_m)v_i$ oder $Av_i = \lambda_m v_i + v_{i-1}$, für $i = 1$ gilt $Av_1 = \lambda_m v_1$. Insbesondere ist die lineare Hülle $W_{t,j}^m := \mathcal{L}(\{v_1, \cdots, v_t\})$ ein invarianter Unterraum. Bezüglich der Basis $B_{t,j}^m = (v_1, \cdots, v_t)$ stellt sich der Homomorphismus

$$f : \begin{cases} W_{t,j}^m & \to W_{t,j}^m \\ x & \mapsto Ax \end{cases}$$

durch die Matrix

$$M^{B^m_{t,j}}_{B^m_{t,j}}(f) = \begin{pmatrix} \lambda_m & 1 & 0 & \cdots & 0 \\ 0 & \lambda_m & 1 & \cdots & 0 \\ 0 & 0 & \lambda_m & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_m \end{pmatrix}$$

dar. Setzen wir nun die Basen aller Räume $W^m_{t,j}$ zu einer Basis B von $\mathbb{K}^n$ zusammen, so ergibt sich bezüglich dieser Basis eine Jordansche Normalform. □

Satz 7.4.15. *Sei $A \in \mathbb{K}^{n\times n}$ ähnlich zu einer Matrix in Jordanscher Normalform*

$$B = T^{-1}AT = \begin{pmatrix} J_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_q \end{pmatrix}.$$

Dabei seien $J_1, \cdots, J_q$ Jordanblöcke der Längen $t_1, \cdots, t_q$ zu den (nicht unbedingt paarweise verschiedenen) Eigenwerten $\lambda_1, \cdots, \lambda_q$. Dann gilt:

1. *Die algebraische Vielfachheit eines Eigenwertes λ von A ist die Summe der Längen der Jordanblöcke zum Eigenwert $\lambda_m = \lambda$.*
2. *Die geometrische Vielfachheit eines Eigenwertes λ von A ist die Anzahl der Jordanblöcke zum Eigenwert $\lambda_m = \lambda$.*
3. *Die kleinste Potenz r, für die $\ker(\lambda - A)^r$ der Hauptraum zum Eigenwert A ist, ist zugleich die größte Länge der Jordanblöcke zum Eigenwert $\lambda_m = \lambda$.*
4. *Die Anzahl der Jordanblöcke mit Länge größer als i zu einem Eigenwert $\lambda_m = \lambda$ ist*
$$\rho((\lambda - A)^i) - \rho((\lambda - A)^{i+1}).$$
5. *Die Jordanblöcke J_m sind, bis auf ihre Reihenfolge auf der Diagonalen, eindeutig bestimmt.*

Beweis. Es ist

$$\det(\lambda - B) = \det(T^{-1}(\lambda - A)T) = \det(T^{-1})\det(\lambda - A)\det(T) = \det(\lambda - A),$$

daher haben A und B dasselbe charakteristische Polynom und dieselben Eigenwerte, mit denselben algebraischen Vielfachheiten. Es ist

$$(\lambda - B)^i = (\lambda - T^{-1}AT)^i = [T^{-1}(\lambda - A)T]^i = T^{-1}(\lambda - A)^i T,$$

also sind $(\lambda - A)^i$ und $(\lambda - B)^i$ ähnliche Matrizen und haben daher dieselben Ränge. Wir können also o.B.d.A. annehmen, dass $A = B$ selbst in Jordanscher Normalform vorliegt.
Dann ist $\lambda - A$ eine obere Dreiecksmatrix, und ihre Determinante ist

$$\det(\lambda - A) = \Pi_{m=1}^q (\lambda - \lambda_m)^{t_m},$$

damit ist die algebraische Vielfachheit eines Eigenwertes λ gegeben durch die Summe der Längen der Jordanblöcke zum Eigenwert λ.
Für jeden Jordanblock J_m ist der Rang

$$\rho((\lambda_m - J_m)^i) = \begin{cases} t_m - i & \text{wenn } i = 0 \cdots t_m - 1 \\ 0 & \text{wenn } i \geq t_m. \end{cases}$$

Damit ist

$$\rho((\lambda_m - J_m)^i) - \rho((\lambda_m - J_m)^{i+1}) = \begin{cases} 1 & \text{wenn } i = 0 \cdots t_m - 1 \\ 0 & \text{wenn } i \geq t_m. \end{cases}$$

Daher ist $\rho((\lambda_m - A)^i) - \rho((\lambda_m - A)^{i+1})$ die Anzahl aller Jordanblöcke zum Eigenwert λ_m mit Länge größer als i. Insbesondere, für $i = 0$, erhalten wir die geometrische

Vielfachheit von λ_m als Anzahl aller Jordanblöcke zum Eigenwert λ_m.
Aus diesen Regeln läßt sich eindeutig die Anzahl der Jordanblöcke jeder bestimmten Länge zu den einzelnen Eigenwerten mit Hilfe der Ränge von $(\lambda_m - A)^i$ feststellen. □

Beispiel 7.4.16. Wir erklären in diesem Beispiel, wie man die Jordansche Normalform einer Matrix bestimmen kann.

Vorgangsweise: Wir nehmen an, wir haben eine Matrix $A \in \mathbb{C}^{10\times 10}$ gegeben. Bekanntlich ist $\mathbb{C}$ algebraisch abgeschlossen, sodass der Satz von der Jordannormalform gilt.

1) Wir bestimmen das charakteristische Polynom und die Eigenwerte von A. Nehmen wir an, wir erhalten $\lambda_1 = 3$ mit algebraischer Vielfachheit 2, und $\lambda_2 = 5$ mit algebraischer Vielfachheit 8. (Probe: Die algebraischen Vielfachheiten müssen sich auf 10 summieren.) Wir wissen nun, dass der Hauptraum zu $\lambda = 3$ die Dimension 2 besitzt, und der Hauptraum zu $\lambda = 5$ die Dimension 8 besitzt.
2) Wir verfolgen nun den Eigenwert $\lambda = 3$: Bisher wissen wir, dass entweder zwei Jordanblöcke der Länge 1, oder ein Jordanblock der Länge 2 vorliegen, da sich die Längen der Jordanblöcke auf die Dimension des Hauptraumes summieren. Wir berechnen nun $\ker(3 - A)$. Nehmen wir an, wir erhalten einen eindimensionalen Kern. Wir wissen nun: Die geometrische Vielfachheit von $\lambda = 3$ beträgt 1. Es gibt einen Jordanblock, und daher muss das ein Block der Länge 2 sein. Der Hauptraum zum Eigenwert 3 ist $\ker(3 - A)^2$, weil 2 die Länge des größten (einzigen) Jordanblocks zu $\lambda = 3$ ist.
3) Wir verfolgen nun den Eigenwert $\lambda = 5$. Bisher wissen wir, dass sich die Längen der Jordanblöcke auf 8 summieren. Wir berechnen $\ker(5 - A)$. Nehmen wir an, wir erhalten einen vierdimensionalen Kern. Dann ist die geometrische Vielfachheit von $\lambda = 5$ also 4, und es gibt 4 Jordanblöcke. Es gibt aber immer noch viele Möglichkeiten, die Gesamtlänge 8 auf vier Blöcke aufzuteilen.
4) Wir berechnen also $\ker(5 - A)^2$. Nehmen wir an, dieser Kern hat die Dimension 7. Die Dimension ist im Vergleich zu $\ker(5 - A)$ um 3 gestiegen. Folglich gibt es drei Blöcke der Länge mindestens 2, der vierte Block hat die Länge 1. Es ist also die Gesamtlänge 7 auf drei Blöcke der Länge ≥ 2 aufzuteilen, und das geht nur in der Form $7 = 2 + 2 + 3$. Wir haben also zwei Blöcke der Länge 2, und einen Block der Länge 3. Der Hauptraum zum Eigenwert $\lambda = 5$ ist $\ker(5 - A)^3$, weil 3 die Länge des größten Jordanblocks ist.
5) Die Jordannormalform ist also

$$\begin{pmatrix} 3 & 1 & & & & & & & & \\ 0 & 3 & & & & & & & & \\ & & 5 & & & & & & & \\ & & & 5 & 1 & & & & & \\ & & & 0 & 5 & & & & & \\ & & & & & 5 & 1 & & & \\ & & & & & 0 & 5 & & & \\ & & & & & & & 5 & 1 & 0 \\ & & & & & & & 0 & 5 & 1 \\ & & & & & & & 0 & 0 & 5 \end{pmatrix}.$$

□

Beispiel 7.4.17. Bestimmen Sie eine Basis, die die folgende Matrix A in Jordansche Normalform überführt.

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

Lösung: Wegen der vielen Nullen ist es am bequemsten, das charakteristische Polynom mit Hilfe des Laplaceschen Entwicklungssatzes zu berechnen:

$$\begin{vmatrix} \lambda-2 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & \lambda-2 & -3 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & \lambda-2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & \lambda-2 & -1 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda-2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & \lambda-2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda-2 & 0 \\ 0 & 4 & 0 & 0 & 0 & 4 & 0 & \lambda-2 \end{vmatrix}$$

Wir sehen zunächst, dass in der ersten, vierten und letzten Spalte je nur ein Element ungleich 0 vorkommt. Wir entwickeln zunächst nach diesen Spalten und erhalten

$$(\lambda-2)^3 \begin{vmatrix} \lambda-2 & -3 & 2 & 0 & 1 \\ 0 & \lambda-2 & 0 & 0 & 0 \\ 0 & 0 & \lambda-2 & 0 & 0 \\ 0 & 2 & 0 & \lambda-2 & -1 \\ 0 & 0 & 0 & 0 & \lambda-2 \end{vmatrix}$$

Nun können wir nach der zweiten, dritten und der letzten Zeile weiterentwickeln:

$$(\lambda-2)^6 \begin{vmatrix} \lambda-2 & 0 \\ 0 & \lambda-2 \end{vmatrix} = (\lambda-2)^8$$

Die Matrix besitzt nur den Eigenwert $\lambda = 2$, mit algebraischer Vielfachheit 8. Wir berechnen nun den Kern von $2 - A$ durch Pivotschritte, wobei wir Nullzeilen sofort weglassen:

$$\begin{pmatrix} 0 & -1* & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -3 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & -1 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 4 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -3 & 0 & 2 & 0 & 1* & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & -1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & -1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & -3 & 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1* & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

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

$$N_1 = \begin{pmatrix} 0 & -1* & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1* & 0 \\ 0 & 0 & 0 & 0 & -1* & 0 & 0 & 0 \\ 0 & 0 & -1* & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nichtbasisch bleiben die erste, vierte, sechste und achte Variable, wir erhalten als Basis des Kerns $\ker(2-A)$

$$w_{1,1} = \begin{pmatrix}1\\0\\0\\0\\0\\0\\0\\0\end{pmatrix},\ w_{1,2} = \begin{pmatrix}0\\0\\0\\1\\0\\0\\0\\0\end{pmatrix},\ w_{1,3} = \begin{pmatrix}0\\-1\\0\\0\\0\\1\\0\\0\end{pmatrix},\ w_{1,4} = \begin{pmatrix}0\\0\\0\\0\\0\\0\\0\\1\end{pmatrix}.$$

Beachten Sie, dass die Matrix N_1 denselben Kern hat wie $(2-A)$, und daher ist $\ker((2-A)^2) = \ker(N_1(2-A))$. Die zweite Matrix ist aber bequemer zu berechnen. Wir führen weitere Pivotschritte durch, wobei wir die Pivotelemente nur aus den bisher basischen Spalten wählen.

$$\begin{pmatrix}0&-1&0&0&0&-1&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&-1&0&0&0\\0&0&-1&0&0&0&0&0\end{pmatrix}\cdot\begin{pmatrix}0&-1&0&0&0&-1&0&0\\0&0&-3&0&2&0&1&0\\0&0&0&0&0&0&0&0\\0&-3&0&0&-1&-3&0&0\\0&0&0&0&0&0&0&0\\0&0&2&0&0&0&-1&0\\0&0&0&0&0&0&0&0\\0&4&0&0&0&4&0&0\end{pmatrix}$$

$$= \begin{pmatrix}0&0&1*&0&-2&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\end{pmatrix}$$

Hier sind keine weiteren Pivotschritte notwendig, wir erhalten nach Weglassen der Nullzeilen

$$N_2 = \begin{pmatrix}0&0&1*&0&-2&0&0&0\end{pmatrix}$$

Neue nichbasische Variable sind die zweite, fünfte und siebente Variable. Wir ergänzen die bisherige Basis von $\ker(2-A)$ durch die dazu gehörigen Vektoren zu einer Basis von $\ker((2-A)^2)$. Neu dazu gekommen sind also

$$w_{2,1} = \begin{pmatrix}0\\1\\0\\0\\0\\0\\0\\0\end{pmatrix},\ w_{2,2} = \begin{pmatrix}0\\0\\2\\0\\1\\0\\0\\0\end{pmatrix},\ w_{2,3} = \begin{pmatrix}0\\0\\0\\0\\0\\0\\1\\0\end{pmatrix}.$$

Für den letzten Schritt multiplizieren wir wieder $N_2(2-A)$:

$$\begin{pmatrix}0&0&1&0&-2&0&0&0\end{pmatrix}\cdot\begin{pmatrix}0&-1&0&0&0&-1&0&0\\0&0&-3&0&2&0&1&0\\0&0&0&0&0&0&0&0\\0&-3&0&0&-1&-3&0&0\\0&0&0&0&0&0&0&0\\0&0&2&0&0&0&-1&0\\0&0&0&0&0&0&0&0\\0&4&0&0&0&4&0&0\end{pmatrix}$$

$$= \begin{pmatrix}0&0&0&0&0&0&0&0\end{pmatrix}$$

Als letzte Variable wird die dritte nichtbasisch, und wir setzen

$$w_{3,1} = \begin{pmatrix}0\\0\\1\\0\\0\\0\\0\\0\end{pmatrix}.$$

Dieser Vektor $w_{3,1}$ wird, zusammen mit $(2-A)w_{3,1}$ und $(2-A)^2w_{3,1}$ unseren längsten Jordanblock der Länge 3 erzeugen. Wir berechnen jetzt $(2-A)w_{3,1}$:

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

Wir tauschen diesen Vektor gegen einen der Vektoren $w_{2,i}$, sodass wir nach wie vor eine Basis des $\mathbb{K}^n$ behalten. Wenn man nicht mit freiem Auge sieht, welchen Vektor man nehmen kann, kann man ansetzen: $(2-A)w_{31} = \sum_{t=1}^{2}\sum_{j=1}^{k_t} \mu t, j w_{t,j}$. Mindestens ein $\mu_{2,j}$ muss ungleich null sein, und nach dem Austauschlemma von Steinitz kann $(2-A)w_{31}$ gegen w_{2j} getauscht werden. Das ergibt folgende Rechnung:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1* & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

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

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

also $(2-A)w_{3,1} = 2w_{1,3} - w_{2,1}$. Daher lassen wir w_{21} weg, es bleiben

$$w_{2,2} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \; w_{2,3} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \; (2-A)w_{3,1} = \begin{pmatrix} 0 \\ -3 \\ 0 \\ 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}.$$

Mit den Vektoren $w_{2,2}$ und $w_{2,3}$ starten unsere beiden Jordanblöcke der Länge 2. Die drei obigen Vektoren multiplizieren wir wieder mit $(2-A)$, der Kürze halber schreiben wir die drei Multiplikationen in eine Matrix:

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

Wir brauchen noch einen Vektor aus dem System $w_{1,1}, \cdots, w_{1,4}$, um diese drei Vektoren zu einer Basis von $\ker(2-A)$ zu ergänzen:

$$w_{1,1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ (2-A)w_{2,2} = \begin{pmatrix} 0 \\ -4 \\ 0 \\ -1 \\ 0 \\ 4 \\ 0 \\ 0 \end{pmatrix}, \ (2-A)w_{2,3} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \ (2-A)^2 w_{3,1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 3 \\ 0 \\ 0 \\ 0 \\ -4 \end{pmatrix}.$$

Nun setzen wir die erhaltenen Basisvektoren in der folgenden Reihenfolge zur Matrix T zusammen:

$$\big((2-A)^2 w_{3,1}, -(2-A)w_{3,1}, w_{3,1}, -(2-A)w_{2,2}, w_{2,2}, -(2-A)w_{2,3}, w_{2,3}, w_{1,1}\big)$$

und erhalten

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

Nachrechnen (mit MATLAB) ergibt die Jordansche Normalform

$$T^{-1}AT = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

□

7.5. **Matrizen in Polynome eingesetzt.**

7.5.1. *Minimalpolynom.*

Definition 7.5.1. Sei $\mathbb{K}$ ein Körper, $p(x) = \gamma_0 + \gamma_1 x + \gamma_2 x^2 + \cdots + \gamma_r x^r$ ein Polynom in $\mathbb{K}[x]$. Sei $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. Wir definieren die Matrix

$$p(A) := \gamma_0 E_n + \gamma_1 A + \gamma_2 A^2 + \cdots + \gamma_r A^r .$$

Weil der Raum $\mathbb{K}^{n\times n}$ die Dimension n^2 hat, muss das System von Matrizen $(E_n, A, \cdots, A^{n^2})$ linear abhängig sein, daher muss es ein Polynom vom Grad höchstens n^2 geben, sodass $p(A) = 0$. Der folgende Satz zeigt, dass in Wirklichkeit ein viel kleinerer Grad ausreicht.

Lemma 7.5.2. *Sei $\mathbb{K}$ ein Körper, $p \in \mathbb{K}[x]$ ein Polynom, $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix, und $T \in \mathbb{K}^{n\times n}$ eine reguläre Matrix. Dann gilt: $p(T^{-1}AT) = T^{-1}p(A)T$.*

Beweis. Zunächst berechnen wir

$$\begin{aligned} & [T^{-1}AT]^n = [T^{-1}AT][T^{-1}AT]\cdots[T^{-1}AT] \\ = \ & T^{-1}A[TT^{-1}]A[TT^{-1}]\cdots[TT^{-1}]AT = T^{-1}A^nT . \end{aligned}$$

Damit ist aber

$$p(T^{-1}AT) = \sum_{i=0}^{r} \gamma_i [T^{-1}AT]^i = \sum_{i=0}^{r} \gamma_i T^{-1}A^iT = T^{-1}\left[\sum_{i=0}^{r} \gamma_i A^i\right] T .$$

□

Satz 7.5.3. *Sei $\mathbb{K}$ ein algebraisch abgeschlossener Körper, sei $A \in \mathbb{K}^{n\times n}$. Es seien $\lambda_1, \cdots, \lambda_k$ die Eigenwerte von A, und es seien $r_1, \cdots, r_k$ die kleinsten Potenzen, sodass $\ker(\lambda_i - A)^{r_i} = \ker(\lambda_i - A)^{r_i+1}$. (Es ist also $\ker(\lambda_i - A)^{r_i}$ der Hauptraum zum Eigenwert λ_i.) Wir setzen*

$$p(x) = (x - \lambda_1)^{r_1}(x - \lambda_2)^{r_2} \cdots (x - \lambda_k)^{r_k} .$$

Dann gilt:

1) $p(A) = 0$.
2) *Ist $q \in \mathbb{K}[x]$ ein Polynom, so dass $q(A) = 0$, so ist p ein Teiler von q.*

Beweis. Nach Satz 7.4.11 gibt es eine reguläre Matrix T, sodass $C = T^{-1}AT$ Blockdiagonalform hat:

$$C = \begin{pmatrix} C_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & C_k \end{pmatrix} .$$

Dabei sind die $C_i \in \mathbb{K}^{m_i\times m_i}$, und die m_i sind die algebraischen Vielfachheiten von λ_i. Die C_i repräsentieren die Einschränkung des Homomorphismus $x \mapsto Ax$ auf die Haupträume $\ker(\lambda_i - A)^{r_i}$. Wegen Lemma 7.5.2 ist $p(A) = 0$ genau dann, wenn $p(C) = 0$. Direktes Nachrechnen ergibt

$$p(C) = \begin{pmatrix} p(C_1) & 0 & \cdots & 0 \\ 0 & p(C_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p(C_k) \end{pmatrix} .$$

Auf Grund der Definition des Hauptraumes und der r_i ist r_i die kleinste Potenz sodass $(\lambda_i - C_i)^{r_i} = 0$. Für $\mu \neq \lambda_i$ gilt $\ker(\mu - A) \cap \ker(\lambda_i - A)^{r_i} = \{0\}$, anders gesagt, der Hauptraum zu λ_i enthält von $\ker(\mu - A)$ nur den Nullvektor. Also gehört die Matrix $(\mu - C_i)$ zu einem injektiven Endomorphismus, deshalb ist sie regulär.

Sei nun $q \in \mathbb{K}[x] \setminus \{0\}$. Weil $\mathbb{K}$ algebraisch abgeschlossen ist, zerfällt q in Linearfaktoren:

$$q(x) = \gamma \, \Pi_{u=1}^{k}(\lambda_u - x)^{s_u} \, \Pi_{j=1}^{l}(\mu_j - x)^{t_j} .$$

Dabei sind $s_u \in \mathbb{N} \cup \{0\}$, $t_j \in \mathbb{N}$, $\mu_j \notin \{\lambda_1, \cdots, \lambda_k\}$, $\gamma \in \mathbb{K} \setminus \{0\}$. Es ist $q(A) = 0$ genau dann, wenn $q(C) = 0$, und das gilt genau dann, wenn für alle i gilt: $q(C_i) = 0$. Nun ist

$$q(C_i) = \gamma \, \Pi_{u\neq i}(\lambda_u - C_i)^{s_u} \, \Pi_{j=1}^{l}(\mu_j - C_i)^{t_j} \, (\lambda_i - C_i)^{s_i} .$$

Da alle Matrizen $(\lambda_u - C_i)$ für $u \neq i$ und alle Matrizen $(\mu_j - C_i)$ regulär sind, ist

$$q(C_i) = 0 \;\Leftrightarrow\; (\lambda_i - C_i)^{s_i} = 0 \;\Leftrightarrow s_i \geq r_i ,$$

und das gilt genau dann, wenn p ein Teiler von q ist. □

Definition 7.5.4. Sei $\mathbb{K}$ ein Körper, $p \in \mathbb{K}[x]$ ein Polynom, $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix. Sei $p(x) = x^r + \gamma_{r-1}x^{r-1} + \cdots + \gamma_0$ in $\mathbb{K}[x]$ so, dass für jedes $q \in \mathbb{K}[x]$ gilt:

$$q(A) = 0 \;\Leftrightarrow\; p \text{ teilt } q .$$

Dann heißt p das Minimalpolynom von A.

Diese Definition funktioniert auch über nicht algebraisch abgeschlossenen Körpern, nur die Beschreibung des Minimalpolynoms mittes Satz 7.5.3 funktioniert nicht mehr immer, weil sich nicht für jede Matrix der Raum $\mathbb{K}^n$ in Haupträume zerlegen lässt.

7.5.2. *Satz von Cayley-Hamilton.*

Satz 7.5.5 (Satz von Cayley-Hamilton)**.** *Sei $\mathbb{K}$ ein algebraisch abgeschlossener Körper, $A \in \mathbb{K}^{n\times n}$ eine quadratische Matrix, und $q \in \mathbb{K}[x]$ ihr charakteristisches Polynom. Dann ist $q(A) = 0$.*

Beweis. Wir wissen bereits aus Satz 7.5.3, dass das Minimalpolynom von A sich in der folgenden Form schreiben lässt:

$$p(x) = \Pi_{i=1}^{k}(x-\lambda_i)^{r_i}\,.$$

Dabei sind $\lambda_1, \cdots, \lambda_k$ die Eigenwerte von A, und r_i jeweils die kleinste Potenz sodass $\ker(\lambda_i - A)^{r_i} = \ker(\lambda_i - A)^{r_i+1}$. Dagegen ist das charakteristische Polynom von A

$$q(x) = \Pi_{i=1}^{k}(x-\lambda_i)^{m_i}\,,$$

wobei m_i die algebraische Vielfachheit von λ_i ist. Zugleich wissen wir aus Satz 7.4.11, dass m_i die Dimension des Hauptraumes $\ker(\lambda_i - A)^{r_i}$ ist. Nun ist

$$1 \le \dim(\ker(\lambda_i - A)) < \dim(\ker(\lambda_i - A)^2) < \cdots < \dim(\ker(\lambda_i - A)^{r_i}) = m_i\,.$$

Daher ist $r_i \le m_i$, und das Minimalpolynom teilt das charakteristische Polynom. Folglich ist $q(A) = 0$. □

Beispiel 7.5.6. Bestimmen Sie das charakteristische Polynom und das Minimalpolynom der folgenden Matrix A über $\mathbb{C}$. Überprüfen Sie, dass A, in beide Polynome eingesetzt, die Nullmatrix ergibt.

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

Lösung: Das charakteristische Polynom erhalten wir durch Entwicklung nach der zweiten Zeile:

$$\begin{vmatrix} \lambda-3 & -1 & -3 \\ 0 & \lambda-2 & 0 \\ 1 & 1 & \lambda+1 \end{vmatrix} = (\lambda-2)\begin{vmatrix} \lambda-3 & -3 \\ 1 & \lambda+1 \end{vmatrix}$$

$$= \quad (\lambda-2)[(\lambda-3)(\lambda+1)+3] = (\lambda-2)(\lambda^2-2\lambda) = (\lambda-2)^2\lambda = \lambda^3 - 4\lambda^2 + 4\lambda\,.$$

Es gibt zwei Eigenwerte $\lambda = 2$ und $\lambda = 0$. Wir bestimmen die geometrische Vielfachheit von $\lambda = 2$, also die Dimension von $\ker(2 - A)$:

$$\begin{pmatrix} -1 & -1 & -3 \\ 0 & 0 & 0 \\ 1 & 1 & 3 \end{pmatrix} \to \begin{pmatrix} -1 & -1 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,.$$

Daher ist der Eigenraum zu $\lambda = 2$ zweidimensional, und das ist bereits die algebraischen Vielfachheit, also die Dimension des Hauptraumes. Folglich ist bereits der Eigenraum $\ker(2 - A)^1$ der Hauptraum. Damit ist das Minimalpolynom

$$(\lambda-2)^1\lambda = \lambda^2 - 2\lambda\,.$$

Um die Probe zu machen, berechnen wir erst die Potenzen von A:

$$A^2 = \begin{pmatrix} 6 & 2 & 6 \\ 0 & 4 & 0 \\ -2 & -2 & -2 \end{pmatrix}\,, \quad A^3 = \begin{pmatrix} 12 & 4 & 12 \\ 0 & 8 & 0 \\ -4 & -4 & -4 \end{pmatrix}\,.$$

Wir setzen in das charakteristische Polynom ein:

$$A^3 - 4A^2 + 4A = \begin{pmatrix} 12 & 4 & 12 \\ 0 & 8 & 0 \\ -4 & -4 & -4 \end{pmatrix} - \begin{pmatrix} 24 & 8 & 24 \\ 0 & 16 & 0 \\ -8 & -8 & -8 \end{pmatrix} + \begin{pmatrix} 12 & 4 & 12 \\ 0 & 8 & 0 \\ -4 & -4 & -4 \end{pmatrix} = 0\,.$$

Wir setzen in das Minimalpolynom ein:

$$A^2 - 2A = \begin{pmatrix} 6 & 2 & 6 \\ 0 & 4 & 0 \\ -2 & -2 & -2 \end{pmatrix} - \begin{pmatrix} 6 & 2 & 6 \\ 0 & 4 & 0 \\ -2 & -2 & -2 \end{pmatrix} = 0\,.$$

□

8. Normale Matrizen

Wir werden in diesem Kapitel stets den Raum $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ mit dem euklidischen inneren Produkt und der dazugehörigen Norm betrachten.

8.1. **Orthogonale und unitäre Matrizen.**

8.1.1. *Unitäre Matrizen.*

Definition 8.1.1.

1) Eine Matrix $T \in \mathbb{R}^{n\times n}$ heißt orthogonal, wenn sie regulär ist und ihre Transponierte zugleich die Inverse ist: $T^{-1} = T^T$.
2) Eine Matrix $T \in \mathbb{C}^{n\times n}$ heißt unitär, wenn sie regulär ist und ihre Adjungierte zugleich die Inverse ist: $T^{-1} = T^*$.

Eine orthogonale Matrix ist also nichts anderes als eine reelle unitäre Matrix. Wenn wir in den folgenden Sätzen immer wieder orthogonale und unitäre Matrizen betrachten, werden wir normalerweise nur den Beweis für unitäre Matrizen über $\mathbb{C}$ führen. Der reelle, orthogonale Fall geht dann genauso.

Satz 8.1.2.

1) *Sei $T \in \mathbb{R}^{n\times n}$. Dann sind äquivalent:*
   (i) *T ist orthogonal.*
   (ii) *Für alle $x \in \mathbb{R}^n$ gilt $\|Tx\| = \|x\|$.*
   (iii) *Für alle $x, y \in \mathbb{R}^n$ gilt $\langle Tx, Ty\rangle = \langle x, y\rangle$.*
2) *Sei $T \in \mathbb{C}^{n\times n}$. Dann sind äquivalent:*
   (i) *T ist unitär.*
   (ii) *Für alle $x \in \mathbb{C}^n$ gilt $\|Tx\| = \|x\|$.*
   (iii) *Für alle $x, y \in \mathbb{C}^n$ gilt $\Re(\langle Tx, Ty\rangle) = \Re(\langle x, y\rangle)$.*
   (iv) *Für alle $x, y \in \mathbb{C}^n$ gilt $\langle Tx, Ty\rangle = \langle x, y\rangle$.*

Beweis. Wir beweisen Teil (2), dann geht Teil (1) ebenso.
(i) $\Rightarrow$ (ii): Sei $T^* = T^{-1}$. Es ist

$$\|Tx\|^2 = \langle Tx, Tx\rangle = (Tx)^*(Tx) = x^*T^*Tx = x^*x = \|x\|^2 .$$

(ii) $\Rightarrow$ (iii): Es ist

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + \|y\|^2 + \langle x, y\rangle + \langle y, x\rangle = \|x\|^2 + \|y\|^2 + 2\Re(\langle x, y\rangle) , \\ \|x-y\|^2 &= \|x\|^2 + \|y\|^2 - \langle x, y\rangle - \langle y, x\rangle = \|x\|^2 + \|y\|^2 - 2\Re(\langle x, y\rangle) . \end{aligned}$$

Wir subtrahieren diese Gleichungen und erhalten

$$4\Re(\langle x, y\rangle) = \|x+y\|^2 - \|x-y\|^2 .$$

Man kann in (ii) den Vektor x einmal durch $x+y$, und dann durch $x-y$ ersetzen, und erhält

$$4\Re(\langle Tx, Ty\rangle) = \|T(x+y)\|^2 - \|T(x-y)\|^2 = \|x+y\|^2 - \|x-y\|^2 = 4\Re(\langle x, y\rangle) .$$

(iii) $\Rightarrow$ (iv): Es gilt:

$$\Re(\langle x, iy\rangle) = \Re(i\langle x, y\rangle) = -\Im(\langle x, y\rangle) .$$

Aus (iii) folgt daher

$$\Im(\langle Tx, Ty\rangle = -\Re(\langle Tx, T(iy)\rangle) = -\Re(\langle x, iy\rangle) = \Im(\langle x, y\rangle) .$$

Nochmals wegen (iii) folgt also $\langle Tx, Ty\rangle = \langle x, y\rangle$.
(iv) $\Rightarrow$ (i): Es gelte (iv). Seien $x, y \in \mathbb{C}^n$. Dann gilt

$$\langle y, T^*Tx - x\rangle = \langle y, T^*Tx\rangle - \langle y, x\rangle = \langle Ty, Tx\rangle - \langle y, x\rangle = 0 .$$

Setzen wir insbesondere $y = T^*Tx - x$, so folgt $\|T^*Tx - x\|^2 = 0$, also $T^*Tx = x$. Damit ist T regulär, und T^* ist die Inverse von T. □

Definition 8.1.3. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $f : \mathbb{K}^n \to \mathbb{K}^n$ eine (nicht notwendig lineare) Abbildung mit der Eigenschaft:

$$(\forall x, y \in \mathbb{K}^n)\, \|f(x) - f(y)\| = \|x - y\| .$$

Dann heißt f eine isometrische Abbildung oder Isometrie.

Eine *lineare* Abbildung $\mathbb{K}^n \to \mathbb{K}^n$ ist also genau dann eine Isometrie, wenn sie durch eine unitäre (bzw. im reellen Fall orthogonale) Matrix beschrieben wird.

Beispiel 8.1.4. Die folgende Matrix beschreibt eine Drehung im $\mathbb{R}^2$ um den Nullpunkt mit einem Drehwinkel von ϕ:

$$D_\phi = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} .$$

Zeigen Sie, dass diese Matrix orthogonal ist.

Lösung: Aus der geometrischen Anschauung wissen wir, dass eine Drehung eine Isometrie ist, also erwarten wir dass D_ϕ orthogonal ist. Wir können das aber auch nachrechnen:

$$\begin{aligned} D_\phi^T D_\phi &= \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \\ = \quad & \begin{pmatrix} \cos^2(\phi) + \sin^2(\phi) & 0 \\ 0 & \sin^2(\phi) + \cos^2(\phi) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \end{aligned}$$

□

Beispiel 8.1.5. Bewegungen des $\mathbb{R}^2$ oder $\mathbb{R}^3$, welche den Nullpunkt festlassen, werden durch orthogonale Matrizen beschrieben. Also z.B. Drehungen um Achsen durch den Nullpunkt, Spiegelungen an Geraden oder Ebenen durch den Nullpunkt, und Hintereinanderausführung solcher Bewegungen. Dagegen ist die Parallelverschiebung entlang eines Vektors a, also $f(x) = x + a$ zwar eine Isometrie, aber keine lineare Abbildung.

Satz 8.1.6.

1) *Das Produkt orthogonaler (bzw. unitärer) Matrizen ist orthogonal (bzw. unitär).*
2) *Eine reguläre Matrix A ist genau dann orthogonal (unitär), wenn A^{-1} orthogonal (unitär) ist.*

Trauen Sie sich und beweisen Sie es selbst! Entweder direkt mit der Definition von Orthogonalität, oder mit Hilfe von Satz 8.1.2.

8.1.2. *Orthogonale Diagonalisierung.*

Satz 8.1.7. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $T \in \mathbb{K}^{n\times n}$. Seien $(s_1, \cdots, s_n)$ die Spalten von T und $z_1, \cdots, z_n$ die Zeilen von T. Dann sind folgende Aussagen äquivalent:*

1) *T ist orthogonal bzw. unitär.*
2) *$(s_1, \cdots, s_n)$ ist ein Orthonormalsystem.*
3) *$(z_1^*, \cdots, z_n^*)$ ist ein Orthonormalsystem.*

Beweis. (1) ⇒ (2): Seien $s_1, \cdots, s_n$ die Spalten von T, und $e_1, \cdots, e_n$ die Einheitsvektoren in $\mathbb{K}^n$. Bekanntlich ist $s_i = T(e_i)$. Wegen Satz 8.1.2 gilt für orthogonale (unitäre) Matrizen

$$\langle s_i, s_j \rangle = \langle Te_i, Te_j \rangle = \langle e_i, e_j \rangle = \delta_{i,j}\,.$$

(Dabei ist $\delta_{i,j}$ das Kronecker-Delta.)

(2) ⇒ (1): Sei nun $(s_1, \cdots, s_n)$ ein Orthonormalsystem. Die Zeilen von T^* sind die Adjungierten der Spalten von T. Damit ist

$$T^*T = \begin{pmatrix} -- & s_1^* & -- \\ & \vdots & \\ -- & s_n^* & -- \end{pmatrix} \begin{pmatrix} | & & | \\ s_1 & \cdots & s_n \\ | & & | \end{pmatrix} = \begin{pmatrix} s_1^*s_1 & \cdots & s_1^*s_n \\ \vdots & & \vdots \\ s_n^*s_1 & \cdots & s_n^*s_n \end{pmatrix} = E_n\,.$$

Also ist T^* die Inverse zu T.

(1) ⇔ (3): Es ist T unitär genau dann, wenn T^* unitär ist. Dies gilt aber wegen Punkt (2) genau dann, wenn die Spalten von T^* ein Orthonormalsystem sind. Die Spalten von T^* sind aber die Adjungierten der Zeilen von T. □

Korollar 8.1.8. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $B = (b_1, \cdots, b_n)$ eine Basis von $\mathbb{K}^n$ und $E = (e_1, \cdots, e_n)$ die Einheitsbasis. Dann gilt: Die Übergangsmatrix T_E^B (und damit auch ihre Inverse T_B^E) ist genau dann orthogonal bzw. unitär, wenn B eine Orthonormalbasis ist.*

Beweis. Die Spalten von T_E^B sind die Basisvektoren $b_1, \cdots, b_n$, in Einheitskoordinaten geschrieben. Nach Satz 8.1.7 sind die Spalten von T_E^B genau dann ein Orthonormalsystem, wenn T_E^B orthogonal bzw. unitär ist. □

Definition 8.1.9. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n\times n}$. Die Matrix A heißt orthogonal (unitär) diagonalisierbar, wenn $\mathbb{K}^n$ eine Orthonormalbasis aus Eigenvektoren der Matrix A besitzt.

Satz 8.1.10. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n\times n}$. Die Matrix A ist genau dann orthogonal (unitär) diagonalisierbar, wenn es eine orthogonale (unitäre) Matrix T gibt, sodass $T^{-1}AT$ eine Diagonalmatrix ist.*

Beweis. Ist $T \in \mathbb{K}^{n\times n}$ regulär, so ist $T^{-1}AT$ genau dann eine Diagonalmatrix, wenn die Spalten von T Eigenvektoren von A sind. Die Spalten bilden genau dann zusätzlich ein Orthonormalsystem, wenn T orthogonal (unitär) ist. □

Satz 8.1.11. *Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann unitär diagonalisierbar, wenn die folgenden beiden Aussagen gelten:*

1) *Sind λ, μ zwei verschiedene Eigenwerte von A und x, y dazugehörige Eigenvektoren, so ist $\langle x, y\rangle = 0$. (Anders ausgedrückt, die Eigenräume stehen aufeinander orthogonal.)*
2) *Ist λ ein Eigenvektor von A, so ist $\ker(\lambda - A)^2 = \ker(\lambda - A)$. (Anders ausgedrückt, die Vielfachheit des Eigenwertes im Minimalpolynom ist 1, oder wieder anders, der Hauptraum zu λ ist der Eigenraum.)*

Beweis. ⇒: Wenn A diagonalisierbar ist, ist $\mathbb{C}^n$ die direkte Summe der Eigenräume. Daher müssen in diesem Fall die Haupträume gleich den Eigenräumen sein. Damit gilt (2). Sei nun A unitär diagonalisierbar mit paarweise verschiedenen Eigenwerten $\lambda_1, \cdots, \lambda_k$, seien $b_{i,1}, \cdots, b_{i,m_i}$ Eigenvektoren zu λ_i sodass sich mit $(b_{1,1}, \cdots, b_{k,m_k})$ eine Orthonormalbasis von $\mathbb{C}^n$ ergibt. Sei x ein beliebiger Eigenvektor zu λ_i und y ein Eigenvektor zu λ_j mit $i \neq j$. Man sieht leicht, dass $x \in \mathcal{L}(\{b_{i,1}, \cdots, b_{i,m_i}\})$ und $y \in \mathcal{L}(\{b_{j,1}, \cdots, b_{j,m_j}\})$. Weil alle $b_{i,s}$ auf alle $b_{j,t}$ orthogonal stehen, folgt $\langle x, y\rangle = 0$. Also gilt (1).

$\Leftarrow$: Wir nehmen nun an, dass (1) und (2) gelten. Wegen der algebraischen Abgeschlossenheit von $\mathbb{C}$ und Satz 7.4.11 ist $\mathbb{C}^n$ die direkte Summe der Haupträume von A. Wegen (1) sind aber die Haupträume zugleich die Eigenräume. Seien wieder $\lambda_1, \cdots, \lambda_k$ die Eigenwerte von A. Für jedes i wählen wir eine Orthonormalbasis $(b_{i,1}, \cdots, b_{i,m_i})$ des Eigenraumes $\ker(\lambda_i - A)$. Wegen (2) stehen für $i \neq j$ die Vektoren $b_{i,s}$ und $b_{j,t}$ orthogonal. Daraus ergibt sich, dass $(b_{1,1}, \cdots, b_{k,m_k})$ eine Orthonormalbasis des $\mathbb{C}^n$ ist. □

Bemerkung 8.1.12. Beachten Sie aber, dass eine reelle $n \times n$-Matrix, für deren reelle Eigenvektoren und Eigenwerte die Bedingungen von Satz 8.1.11 gelten, nicht diagonalisierbar sein muss. Sie kann ja komplexe Eigenwerte haben.

8.2. **Normale Matrizen.**

8.2.1. *Definition und Beispiele.*

Definition 8.2.1. Sei $A \in \mathbb{C}^{n\times n}$ eine quadratische Matrix.

1) A heißt normal, wenn gilt $AA^* = A^*A$.
2) A heißt selbstadjungiert oder hermitesch, wenn gilt: $A = A^*$. Eine selbstadjungierte Matrix $A \in \mathbb{R}^{n\times n}$ heißt auch symmetrisch.
3) A heißt schiefadjungiert, wenn gilt: $A = -A^*$.

Bemerkung 8.2.2. Unitäre, selbstadjungierte und schiefadjungierte Matrizen sind Spezialfälle von normalen Matrizen.

Beweis den geneigten LeserInnen mit besten Empfehlungen überlassen.

8.2.2. *Orthogonale Diagonalisierbarkeit normaler Matrizen.*

Satz 8.2.3. *Sei $A \in \mathbb{C}^{n\times n}$ eine normale Matrix. Dann gilt:*

1) $\ker(\lambda - A) = \ker(\overline{\lambda} - A^*)$
2) $\ker(\lambda - A) \perp \ker(\mu - A)$ *falls* $\lambda \neq \mu$.
3) $\ker(\lambda - A)^2 = \ker(\lambda - A)$

Beweis. Teil (1): Sei $(\lambda - A)x = 0$. Dann ist

$$\begin{aligned} &\|(\overline{\lambda} - A^*)x\|^2 = \langle(\overline{\lambda} - A^*)x, (\overline{\lambda} - A^*)x\rangle = \langle x, (\lambda - A)(\overline{\lambda} - A^*)x\rangle \\ = \;& \langle x, (\overline{\lambda} - A^*)(\lambda - A)x\rangle = \langle(\lambda - A)x, (\lambda - A)x\rangle = \|(\lambda - A)x\|^2 = 0\,. \end{aligned}$$

Also ist $\ker(\lambda - A) \subset \ker(\overline{\lambda} - A^*)$. Aus Gründen der Symmetrie folgt nun auch $\ker(\overline{\lambda} - A^*) \subset \ker(\lambda - A)$.

Teil (2): Sei $(\lambda - A)x = 0$, $(\mu - A)y = 0$ für $\lambda \neq \mu$. Wegen Teil (1) gilt auch $(\overline{\lambda} - A^*)x = 0$. Daher ist

$$\mu\langle x, y\rangle = \langle x, \mu y\rangle = \langle x, Ay\rangle = \langle A^*x, y\rangle = \langle\overline{\lambda}x, y\rangle = \lambda\langle x, y\rangle\,.$$

Da $\lambda \neq \mu$, folgt $\langle x, y\rangle = 0$.

Teil (3): Sei $(\lambda - A)^2x = 0$. Dann ist also $(\lambda - A)x \in \ker(\lambda - A)$. Wegen Teil (1) ist auch $(\lambda - A)x \in \ker(\overline{\lambda} - A^*)$. Es ist

$$\|(\lambda - A)x\|^2 = \langle(\lambda - A)x, (\lambda - A)x\rangle = \langle x, (\overline{\lambda} - A^*)(\lambda - A)x\rangle = \langle x, 0\rangle = 0\,.$$

□

Satz 8.2.4. *Sei $A \in \mathbb{C}^{n\times n}$. Dann gilt: Die Matrix A ist genau dann normal, wenn A unitär diagonalisierbar ist.*

Beweis. Ist A normal, so folgt aus Satz 8.2.3 und Satz 8.1.11, dass A unitär diagonalisierbar ist.
Sei nun $D = T^{-1}AT$ mit einer unitären Matrix T und einer Diagonalmatrix D. Man überprüft leicht, dass Diagonalmatrizen miteinander kommutieren, sodass $D^*D = DD^*$. Es ist dann $A = TDT^{-1}$ und $T^{-1} = T^*$. Daher ist $A^* = (T^{-1})^*D^*T^* = TD^*T^{-1}$ und daher

$$AA^* = TDT^{-1}TD^*T^{-1} = TDD^*T^{-1} = TD^*DT^{-1} = TD^*T^{-1}TDT^{-1} = A^*A\,.$$

□

8.3. **Selbstadjungierte Matrizen und Definitheit.**

8.3.1. *Diagonalisierung selbstadjungierter Matrizen.*

Satz 8.3.1. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $A \in \mathbb{K}^{n\times n}$ selbstadjungiert.*

1) *Alle Eigenwerte von A sind reell.*
2) *A ist über $\mathbb{K}$ unitär (orthogonal) diagonalisierbar.*

Beweis. Sei $\lambda x = Ax$ für einen Vektor $x \neq 0$. Dann ist wegen Satz 8.2.3 auch $\overline{\lambda}x = A^*x = Ax = \lambda x$. Da $x \neq 0$ folgt $\lambda = \overline{\lambda}$.

Die Matrix A ist jedenfalls diagonalisierbar über $\mathbb{C}$ nach Satz 8.2.4. Ist $A \in \mathbb{R}^{n\times n}$, so können die Eigenvektoren reell gewählt werden, weil auch die Eigenwerte reell sind. Es folgt, dass A eine Orthonormalbasis aus rellen Eigenvektoren hat. □

8.3.2. *Definitheit.*

Definition 8.3.2. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n\times n}$ selbstadjungiert. Die Matrix A heißt

1) positiv (negativ) definit, wenn für alle $x \in \mathbb{K}^n \setminus \{0\}$ gilt $x^*Ax > 0$ (bzw., $x^*Ax < 0$).
1) positiv (negativ) semidefinit, wenn für alle $x \in \mathbb{K}^n$ gilt $x^*Ax \geq 0$ (bzw., $x^*Ax \leq 0$).

Satz 8.3.3. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n\times n}$ eine selbstadjungierte Matrix. Dann gilt:*

1) *A ist genau dann positiv definit, wenn alle Eigenwerte von A größer als 0 sind.*
2) *A ist genau dann positiv semidefinit, wenn alle Eigenwerte von A größer oder gleich 0 sind.*

(Ein analoges Kriterium gilt für negativ definit und negativ semidefinit.)

Beweis. Sei A positiv definit und λ ein Eigenwert von A. Sei $x \neq 0$ ein Eigenvektor zum Eigenwert λ. Dann ist

$$\lambda\|x\|^2 = \lambda\langle x, x\rangle = \langle x, \lambda x\rangle = \langle x, Ax\rangle > 0\,.$$

Seien nun umgekehrt alle Eigenwerte von A positiv. Sei $x \neq 0$. Weil A unitär diagonalisierbar ist, gilt $x = \sum_{i=1}^k x_i$ mit $x_i \in \ker(\lambda_i - A)$, wobei die λ_i paarweise verschiedene Eigenwerte von A sind. Dabei sind nicht alle $x_i = 0$. Wir nützen nun aus, dass die Eigenräume aufeinander orthogonal stehen und erhalten

$$\begin{aligned}\langle x, Ax\rangle &= \sum_{i=1}^k\sum_{j=1}^k \langle x_i, Ax_j\rangle = \sum_{i=1}^k\sum_{j=1}^k \langle x_i, \lambda_j x_j\rangle \\ &= \sum_{i=1}^k\sum_{j=1}^k \lambda_j\langle x_i, x_j\rangle = \sum_{i=1}^k \lambda_i\|x_i\|^2 > 0.\end{aligned}$$

Der gleiche Beweis funktioniert auch für Semidefinitheit. □

Lemma 8.3.4. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n\times n}$ eine hermitesche, positiv definite Matrix. Sei $b \in \mathbb{K}^n$ und $\gamma \in \mathbb{R}$ (reell!). Wir betrachten die hermitesche $(n+1) \times (n+1)$-Blockmatrix*

$$M := \begin{pmatrix} A & b \\ b^* & \gamma \end{pmatrix}.$$

Dann ist M ist genau dann positiv definit, wenn $\det(M) > 0$ gilt.

Beweis. Weil A positiv definit ist, ist A regulär. Wir betrachten die reguläre Blockmatrix

$$T = \begin{pmatrix} E_n & -A^{-1}b \\ 0 & 1 \end{pmatrix}$$

und

$$\begin{aligned} \tilde{M} &:= T^*MT = \begin{pmatrix} E_n & 0 \\ -b^*A^{-1} & 1 \end{pmatrix} \begin{pmatrix} A & b \\ b^* & \gamma \end{pmatrix} \begin{pmatrix} E_n & -A^{-1}b \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} E_n & 0 \\ -b^*A^{-1} & 1 \end{pmatrix} \begin{pmatrix} A & 0 \\ b^* & -b^*A^{-1}b + \gamma \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & \gamma - b^*A^{-1}b \end{pmatrix}. \end{aligned}$$

Man sieht leicht dass $\det(T) = 1$. Damit ist auch $\det(T^*) = 1$, und

$$\det(\tilde{M}) = \det(T^*)\det(M)\det(T) = \det(M).$$

Damit ist

$$\det(M) = \det(\tilde{M}) = (\gamma - b^*A^{-1}b)\det(A) > 0$$

genau dann wenn $\gamma - b^*A^{-1}b > 0$.
Andererseits gilt für alle $x \in \mathbb{K}^{n+1}$

$$x^*Mx = [T^{-1}x]^*T^*MT[T^{-1}x] = [T^{-1}x]^*\tilde{M}[T^{-1}x].$$

Daher ist M genau dann positiv definit, wenn $\tilde{M}$ positiv definit ist. Die Blockdiagonalmatrix $\tilde{M}$ ist aber genau dann positiv definit, wenn ihre Diagonalblöcke A und $\gamma - b^*A^{-1}b$ positiv definit sind. Nun ist A nach Voraussetzung positiv definit, und es bleibt die Bedingung dass $\gamma - b^*A^{-1}b > 0$, d.h., $\det(M) > 0$. □

Satz 8.3.5 (Hauptminorenkriterium für Definitheit)**.** *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A = (\alpha_{i,j})_{i,j=1,\cdots,n} \in \mathbb{K}^{n\times n}$ eine hermitesche Matrix. Wir definieren für $k = 1, \cdots, n$*

$$A_k := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,k} \\ \vdots & & \vdots \\ \alpha_{k,1} & \cdots & \alpha_{k,k} \end{pmatrix}.$$

Dann ist A genau dann positiv definit, wenn für alle $k = 1, \cdots, n$ gilt $\det(A_k) > 0$.

Beweis. Sei zunächst A positiv definit. Für $y = (\eta_1, \cdots, \eta_k)^T \in \mathbb{K}^k \setminus \{0\}$ setzen wir $\tilde{y} = (\eta_1, \cdots, \eta_k, 0, \cdots, 0)^T \in \mathbb{K}^n$. Dann ist $y^*A_ky = \tilde{y}^*A\tilde{y} > 0$. Also ist A_k positiv definit und hat nach Satz 8.3.3 lauter positive Eigenwerte. Die Determinante von A_k ist aber das Produkt der Eigenwerte und muß folglich auch positiv sein.

Es sei nun $\det(A_k) > 0$ für alle k. Wir zeigen durch Induktion nach k: Die Matrix A_k ist positiv definit. Für $k = n$ ergibt sich dann: $A = A_n$ ist positiv definit.
Verankerung $k = 1$. Es ist $A_1 = (\alpha_{1,1})$. Offensichtlich ist diese 1×1-Matrix genau dann positiv definit, wenn $\alpha_{1,1} > 0$. Es ist aber $\det(A_1) = \alpha_{1,1}$.
Induktionsschritt: Sei A_k positiv definit. Wir können A_{k+1} als Blockmatrix mit einem geeigneten Vektor $b \in \mathbb{K}^n$ und $\gamma = \alpha_{k+1,k+1} \in \mathbb{R}$ schreiben:

$$A_{k+1} = \begin{pmatrix} A_k & b \\ b^* & \gamma \end{pmatrix}.$$

Nach Lemma 8.3.4 ist A_{k+1} genau dann positiv definit, wenn $\det(A_{k+1}) > 0$. □

Bemerkung 8.3.6. Beachten Sie aber, dass sich das Hauptminorenkriterium nicht auf semidefinite Matrizen verallgemeinern lässt. Es sind alle Hauptminoren der folgenden Matrix nichtnegativ, aber die Matrix ist weder positiv noch negativ semidefinit:

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

8.3.3. *Sesquilinearformen und Matrizen.*

In diesem Kapitel geht es um Sesquilinearformen. Zur Erinnerung zitieren wir hier nochmals Definition 6.2.9:

Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Abbildung

$$a : \begin{cases} V \times V & \to \mathbb{K}, \\ (x, y) & \mapsto a(x, y) \end{cases}$$

heißt

Bilinearform, wenn gilt:	Sesquilinearform, wenn gilt:
$a(z, \lambda x + \mu y) = \lambda a(z, x) + \mu a(z, y)$	$a(z, \lambda x + \mu y) = \lambda a(z, x) + \mu a(z, y)$
$a(\lambda x + \mu y, z) = \lambda a(x, z) + \mu a(y, z)$	$a(\lambda x + \mu y, z) = \overline{\lambda} a(x, z) + \overline{\mu} a(y, z)$

Eine Bilinearform bzw. Sesquilinearform a heißt

symmetrisch, wenn gilt:	hermitesch, wenn gilt:
$a(y, x) = a(x, y)$,	$a(y, x) = \overline{a(x, y)}$.

Eine Bilinearform bzw. Sesquilinearform a heißt
positiv semidefinit, wenn für alle $x \in V$ gilt: $a(x, x) \in [0, \infty)$,
positiv definit, wenn für alle $x \in V \setminus \{0\}$ gilt $a(x, x) \in (0, \infty)$.

Satz 8.3.7. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.*

1) *Sei $a : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ eine Sesquilinearform. Dann gibt es eine eindeutig bestimmte Matrix $A \in \mathbb{K}^{n \times n}$ sodass für alle $x, y \in \mathbb{K}^n$ gilt $a(x, y) = x^* A y$.*
2) *Ist $A \in \mathbb{K}^{n \times n}$, so ist die Abbildung*

$$a : \begin{cases} \mathbb{K}^n \times \mathbb{K}^n & \to \mathbb{K} \\ (x, y) & \mapsto x^* A y \end{cases}$$

   eine Sesquilinearform.
3) *Sind e_i die Einheitsvektoren in $\mathbb{K}^n$, so ist $A = (\alpha_{i,j})_{i,j=1,\cdots,n}$ gegeben durch $\alpha_{i,j} = a(e_i, e_j)$.*
4) *Die Matrix A ist genau dann selbstadjungiert, wenn a hermitesch ist.*
5) *Die Matrix A ist genau dann positiv (semi)-definit, wenn a positiv (semi)-definit ist.*

Beweis. Beweis als Übung. □

Satz 8.3.8. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $B = (b_1, \cdots, b_n)$ eine Basis des $\mathbb{K}^n$. Mit T_E^B bezeichnen wir wie immer die Übergangsmatrix von der Basis B zur Einheitsbasis, d.i. die Matrix, deren Spalten die Basisvektoren b_i sind. Sei $A \in \mathbb{K}^{n \times n}$ und a die Sesquilinearform $a(x, y) = x^* A y$.*
Sind $\tilde{x}$, $\tilde{y}$ die Koordinaten bezüglich B von Vektoren $x, y \in \mathbb{K}^n$, so ist

$$a(x, y) = \tilde{x}^* C \tilde{y} \quad \text{mit } C = (T_E^B)^* A T_E^B .$$

Beweis. Weil $x = T_E^B \tilde{x}$ und $y = T_E^B \tilde{y}$, gilt

$$x^* A y = (T_E^B \tilde{x})^* A (T_E^B \tilde{y}) = \tilde{x}^* [T_E^B]^* A T_E^B \tilde{y} \,.$$

□

Beachten Sie: Wenn A als die Matrix einer linearen Abbildung $f : x \mapsto Ax$ aufgefasst wird, geschieht der Basiswechsel durch Ähnlichkeit: $[T_E^B]^{-1} A T_E^B$.
Wenn aber A als die Matrix einer Sesquilinearform $a(x, y) = x^* A y$ aufgefasst wird, geschieht der Basiswechsel mit Hilfe der Adjungierten: $[T_E^B]^* A T_E^B$.
Wenn B eine Orthonormalbasis ist, ist $[T_E^B]^* = [T_E^B]^{-1}$, und die Basiswechsel für lineare Abbildungen und für Sesquilinearformen sind gleich.

8.3.4. *Trägheitssatz von Sylvester.*

Definition 8.3.9. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $a : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ eine hermitesche Sesquilinearform. Der Ausartungsraum von a ist definiert durch $\{x \in \mathbb{K}^n \mid (\forall y \in \mathbb{K}^n)\, a(y, x) = 0\}$.

Bemerkung 8.3.10. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n \times n}$ eine selbstadjungierte Matrix. Dann hat die Sesquilinearform $a(x, y) = x^* A y$ den Ausartungsraum $\ker(A)$.

Beweis. Liegt x im Ausartungsraum von a, so gilt $y^* A x = 0$ für alle $y \in \mathbb{K}^n$. Setzt man insbesondere $y = Ax$, so erhält man $\|Ax\|^2 = 0$, und daher ist $x \in \ker(A)$.
Ist umgekehrt $x \in \ker(A)$, so gilt für alle $y \in \mathbb{K}^n$ dass $y^* A x = y^* 0 = 0$. □

Satz 8.3.11. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $a : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ eine hermitesche Sesquilinearform. Dann gilt:*

1) *$\mathbb{K}^n$ lässt sich als direkte Summe dreier Unterräume schreiben:*

$$\mathbb{K}^n = W_0 + W_+ + W_- \,,$$

wobei W_0 der Ausartungsraum von a ist, und

$$\begin{aligned} &(\forall x \in W_+ \setminus \{0\})\ a(x, x) > 0 \,, \\ &(\forall x \in W_- \setminus \{0\})\ a(x, x) < 0 \,. \end{aligned} \tag{8.3.1}$$

Dabei können die Räume W_+, W_- so gewählt werden, dass W_+, W_- und W_0 aufeinander orthogonal stehen.

2) *Die Zerlegung aus Punkt (1) ist nicht notwendigerweise eindeutig. Sind aber $\mathbb{K}^n = W_0 + W_+^1 + W_-^1$ und $\mathbb{K}^n = W_0 + W_+^2 + W_-^2$ zwei Zerlegungen in direkte Summen (orthogonal oder nicht), sodass jeweils* (8.3.1) *gilt, dann ist $\dim(W_+^1) = \dim(W_+^2)$ und $\dim(W_-^1) = \dim(W_-^2)$.*

Beweis. Teil (1): Sei $a(x, y) = x^* A y$ mit der selbstadjungierten Matrix $A \in \mathbb{K}^{n \times n}$. Wir wissen, dass A unitär diagonalisierbar ist. Seien $\lambda_i : i = 1, \cdots, m$ die positiven Eigenwerte von A und $\mu_j : j = 1, \cdots, k$ die negativen Eigenwerte.

$$\begin{aligned} W_0 &= \ker(A) \\ W_+ &= \begin{cases} \{0\} \text{ wenn es keine positiven Eigenwerte gibt,} \\ \ker(\lambda_1 - A) + \cdots + \ker(\lambda_m - A) \text{ sonst.} \end{cases} \\ W_- &= \begin{cases} \{0\} \text{ wenn es keine negativen Eigenwerte gibt,} \\ \ker(\mu_1 - A) + \cdots + \ker(\mu_k - A) \text{ sonst.} \end{cases} \end{aligned}$$

Natürlich ist W_0 der Eigenraum von A zum Eigenwert 0. Weil sich $\mathbb{K}^n$ als die direkte Summe der Eigenräume von A schreiben lässt, ist $\mathbb{K}^n = W_0 + W_+ + W_-$. Da die Eigenräume von A aufeinander orthogonal stehen, stehen auch W_0, W_+ und W_- aufeinander orthogonal. Wegen Satz 8.3.3 ist die Einschränkung von a auf W_+ positiv definit, und die Einschränkung von a auf W_- negativ definit.

Teil (2): Nehmen wir nun zwei Zerlegungen in direkte Summen (orthogonal oder auch nicht)

$$\mathbb{K} = W_0 + W_+^1 + W_-^1 = W_0 + W_+^2 + W_-^2\,,$$

sodass (8.3.1) gilt. Wir nehmen an, es wäre $\dim(W_+^1) \neq \dim(W_+^2)$, o.B.d.A. $\dim(W_+^1) > \dim(W_+^2)$. Weil $\dim(W_+^1) + \dim(W_0) + \dim(W_-^2) > n$, kann der Durchschnitt $W_+^1 \cap (W_0 + W_-^2)$ nicht der Nullraum sein. Sei $x \neq 0$ aus diesem Durchschnitt. Wir setzen $x = y + z$ mit $y \in W_0$, $z \in W_-^2$. Dann ist

$$\begin{aligned} a(x,x) &> 0 \qquad \text{weil } x \in W_+^1 \setminus \{0\}\,, \\ a(x,x) &\leq 0 \qquad \text{weil } a(x,x) = a(y,y) + a(y,z) + a(z,y) + a(z,z) = a(z,z)\,. \end{aligned}$$

Damit erhalten wir einen Widerspruch, und es muss gelten $\dim(W_+^1) = \dim(W_+^2)$. Daraus folgt automatisch

$$\dim(W_-^1) = n - \dim(W_0) - \dim(W_+^1) = n - \dim(W_0) - \dim(W_+^2) = \dim(W_-^2)\,.$$

□

Korollar 8.3.12 (Trägheitssatz von Sylvester)**.** *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n\times n}$ eine selbstadjungierte Matrix. Sei $T \in \mathbb{K}^{n\times n}$ regulär. Sei $B = T^*AT$, seien λ_i die Eigenwerte von A und μ_i die Eigenwerte von B.*

$$\begin{aligned} \alpha_0 &= \dim(\ker(A))\,, \\ \alpha_+ &= \sum_{i:\lambda_i>0} \dim(\ker(\lambda_i - A))\,, \\ \alpha_- &= \sum_{i:\lambda_i<0} \dim(\ker(\lambda_i - A))\,. \end{aligned}$$

Ebenso definieren wir β_0, β_+ und β_- für B. Dann gilt:

$$\alpha_0 = \beta_0\,, \quad \alpha_+ = \beta_+\,, \quad \alpha_- = \beta_-\,.$$

Beweis. Sei W_0 der Ausartungsraum von A, und W_+, W_- die Summenräume aus den Eigenräumen von A zu den positiven bzw. negativen Eigenwerten. Ebenso konstruieren wir V_0, V_+, V_- für die Matrix B. Wir setzen dann $U_0 := T(V_0)$, $U_+ := T(V_+)$, und $U_- := T(V_-)$. Weil die Abbildung $x \mapsto Tx$ ein Isomorphismus ist, ist auch $\mathbb{K}^n$ die direkte Summe $\mathbb{K}^n = U_0 + U_+ + U_-$, und die Dimensionen der U_i sind die Dimensionen der entsprechenden V_i.

Nun ist $x \in V_0 = \ker(B)$ genau dann wenn $T^{-1}ATx = 0$, weil aber T regulär ist, ist das äquivalent zu $Tx \in \ker(A)$. Daher ist U_0 der Ausartungsraum von A, also $U_0 = W_0$. Ist $x \in V_+ \setminus \{0\}$ (also $Tx \in U_+$), so gilt $(Tx)^*A(Tx) = x^*Bx > 0$, daher die Einschränkung der Sesquilinearform y^*Ax auf den Raum U_+ positiv definit, ebenso ist y^*Ax auf U_- negativ definit. Wegen Satz 8.3.11 (Teil 2) gilt also für $i = 0, +, -$, dass $\dim(W_i) = \dim(U_i) = \dim(V_i)$. □

8.3.5. *Grösster Eigenwert und Matrixnorm.*

Satz 8.3.13. *Sei $A = A^* \in \mathbb{C}^{n\times n}$, dann sind nach Satz 8.3.1 alle Eigenwerte reell. Sei λ_+ der größte Eigenwert, und λ_- der kleinste Eigenwert von A. Dann gilt:*

$$\begin{aligned} \lambda_+ &= \max\{x^*Ax \mid x \in \mathbb{C}^n,\ \|x\| = 1\}\,, \\ \lambda_- &= \min\{x^*Ax \mid x \in \mathbb{C}^n,\ \|x\| = 1\}\,. \end{aligned}$$

Ist $A \in \mathbb{R}^{n\times n}$, so kann das Maximum oder Minimum über $x \in \mathbb{R}^n$ mit $\|x\| = 1$ genommen werden.

Beweis. Sei $(v_1, \cdots, v_n)$ eine Orthonormalbasis von $\mathbb{C}^n$ (bzw. $\mathbb{R}^n$) aus Eigenvektoren von A, also $Av_i = \lambda_i v_i$. Jedes $x \in \mathbb{C}^n$ hat eine Darstellung

$$x = \sum_{i=1}^{n} \xi_i v_i \,.$$

Dann ist

$$Ax = \sum_{i=1}^{n} \lambda_i v_i \quad \text{und } \langle x, Ax\rangle = \sum_{i=1}^{n} \lambda_i |\xi_i|^2 \,.$$

Nach Satz 6.4.7 ist

$$\|x\|^2 = \sum_{i=1}^{n} |\xi_i|^2 \,.$$

Wenn $\|x\| = 1$, dann gilt also

$$\lambda_- = \lambda_- \sum_{i=1}^{n} |\xi_i|^2 \le \sum_{i=1}^{n} \lambda_i |\xi_i|^2 (= x^T Ax) \le \sum_{i=1}^{n} \lambda_+ |\xi_i|^2 = \lambda_+ \,.$$

Ist andererseits $\lambda_i = \lambda_+$, so ist $\|v_i\| = 1$ und

$$v_i^T A v_i = \lambda_i \|v_i\|^2 = \lambda_i \,.$$

Ebenso ist für $\lambda_j = \lambda_-$

$$v_j^T A v_j = \lambda_j \|v_j\|^2 = \lambda_j \,.$$

Also ist insgesamt (für $\|x\| = 1$)

$$\lambda_- = v_j^T A v_j \le x^T A x \le v_i^T A v_i = \lambda_+ \,.$$

□

Definition 8.3.14. Sei $B \in \mathbb{C}^{m\times n}$ und seien $\|\cdot\|_{\mathbb{C}^m}$ und $\|\cdot\|_{\mathbb{C}^n}$ Normen auf $\mathbb{C}^m$ bzw. $\mathbb{C}^n$. Diese müssen nicht unbedingt die euklidischen Normen sein.
Die Matrixnorm von B bezüglich der Normen $\|\cdot\|_{\mathbb{C}^m}$, $\|\cdot\|_{\mathbb{C}^n}$ ist

$$\|B\| := \sup\{\|Bx\|_{\mathbb{C}^m} \mid x \in \mathbb{C}^n, \|x\|_{\mathbb{C}^n} = 1\} \,.$$

Bemerkung 8.3.15. Tatsächlich ist

$$\|B\| := \max\{\|Bx\|_{\mathbb{C}^m} \mid x \in \mathbb{C}^n, \|x\|_{\mathbb{C}^n} = 1\} \,,$$

also das Maximum. Das heißt, es gibt ein $x \in \mathbb{C}^n$ mit $\|x\|_{\mathbb{C}^n} = 1$ und $\|Bx\|_{\mathbb{C}^m} = \|B\|$. Das folgt mit Hilfe von Kompaktheitsargumenten, denn im $\mathbb{R}$ besitzt jede abgeschlossene, beschränkte Menge ein Maximum[26].

Bemerkung 8.3.16. In Definition 8.3.14 geht wesentlich ein, dass wir in endlichdimensionalen Räumen arbeiten. Sind V und W unendlichdimensionale Räume mit Normen $\|\cdot\|_V$ und $\|\cdot\|_W$ und ist $f : V \to W$ eine lineare Abbildung, dann muss die Menge

$$\{\|f(v)\|_W \mid \|v\|_V = 1\}$$

durchaus nicht beschränkt sein. Wenn sie beschränkt ist, sprechen wir von einer beschränkten oder stetigen linearen Abbildung und definieren

$$\|f\| := \sup\{\|f(x)\|_W \mid x \in V, \|x\|_V = 1\} \,.$$

[26]Mehr darüber hören Sie in einer Einführungsvorlesung über Analysis.

Satz 8.3.17. *Sei $B \in \mathbb{C}^{m\times n}$ und seien $\|\cdot\|_{\mathbb{C}^m}$ und $\|\cdot\|_{\mathbb{C}^n}$ Normen auf $\mathbb{C}^m$ bzw. $\mathbb{C}^n$. Diese müssen nicht unbedingt die euklidischen Normen sein.*
Die Matrixnorm von B bezüglich der Normen $\|\cdot\|_{\mathbb{C}^m}$, $\|\cdot\|_{\mathbb{C}^n}$ ist

$$\|B\| = \inf\{M \geq 0 \mid (\forall x \in \mathbb{C}^n)\, \|Bx\|_{\mathbb{C}^m} \leq M\|x\|_{\mathbb{C}^n}\}\,,$$

also insbesondere $\|Bx\|_{\mathbb{C}^m} \leq \|B\|\,\|x\|_{\mathbb{C}^n}$.

Beweis. Es sei

$$\mathcal{M} := \{M \geq 0 \mid (\forall x \in \mathbb{C}^n)\, \|Bx\|_{\mathbb{C}^m} \leq M\|x\|_{\mathbb{C}^n}\}\,.$$

Ist $M \in \mathcal{M}$ und $x \in \mathbb{C}^n$ mit $\|x\|_{\mathbb{C}^n} = 1$, so ist also $\|Bx\|_{\mathbb{C}^m} \leq M\|x\|_{\mathbb{C}^n} = M$. Also ist für alle $M \in \mathcal{M}$

$$\|B\| = \sup\{\|Bx\|_{\mathbb{C}^m} \mid \|x\|_{\mathbb{C}^n} = 1\} \leq M\,.$$

Sei umgekehrt $x \in \mathbb{C}^n \setminus \{0\}$. Wir setzen $\gamma = \|x\|_{\mathbb{C}^n}$ und $y := \frac{1}{\gamma}\,x$, sodass $\|y\|_{\mathbb{C}^n} = 1$. Es ist dann

$$\|Bx\|_{\mathbb{C}^m} = \|\gamma By\|_{\mathbb{C}^m} = \gamma\,\|By\|_{\mathbb{C}^m} \leq \gamma\,\|B\| = \|B\|\,\|x\|_{\mathbb{C}^n}\,.$$

□

Satz 8.3.18. *Wir betrachten beliebige Normen $\|\cdot\|_{\mathbb{C}^j}$ auf $\mathbb{C}^j$ mit $j = k, m, n$, und dazu die passenden Matrixnormen. Seien $\alpha \in \mathbb{C}$, $A, B \in \mathbb{C}^{m\times n}$, $C \in \mathbb{C}^{k\times m}$. Dann gilt*

(1) *Wenn $\|A\| = 0$, dann ist $A = 0$.*
(2) $\|\alpha A\| = |\alpha|\,\|A\|$.
(3) $\|A + B\| \leq \|A\| + \|B\|$.
(4) $\|CA\| \leq \|C\|\,\|A\|$.

Insbesondere bildet $\mathbb{C}^n$ mit der Matrixnorm einen normierten Vektorraum.

Beweis. Ist $\|A\| = 0$, so gilt $Ax = 0$ für alle $x \in \mathbb{C}^n$. Also muss A die Nullmatrix sein. Ferner ist

$$\begin{aligned}
\|\alpha A\| &= \sup_{\|x\|=1} \|\alpha Ax\| = |\alpha| \sup_{\|x\|=1} \|Ax\| = |\alpha|\,\|A\|\,,\\
\|A + B\| &= \sup_{\|x\|=1} \|(A+B)x\| \leq \sup_{\|x\|=1} (\|Ax\| + \|Bx\|)\\
&\leq \sup_{\|x\|=1} \|Ax\| + \sup_{\|x\|=1} \|Bx\| = \|A\| + \|B\|\,,\\
\|CA\| &= \sup_{\|x\|=1} \|CAx\| \leq \sup_{\|x\|=1} \|C\|\,\|Ax\| = \|C\| \sup_{\|x\|=1} \|Ax\| = \|C\|\,\|A\|\,.
\end{aligned}$$

□

Satz 8.3.19. *Sei $A \in \mathbb{C}^{n\times n}$ regulär und $\|\cdot\|$ eine beliebige Norm auf $\mathbb{C}^n$. Sei $\|A\|$ die dazu passende Matrixnorm. Dann ist für alle $b \in \mathbb{C}^n$ das Gleichungssystem $Ax = b$ eindeutig lösbar. Wenn b einer (kleinen) "Störung" $\Delta b \in \mathbb{C}^n$ unterworfen ist, also statt $Ax = b$ das Gleichungssystem $Ay = (b + \Delta b)$ gelöst wird, dann ist $y = x + \Delta x$ mit $\Delta x = A^{-1}\Delta b$.*
Der "relative Fehler" $\|\Delta x\| / \|x\|$ läßt sich dann abschätzen:

$$\frac{\|\Delta x\|}{\|x\|} \leq \|A\|\,\|A^{-1}\|\,\frac{\|\Delta b\|}{\|b\|}\,.$$

Beweis. Es ist

$$\begin{aligned} \|\Delta x\| &\leq \|A^{-1}\|\,\|\Delta b\|\,, \\ \|b\| &\leq \|A\|\,\|x\|\,, \quad \text{also} \\ \frac{\|\Delta x\|}{\|x\|} &\leq \frac{\|A^{-1}\|\,\|\Delta b\|}{\|b\|/\|A\|} = \|A^{-1}\|\,\|A\|\,\frac{\|\Delta b\|}{\|b\|}\,. \end{aligned}$$

□

Definition 8.3.20. Sei $A \in \mathbb{C}^{n\times n}$ regulär. Sei $\|\cdot\|$ eine Norm auf $\mathbb{C}^n$ und $\|A\|$ die dazu passende Matrixnorm.
Die Zahl $\|A^{-1}\|\,\|A\|$ heißt die Konditionszahl von A.

Bemerkung 8.3.21. Wegen Satz 8.3.19 ist also die Konditionszahl ein Maß, das abschätzt, wie stark sich relative Fehler fortpflanzen. Hat eine Matrix A eine hohe Konditionszahl, so kann in der Gleichung $Ax = b$ auch eine (im Verhältnis zu $\|b\|$) kleine Änderung von b eine (im Verhältnis zu $\|x\|$) große Änderung von x bewirken.

Bemerkung 8.3.22. Die Konditionszahl einer regulären Matrix ist immer größer oder gleich 1.

Beweis.

$$\|E_n\| = \|A^{-1}A\| \leq \|A^{-1}\|\,\|A\|\,.$$

□

Wenn man die euklidischen Normen zugrunde legt, so läßt sich die Matrixnorm berechnen:

Satz 8.3.23. *Sei $B \in \mathbb{C}^{m\times n}$. Sei λ_+ und λ_- der größte und kleinste Eigenwert von B^*B. Wir betrachten die Matrixnormen bezüglich der euklidischen Normen auf $\mathbb{C}^m$ und $\mathbb{C}^n$.*

1. *Die Matrixnorm $\|B\|$ ist die Wurzel $\sqrt{\lambda_+}$ des größten Eigenwertes von B^*B.*
2. *Die Konditionszahl von B ist die Wurzel $\sqrt{\lambda_+/\lambda_-}$ des Verhältnisses zwischen dem größten und kleinsten Eigenwert von B^*B.*

Beweis. Sei λ_+ der größte Eigenwert von B^*B. Dann ist nach Satz 8.3.13 (mit $A = B^*B$)

$$\begin{aligned} \|B\|^2 &= \sup\{\|Bx\|^2 \mid x \in \mathbb{C}^n,\, \|x\| = 1\} \\ &= \sup\{x^*B^*Bx \mid x \in \mathbb{C}^n,\, \|x\| = 1\} = \lambda_+\,. \end{aligned}$$

Die Eigenwerte von B^*B sind zugleich die Eigenwerte von BB^*, denn ist $B^*Bx = \lambda x$, so setzen wir $y = Bx$ und erhalten $BB^*y = BB^*Bx = B(\lambda x) = \lambda y$. Die Eigenwerte von $(BB^*)^{-1}$ sind die Reziprokwerte der Eigenwerte von BB^*, denn ist $BB^*x = \lambda x$, so ist $\lambda^{-1}x = (BB^*)^{-1}x$. Damit ist der größte Eigenwert von $(B^{-1})^*B^{-1} = (BB^*)^{-1}$ der Reziprokwert des kleinsten Eigenwertes von BB^*, also zugleich der Reziprokwert des kleinsten Eigenwertes von B^*B. Damit ist

$$\|B\|\,\|B^{-1}\| = \sqrt{\lambda_+}\,\sqrt{\frac{1}{\lambda_-}} = \sqrt{\frac{\lambda_+}{\lambda_-}}\,.$$

□

8.4. **Quadriken.**

Quadriken sind Kurven in $\mathbb{R}^2$ bzw. Hyperflächen in $\mathbb{R}^n$, welche durch quadratische Gleichungen beschrieben werden. Im zweidimensionalen euklidischen Raum sind das also die Kegelschnitte.

8.4.1. *Definition und Schreibweisen für Quadriken.*

Definition 8.4.1. Sei $A \in \mathbb{R}^{n\times n}$ selbstadjungiert, $b \in \mathbb{R}^n$ und $\gamma \in \mathbb{R}$. Dann heißt die Lösungsmenge der Gleichung

$$x^T Ax + 2b^T x + \gamma = 0 \tag{8.4.1}$$

eine Quadrik in $\mathbb{R}^n$.

Bemerkung 8.4.2. Ist $A = (\alpha_{ij})_{i,j=1\cdots n}$ symmetrisch, $b = (\beta_i)_{i=1\cdots n}$ und $x = (\xi_i)_{i=1\cdots n}$ so lautet (8.4.1) ausgeschrieben:

$$\begin{aligned} 0 &= \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij}\xi_i\xi_j + \sum_{i=1}^n 2\beta_i\xi_i + \gamma \\ &= \sum_{i=1}^n \alpha_{ii}\xi_i^2 + \sum_{i,j=1\cdots n,\, i<j} (\alpha_{ij}+\alpha_{ji})\xi_i\xi_j + \sum_{i=1}^n 2\beta_i\xi_i + \gamma\,. \end{aligned}$$

Wird eine quadratische Gleichung in $\mathbb{R}^n$ in Matrixform umgeschrieben, so ist also zu beachten, dass im Matrixprodukt $x^T Ax$ die gemischten Terme immer doppelt auftreten.

Bemerkung 8.4.3. Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch, $b \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$. Die Quadrik

$$x^T Ax + 2b^T x + \gamma = 0$$

lässt sich umschreiben in eine Form mit einer einzigen $(n+1)\times(n+1)$-Matrix:

$$\begin{pmatrix} x \\ 1 \end{pmatrix}^T \begin{pmatrix} A & b \\ b^T & \gamma \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = 0\,. \tag{8.4.2}$$

Beweis.

$$\begin{aligned} &\begin{pmatrix} x \\ 1 \end{pmatrix}^T \begin{pmatrix} A & b \\ b^T & \gamma \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ 1 \end{pmatrix}^T \begin{pmatrix} Ax+b \\ b^T x + \gamma \end{pmatrix} \\ = \;& x^T(Ax+b) + (b^T x + \gamma) = x^T Ax + 2b^T x + \gamma\,. \end{aligned}$$

□

Beispiel 8.4.4. Wir betrachten in $\mathbb{R}^2$ die Gleichung

$$7\xi_1^2 - 4\xi_2^2 + 6\xi_1\xi_2 - 8\xi_1 + 2\xi_2 + 15 = 0\,.$$

Wir erhalten in Matrixform

$$x^T \begin{pmatrix} 7 & 3 \\ 3 & -4 \end{pmatrix} x + 2 \begin{pmatrix} -4 \\ 1 \end{pmatrix}^T x + 15 = 0$$

oder

$$\begin{pmatrix} x \\ 1 \end{pmatrix}^T \begin{pmatrix} 7 & 3 & -4 \\ 3 & -4 & 1 \\ -4 & 1 & 15 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = 0\,.$$

Lemma 8.4.5. *Sei $M \in \mathbb{R}^{m\times n}$, $p \in \mathbb{R}^m$ und f die Abbildung*

$$f : \begin{cases} \mathbb{R}^n & \to \mathbb{R}^m\,, \\ x & \mapsto Mx + p\,. \end{cases}$$

So lässt sich f in der Form schreiben:

$$\begin{pmatrix} f(x) \\ 1 \end{pmatrix} = \begin{pmatrix} M & p \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix}\,.$$

Beweis. Einfach nachrechnen. □

8.4.2. *Hauptachsentransformation.*

Satz 8.4.6 (Hauptachsentransformation von Quadriken)**.** *Sei $A = A^T \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$ und $Q := \{x \in \mathbb{R}^n \mid x^T Ax + 2b^T x + c = 0\}$ die dazu gehörende Quadrik. Dann gibt es eine orthogonale Matrix $T \in \mathbb{R}^{n\times n}$, eine Diagonalmatrix $D \in \mathbb{R}^{n\times n}$, Vektoren $s, q \in \mathbb{R}^n$ und eine Zahl $\delta \in \mathbb{R}$, sodass die Koordinatentransformation*

$$\begin{cases} \mathbb{R}^n & \to \mathbb{R}^n\,, \\ y & \mapsto Ty + q\,, \end{cases}$$

die Quadrik Q folgendermaßen diagonalisiert:

$$Q = \{Ty + q \mid y^T Dy + 2s^T y + \delta = 0\}\,.$$

- *Dabei sind die Spalten von T eine Orthonormalbasis von Eigenvektoren von A, und die Diagonalelemente von D sind die Eigenwerte λ_i von A.*
- *Der Vektor $s = (\sigma_i)_{i=1\cdots n}$ kann so konstruiert werden, dass $\sigma_i = 0$ falls $\lambda_i \neq 0$.*
- *Wenn es ein i gibt mit $\sigma_i \neq 0$ und $\lambda_i = 0$, dann kann $\delta = 0$ gesetzt werden.*

Beweis. Der Beweis wird durch den folgenden Algorithmus gegeben: □

Algorithmus 8.4.7 (Hauptachsentransformation von Quadriken)**.** Zur Diagonalisierung der Quadrik

$$x^T Ax + 2b^T x + \gamma$$

im Sinne von Satz 8.4.6 in Form

$$y^T Dy + 2s^T y + \delta$$

mit der Koordinatentransformation $x = Ty + q$ wird folgendermaßen vorgegangen:

(1) Bestimme eine Orthonormalbasis aus Eigenvektoren von A und setze diese Basis zu einer Matrix T zusammen. Setze $r = T^T b$. Ist dann $x = Tz$, so ist

$$Q = \{Tz \mid z^T Dz + 2r^T z + c = 0\}\,.$$

(2) Wir schreiben $r = (\rho_i)_{i=1\cdots n}$. Konstruiere $p = (\phi_i)_{i=1\cdots n}, s = (\sigma_i)_{i=1\cdots n} \in \mathbb{K}^n$ komponentenweise:

$$\begin{aligned} \phi_i &:= \begin{cases} -\frac{\rho_i}{\lambda_i} & \text{wenn } \lambda_i \neq 0\,, \\ 0 & \text{wenn } \lambda_i = 0\,. \end{cases} \\ \sigma_i &:= \begin{cases} 0 & \text{wenn } \lambda_i \neq 0\,, \\ r_i & \text{wenn } \lambda_i = 0\,. \end{cases} \\ \delta &:= \gamma - \sum_{i:\lambda_i\neq 0} \frac{\rho_i^2}{\lambda_i}\,. \end{aligned}$$

Ist dann $z = y + p$ und $q = Tp$, so ist

$$\begin{aligned} Q &= \{T(y+p) \mid y^T Dy + 2s^T y + \delta = 0\} \\ &= \{Ty + q \mid y^T Dy + 2s^T y + \delta = 0\}\,. \end{aligned}$$

(3) Falls es k mit der Eigenschaft $\lambda_k = 0$ und $\sigma_k \neq 0$ gibt, dann wähle davon ein beliebiges (nur eines!). Konstruiere $\tilde{p} = (\psi_i)_{i=1\cdots n}$ komponentenweise:

$$\psi_i := \begin{cases} -\frac{\delta}{2\sigma_k} & \text{wenn } i = k\,, \\ 0 & \text{sonst}\,. \end{cases}$$

Ist dann $y = \tilde{y} + \tilde{p}$ und $\tilde{q} = T(\tilde{p} + p)$, dann ist

$$\begin{aligned} Q &= \{T(\tilde{y} + \tilde{p} + p) \mid \tilde{y}^T D\tilde{y} + 2s^T \tilde{y} = 0\} \\ &= \{T\tilde{y} + \tilde{q} \mid \tilde{y}^T D\tilde{y} + 2s^T \tilde{y} = 0\}\,. \end{aligned}$$

Beweis. Der erste Schritt ist einfach die Diagonalisierung von A, wobei wir anschließend Tz für x einsetzen. Wir haben dann

$$x^T Ax + 2b^T x + c = z^T T^T ATz + 2b^T Tz + c = z^T Dz + 2(T^T b)^T z + c\,.$$

Für den zweiten Schritt schreiben wir die Quadrik in Form (8.4.2):

$$\begin{pmatrix} z \\ 1 \end{pmatrix}^T \begin{pmatrix} D & r \\ r^T & \gamma \end{pmatrix} \begin{pmatrix} z \\ 1 \end{pmatrix} = 0\,.$$

Der Kürze halber setzen wir $I = \{i \mid \lambda_i \neq 0\}$. Wir können o.B.d.A. annehmen, dass die Indices so sortiert sind, dass die $i \in I$ zuerst kommen. Unsere Form läßt sich also in weitere Blöcke aufschließen, z.B. $z^T = \begin{pmatrix} z_I & z_{I\!\!I} \end{pmatrix}$. Das ergibt

$$\begin{pmatrix} z_I \\ z_{I\!\!I} \\ 1 \end{pmatrix}^T \begin{pmatrix} D_I & 0 & r_I \\ 0 & 0 & r_{I\!\!I} \\ r_I^T & r_{I\!\!I}^T & \gamma \end{pmatrix} \begin{pmatrix} z_I \\ z_{I\!\!I} \\ 1 \end{pmatrix} = 0\,.$$

Nach unserer Konstruktion ist

$$p = \begin{pmatrix} -D_I^{-1} r_I \\ 0 \end{pmatrix}.$$

Letztlich ist nach Lemma 8.4.5

$$\begin{pmatrix} z+p \\ 1 \end{pmatrix} = \begin{pmatrix} E_n & p \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} E & 0 & -D_I^{-1} r_I \\ 0 & E & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z_I \\ z_{I\!\!I} \\ 1 \end{pmatrix}.$$

Also ist

$$\begin{aligned}
0 &= \begin{pmatrix} z_I \\ z_{I\!\!I} \\ 1 \end{pmatrix}^T \begin{pmatrix} D_I & 0 & r_I \\ 0 & 0 & r_{I\!\!I} \\ r_I^T & r_{I\!\!I}^T & \gamma \end{pmatrix} \begin{pmatrix} z_I \\ z_{I\!\!I} \\ 1 \end{pmatrix}^T \\
&= \begin{pmatrix} y_I \\ y_{I\!\!I} \\ 1 \end{pmatrix}^T \begin{pmatrix} E & 0 & 0 \\ 0 & E & 0 \\ -r_I^T D_I^{-1} & 0 & 1 \end{pmatrix} \begin{pmatrix} D_I & 0 & r_I \\ 0 & 0 & r_{I\!\!I} \\ r_I^T & r_{I\!\!I}^T & \gamma \end{pmatrix} \begin{pmatrix} E & 0 & D_I^{-1} r_I \\ 0 & E & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_I \\ y_{I\!\!I} \\ 1 \end{pmatrix}^T \\
&= \begin{pmatrix} y_I \\ y_{I\!\!I} \\ 1 \end{pmatrix}^T \begin{pmatrix} D_I & 0 & r_I \\ 0 & 0 & r_{I\!\!I} \\ 0 & r_{I\!\!I}^T & -r_I^T D_I^{-1} r_I + \gamma \end{pmatrix} \begin{pmatrix} E & 0 & -D_I^{-1} r_I \\ 0 & E & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_I \\ y_{I\!\!I} \\ 1 \end{pmatrix}^T \\
&= \begin{pmatrix} y_I \\ y_{I\!\!I} \\ 1 \end{pmatrix}^T \begin{pmatrix} D_I & 0 & 0 \\ 0 & 0 & r_{I\!\!I} \\ 0 & r_{I\!\!I}^T & -r_I^T D_I^{-1} r_I + \gamma \end{pmatrix} \begin{pmatrix} y_I \\ y_{I\!\!I} \\ 1 \end{pmatrix}^T \\
&= \begin{pmatrix} y \\ 1 \end{pmatrix}^T \begin{pmatrix} D & s \\ s^T & \delta \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix},
\end{aligned}$$

mit s und δ wie oben definiert.

Für den letzten Schritt sei o.B.d.A. $k = n$ und schreiben in der Blockmatrix die k-te Komponente separat, z.B. $y^T = \begin{pmatrix} y_0^T & \eta_n \end{pmatrix}$. Nach Konstruktion ist $\tilde{p} = \begin{pmatrix} 0 & -\delta/(2\sigma_n) \end{pmatrix}^T$. Es ist $y = \tilde{y} + \tilde{p}$, also nach Lemma 8.4.5

$$\begin{pmatrix} y \\ 1 \end{pmatrix} = \begin{pmatrix} y_0 \\ \eta_n \\ 1 \end{pmatrix} = \begin{pmatrix} E & 0 & \tilde{p}_0 \\ 0 & 1 & -\delta/(2\sigma_n) \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{y}_0 \\ \tilde{\eta}_n \\ 1 \end{pmatrix}.$$

Also ist

$$\begin{aligned}
0 &= \begin{pmatrix} y_0 \\ \eta_n \\ 1 \end{pmatrix}^T \begin{pmatrix} D_0 & 0 & s_0 \\ 0 & 0 & \sigma_n \\ s_0^T & \sigma_n & \delta \end{pmatrix} \begin{pmatrix} y_0 \\ \eta_n \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} \tilde{y}_0 \\ \tilde{\eta}_n \\ 1 \end{pmatrix}^T \begin{pmatrix} E & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\delta/(2\sigma_n) & 1 \end{pmatrix} \begin{pmatrix} D_0 & 0 & s_0 \\ 0 & 0 & \sigma_n \\ s_0^T & \sigma_n & \delta \end{pmatrix} \begin{pmatrix} E & 0 & 0 \\ 0 & 1 & -\delta/(2\sigma_n) \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{y}_0 \\ \tilde{\eta}_n \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} \tilde{y}_0 \\ \tilde{\eta}_n \\ 1 \end{pmatrix}^T \begin{pmatrix} D_0 & 0 & s_0 \\ 0 & 0 & \sigma_n \\ s_0^T & \sigma_n & \delta/2 \end{pmatrix} \begin{pmatrix} E & 0 & 0 \\ 0 & 1 & -\delta/(2\sigma_n) \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{y}_0 \\ \tilde{\eta}_n \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} \tilde{y}_0 \\ \tilde{\eta}_n \\ 1 \end{pmatrix}^T \begin{pmatrix} D_0 & 0 & s_0 \\ 0 & 0 & \sigma_n \\ s_0^T & \sigma_n & 0 \end{pmatrix} \begin{pmatrix} \tilde{y}_0 \\ \tilde{\eta}_n \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} \tilde{y} \\ 1 \end{pmatrix}^T \begin{pmatrix} D & s \\ s^T & 0 \end{pmatrix} \begin{pmatrix} \tilde{y} \\ 1 \end{pmatrix} .
\end{aligned}$$

□

Beispiel 8.4.8. Führen Sie die Hauptachsentransformation für die folgende Quadrik in $\mathbb{R}^3$ durch und versuchen Sie, sich die Fläche anschaulich vorzustellen.

$$-4\xi_1^2 - 7\xi_2^2 + 2\xi_3^2 - 20\xi_1\xi_2 + 16\xi_1\xi_3 + 4\xi_2\xi_3 + 50\xi_1 + 26\xi_2 - 10\xi_3 + 90 = 0\,.$$

Lösung: Wir übersetzen die Gleichung zuerst in Matrixform, dabei bedenken wir, dass gemischte Terme $\xi_i\xi_j$ mit $i \neq j$ in der Matrix je halbe-halbe auf zwei Plätze aufgeteilt werden, und dass in der üblichen Schreibweise vor dem Term $b^T x$ eine zwei steht.

$$\begin{aligned}
x^T \begin{pmatrix} -4 & -10 & 8 \\ -10 & -7 & 2 \\ 8 & 2 & 2 \end{pmatrix} + 2 \begin{pmatrix} 25 \\ 13 \\ -5 \end{pmatrix}^T x + 90 &= 0\,, \\
x^T A x + 2b^T x + x &= 0\,.
\end{aligned}$$

Wir diagonalisieren A. Das charakteristische Polynom ist (z.B. mittels Satz 7.2.12)

$$\lambda^3 + 9\lambda^2 - 162\lambda = 0\,.$$

Man sieht sofort die Nullstelle $\lambda_1 = 0$, nach Abdividieren dieser Nullstelle erhält man die quadratische Gleichung

$$\lambda^2 + 9\lambda^2 - 162 = 0$$

mit den Lösungen $\lambda_2 = 9$, $\lambda_3 = -18$.

Durch Lösen der linearen Gleichungssysteme $(A - \lambda_i)x = 0$ für die drei Eigenwerte erhalten wir die Eigenvektoren

$$\begin{pmatrix} 1 \\ -2 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -0.5 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ -0.5 \end{pmatrix}.$$

Man kann sich leicht überzeugen, dass diese drei Vektoren aufeinander orthogonal stehen, wie es die Theorie ja voraussagt. Damit sie auch ein Ortho*normal*system bilden, skalieren wir sie auf Norm eins um und fügen sie dann zu einer Matrix T zusammen. Es ist dann $T^T A T = D$ eine Diagonalmatrix mit den Eigenwerten von

A auf der Diagonalen:

$$T = \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ -2 & 2 & -1 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -18 \end{pmatrix}.$$

Wir setzen nun

$$x = Tz$$

und berechnen

$$r = T^T b = \frac{1}{3}\begin{pmatrix} 1 & -2 & -2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix}\begin{pmatrix} 25 \\ 13 \\ -5 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \\ 27 \end{pmatrix}.$$

In den neuen Koordinaten erhalten wir die Quadrik

$$z^T \begin{pmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -18 \end{pmatrix} z + 2\begin{pmatrix} 3 \\ 9 \\ 27 \end{pmatrix} x + 90 = 0\,.$$

Wir annullieren nun die Terme in r für die Koordinaten, zu denen Eigenwerten ungleich Null gehören. Dazu setzen wir

$$z = y + p \quad \text{mit} \quad p = \begin{pmatrix} 0 \\ -9/9 \\ 18/27 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1.5 \end{pmatrix}.$$

In r werden nun die zweite und dritte Koordinate auf Null gesetzt, die neue Konstante ist

$$d = 90 - 9^2/9 - (27)^2/(-18) = 121.5\,.$$

Wir erhalten

$$y^T \begin{pmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -18 \end{pmatrix} y + 2\begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} y + 121.5 = 0\,.$$

Letztlich können wir den konstanten Term auf Null bringen, indem wir die erste Koordinate verschieben:

$$y = \tilde{y} + \tilde{p} \quad \text{mit } \tilde{p} = \begin{pmatrix} -121.5/(2\cdot 3) \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -20.25 \\ 0 \\ 0 \end{pmatrix}.$$

Wir erhalten

$$\tilde{y}^T \begin{pmatrix} 0 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -18 \end{pmatrix} \tilde{y} + 2\begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} \tilde{y} = 0\,.$$

Nun rekonstruieren wir die alten Koordinaten aus den neuen:

$$\begin{aligned} x &= Tz = T(\tilde{y} + p + \tilde{p}) = T\tilde{y} + T(p + \tilde{p}) = T\tilde{y} + q\,, \\ q &= \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ -2 & 2 & -1 \end{pmatrix}\begin{pmatrix} -20.25 \\ -1 \\ 1.5 \end{pmatrix} \approx \begin{pmatrix} -6.42 \\ 14.83 \\ 12.33 \end{pmatrix}. \end{aligned}$$

Das heißt, das neue Koordinatensystem, in alten Koordinaten angegeben, ist:

$$\text{Ursprung: } \begin{pmatrix} -6.42 \\ 14.83 \\ 12.33 \end{pmatrix}, \text{ Richtungen der Achsen: } \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \ \frac{1}{3}\begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}, \ \frac{1}{3}\begin{pmatrix} -2 \\ 2 \\ -1 \end{pmatrix}.$$

Um uns ein Bild von der Quadrik zu machen, schreiben wir sie nun in einzelne (neue) Koordinaten um:

$$9\eta_2^2 - 18\eta_3^2 + 6\eta_1 = 0\,, \text{das heisst}$$
$$9\eta_2^2 - 18\eta_3^2 = -6\eta_1\,.$$

Fixieren wir nun jeweils η_1 und betrachten die Quadrik für η_2, η_3.

- Für $\eta_1 = 0$ erhalten wir

  $$9\eta_2^2 - 18\eta_3^2 = 0\,,\ \text{d.h. } \frac{1}{2}\eta_2^2 - \eta_3^2 = 0\,,\ \text{d.h. } \eta_3 = \pm\frac{1}{\sqrt{2}}\,\eta_2\,.$$

  die Lösungsmenge ist also ein Geradenpaar, das sich im Nullpunkt schneidet.
- Für $\eta_1 > 0$ setzen wir $\alpha = \sqrt{\eta_1/3}$, also $6\eta_1 = 18\alpha^2$. (Der Faktor $1/3$ dient nur dazu, dass die Zahlenwerte glatter aufgehen.)

  $$\frac{\eta_2^2}{2\alpha^2} - \frac{\eta_3^2}{\alpha^2} = 1\,,$$

  das ist eine Hyperbel mit den Halbachsen $\alpha\sqrt{2}$ und α, deren Scheitelpunkte bei $\eta_2 = \pm\alpha\sqrt{2}$, $\eta_3 = 0$ liegen. Die Zweige öffnen sich also in η_2-Richtung.
- Für $\eta_1 = -3\alpha^2 < 0$ erhalten wir

  $$-\frac{\eta_2^2}{2\alpha^2} + \frac{\eta_3^2}{\alpha^2} = 1\,,$$

  das ist eine Hyperbel mit den Halbachsen $\alpha\sqrt{2}$ und α, deren Scheitelpunkte bei $\eta_3 = \pm\alpha$, $\eta_2 = 0$ liegen. Die Zweige öffnen sich also in η_3-Richtung.

Wir erhalten also eine Schar von Hyperbeln, deren Öffnungsrichtung bei $\eta_1 = 0$ umschaltet, und deren Halbachsen proportional zu $\sqrt{|\eta_1|}$ sind. □

8.4.3. *Schnitt einer Quadrik mit einem affinen Unterraum.*

Satz 8.4.9. *Der Schnitt einer Quadrik in $\mathbb{R}^n$ mit einem affinen Unterraum U des $\mathbb{R}^n$ ist wieder eine Quadrik in U.*

Beweis. Sei Q eine Quadrik, die wir in der Schreibweise (8.4.2) angeben:

$$Q = \{x \in \mathbb{R}^n \mid \begin{pmatrix} x \\ 1 \end{pmatrix}^T \begin{pmatrix} A & b \\ b^T & \gamma \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = 0\}\,.$$

Sei U ein m-dimensionaler affiner Unterraum, $p \in U$ beliebig, und $(w_1, \cdots, w_m)$ eine Basis des Tangentialraumes von U, d.h.,

$$U = \{p + \sum_{i=1}^{n} \eta_i w_i \mid \eta_i \in \mathbb{R}\}\,.$$

Setzen wir die Vektoren w_i zu einer Matrix $W \in \mathbb{R}^{n\times m}$ zusammen, so erhalten wir

$$U = \{x \in \mathbb{R}^n \mid (\exists y \in \mathbb{R}^m) \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} W & p \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix}\}\,.$$

Wir betrachten den Vektor y in der obigen Gleichung als den Koordinatenvektor von x in U.

Nun ist y genau dann der U-Koordinatenvektor eines Vektors $x \in U \cap Q$, wenn gilt

$$\begin{aligned} 0 &= \begin{pmatrix} x \\ 1 \end{pmatrix}^T \begin{pmatrix} A & b \\ b^T & \gamma \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} y \\ 1 \end{pmatrix}^T \begin{pmatrix} W^T & 0 \\ p^T & 1 \end{pmatrix} \begin{pmatrix} A & b \\ b^T & \gamma \end{pmatrix} \begin{pmatrix} W & p \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} y \\ 1 \end{pmatrix}^T \begin{pmatrix} W^T A W & W^T(Ap+b) \\ (p^T A + b^T)W & p^T Ap + 2b^T p + \gamma \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix} . \end{aligned}$$

Das ist aber die Gleichung einer Quadrik. □

Beispiel 8.4.10. Wir schneiden die Quadrik Q

$$x^T \begin{pmatrix} 2.6 & 0 & 4.8 \\ 0 & 16 & 0 \\ 4.8 & 0 & 5.4 \end{pmatrix} x + 2 \begin{pmatrix} 10 \\ -16 \\ 15 \end{pmatrix}^T x + 51 = 0$$

mit der Ebene E in Parameterform

$$x = \begin{pmatrix} 7.6 \\ 1 \\ -8.2 \end{pmatrix} + \eta_1 \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} + \eta_2 \begin{pmatrix} 16 \\ 15 \\ -12 \end{pmatrix} .$$

Lösung: Damit unsere neuen Koordinaten in der Ebene die Längen richtig wiedergeben, konstruieren wir uns ein Orthonormalsystem aus den Basisvektoren des Tangentialraumes. In unserem Fall stehen die Vektoren bereits orthogonal, wir müssen sie nur auf Norm Eins skalieren.

$$x = \begin{pmatrix} 7.6 \\ 1 \\ -8.2 \end{pmatrix} + \eta_1 \begin{pmatrix} 0.6 \\ 0 \\ 0.8 \end{pmatrix} + \eta_2 \begin{pmatrix} 0.64 \\ 0.6 \\ -0.48 \end{pmatrix} .$$

Wir erhalten dann für die Koordinaten der Punkte im Durchschnitt $Q \cap E$ (MATLAB!)

$$\begin{pmatrix} y \\ 1 \end{pmatrix}^T \begin{pmatrix} 0.6 & 0 & 0.8 & 0 \\ 0.64 & 0.6 & -0.48 & 0 \\ 7.6 & 1 & -8.2 & 1 \end{pmatrix} \begin{pmatrix} 2.6 & 0 & 4.8 & 10 \\ 0 & 16 & 0 & -16 \\ 4.8 & 0 & 5.4 & 15 \\ 10 & -16 & 15 & 51 \end{pmatrix} \begin{pmatrix} 0.6 & 0.64 & 7.6 \\ 0 & 0.6 & 1 \\ 0.8 & -0.48 & -8.2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} y \\ 1 \end{pmatrix}^T \begin{pmatrix} 9 & 0 & 0 \\ 0 & 5.12 & -9.6 \\ 0 & -9.6 & 144 \end{pmatrix} \begin{pmatrix} y \\ 1 \end{pmatrix} = 0 .$$

Mit dem Ansatz

$$y = \tilde{y} + \begin{pmatrix} 0 \\ 9.6/5.12 \end{pmatrix} = \tilde{y} + \begin{pmatrix} 0 \\ 1.875 \end{pmatrix}$$

erhalten wir :

$$\begin{aligned} 0 &= \begin{pmatrix} \tilde{y} \\ 1 \end{pmatrix}^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1.875 & 1 \end{pmatrix} \begin{pmatrix} 9 & 0 & 0 \\ 0 & 5.12 & -9.6 \\ 0 & -9.6 & -144 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1.875 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tilde{y} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \tilde{y} \\ 1 \end{pmatrix}^T \begin{pmatrix} 9 & 0 & 0 \\ 0 & 5.12 & 0 \\ 0 & 0 & -162 \end{pmatrix} \begin{pmatrix} \tilde{y} \\ 1 \end{pmatrix} = 0 , \end{aligned}$$

das ist eine Ellipse in Hauptachsenform

$$9\tilde{\eta}_1^2 + 5.12\tilde{\eta}_2^2 = 162 .$$

□

8.4.4. *Tangentialhyperebenen an eine Quadrik.*

Ein Kapitel über Quadriken ist unvollständig ohne den Begriff der Polaren, und diese wiederum benötigt den Begriff der Tangentialhypereben. Dazu braucht man einige Grundkenntnisse aus der Analysis in $\mathbb{R}^n$, die wir hier natürlich nur skizzieren können.

Definition 8.4.11. Es sei $\Phi : \mathbb{R}^n \to \mathbb{R}$ eine Funktion. Die partielle Ableitung $\frac{\partial}{\partial \xi_k}\Phi(x)$ von Φ nach ξ_k an einem Punkt $x = (\xi_1, \cdots, \xi_n)^T$ erhält man, indem man die Variablen ξ_i für $i \neq k$ als Konstanten behandelt, und die Funktion Φ nach der Variablen ξ_k differenziert.

$$\frac{\partial}{\partial \xi_i}\Phi(x) = \lim_{h\to 0} \frac{1}{h}\left[\Phi(\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k + h \\ \vdots \\ \xi_n \end{pmatrix}) - \Phi(\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k \\ \vdots \\ \xi_n \end{pmatrix})\right] .$$

Wenn in allen Punkten $x \in \mathbb{R}^n$ alle partiellen Ableitungen von Φ existieren und stetige Funktionen von x sind, dann heißt Φ stetig differenzierbar.

Beispiel 8.4.12. Sei

$$\Phi(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}) := 5\xi_1^2 + 9\xi_1\xi_2^4 + \xi_2 ,$$

dann ist

$$\begin{aligned} \frac{\partial}{\partial \xi_1}\Phi(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}) &= 10\xi_1^2 + 9\xi_2^4 , \\ \frac{\partial}{\partial \xi_2}\Phi(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}) &= 36\xi_1\xi_2^3 + 1 . \end{aligned}$$

Definition 8.4.13. Sei $\Phi : \mathbb{R}^n \to \mathbb{R}$ eine Funktion, $x \in \mathbb{R}^n$ so, dass alle partiellen Ableitungen von Φ im Punkt x existieren. Dann ist der Gradient von Φ an der Stelle x

$$\nabla\Phi(x) := \begin{pmatrix} \frac{\partial}{\partial x_1}\Phi(x) \\ \vdots \\ \frac{\partial}{\partial x_n}\Phi(x) \end{pmatrix} .$$

(Manche Autoren schreiben den Gradienten auch als Zeilenvektor.)

Satz 8.4.14 (Nicht präzise ausformuliert![27])**.** *Sei $\Phi : \mathbb{R}^n \to \mathbb{R}$ stetig differenzierbar, sei $y \in \mathbb{R}^n$ und $\Phi(y) = \gamma$. Sei $C = \{x \in \mathbb{R}^n \mid \Phi(x) = \gamma\}$.*

1. *Der Gradient $\nabla\Phi(y)$ zeigt in die Richtung, in die Φ am stärksten ansteigt.*
2. *Die Tangentialhyperebene T an die Fläche C im Punkt y steht orthogonal auf $\nabla\Phi(y)$, d.h.*

$$T = \{x \in \mathbb{R}^n \mid \langle \nabla\Phi(y), x - y\rangle = 0\} .$$

Beispiel 8.4.15. Sei C eine Ellipse in Hauptlage, mit Halbachsen $a = 4$, $b = 3$, d.h.

$$C = \{x \in \mathbb{R}^2 \mid 9\xi_1^2 + 16\xi_2^2 = 144\} .$$

Setzen wir $\Phi(x) := 9\xi_1^2 + 16\xi_2^2$, so ist $C = \{x \in \mathbb{R}^2 \mid \Phi(x) = 144\}$. Der Punkt $y = (3.2, 1.8)^T$ liegt, wie man leicht nachrechnet, auf C. Wir bestimmen die Tangente an C durch den Punkt y.

[27]Um diesen Satz scharf zu formulieren, müsste man erst den Begriff der Tangentialhyperebene und des Anstiegs definieren. Wir verlassen uns hier auf unsere Anschauung. Eine saubere Formulierung erhalten Sie in einer Analysis-Vorlesung.

Lösung: Wir bestimmen zuerst den Gradienten von Φ. Es ist

$$\nabla\Phi(x) = \begin{pmatrix} \frac{\partial}{\partial\xi_1}(9\xi_1^2+16\xi_2^2) \\ \frac{\partial}{\partial\xi_2}(9\xi_1^2+16\xi_2^2) \end{pmatrix} = \begin{pmatrix} 18\xi_1 \\ 32\xi_2 \end{pmatrix},$$

also

$$\nabla\Phi(y) = \nabla\Phi(\begin{pmatrix} 3.2 \\ 1.8 \end{pmatrix}) = \begin{pmatrix} 57.6 \\ 57.6 \end{pmatrix}.$$

Die Tangente an die Ellipse durch den Punkt y hat also die Gleichung

$$\langle \begin{pmatrix} 57.6 \\ 57.6 \end{pmatrix}, x\rangle = \langle \begin{pmatrix} 57.6 \\ 57.6 \end{pmatrix}, \begin{pmatrix} 3.2 \\ 1.8 \end{pmatrix}\rangle = 288\,.$$

Die Gleichung läßt sich bequem durch 57.6 kürzen und liefert

$$\xi_1 + \xi_2 = 5\,.$$

□

Satz 8.4.16. *Sei $A = A^T \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$. Die Tangentialhyperebene an die Quadrik $Q = \{x \in \mathbb{R}^n \mid x^TAx + 2b^Tx + \gamma = 0\}$ im Punkt $y \in Q$ ist*

$$\{x \in \mathbb{R}^n \mid y^TAx + b^Tx + b^Ty + \gamma = 0\}, \quad d.i.$$

$$\{x \in \mathbb{R}^n \mid \begin{pmatrix} y \\ 1 \end{pmatrix}^T \begin{pmatrix} A & b \\ b^T & \gamma \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = 0\}\,.$$

Beweis. Es sei $A = (\alpha_{ij})_{i,j=1\cdots n}$, $b = (\beta_i)_{i=1\cdots n}$. Wir setzen $\Phi(x) = x^TAx + 2b^Tx + \gamma$. Sei $k \in \{1, \cdots, n\}$. Wir differenzieren

$$\begin{aligned}
&\frac{\partial}{\partial\xi_k}\Phi(x) \\
= \;& \frac{\partial}{\partial\xi_k}\big[\sum_{i=1}^n\sum_{j=1}^n \alpha_{ij}\xi_i\xi_j + 2\sum_{i=1}^n \beta_i\xi_i + \gamma\big] \\
= \;& \frac{\partial}{\partial\xi_k}\big[\alpha_{kk}\xi_k^2 + \sum_{i\neq k}\alpha_{ik}\xi_i\xi_k + \sum_{j\neq k}\alpha_{kj}\xi_k\xi_i + \sum_{i\neq k}\sum_{j\neq k}\alpha_{ij}\xi_i\xi_j \\
& \quad +2\beta_k\xi_k + 2\sum_{i\neq k}\beta_i\xi_i + \gamma\big] \\
= \;& 2\alpha_{kk}\xi_k + 2\sum_{i\neq k}\alpha_{ik}\xi_i + 2\beta_2 \\
= \;& 2[\begin{pmatrix}\alpha_{k1} & \cdots & \alpha_{kn}\end{pmatrix} x + \beta_k]\,.
\end{aligned}$$

Wir setzen nun alle partiellen Ableitungen von Φ an der Stelle y zum Gradienten zusammen und erhalten

$$\nabla\Phi(y) = 2Ay + 2b\,.$$

Dann ist die Gleichung der Tangentialhyperebene im Punkt y

$$\begin{aligned}
0 \;&=\; \frac{1}{2}\langle\nabla\Phi(y), x-y\rangle = (y^TA + b^T)(x-y) = y^TAx + b^Tx - (y^TAy + b^Ty) \\
&=\; y^TAx + b^Tx - (y^TAy + 2b^Ty + \gamma) + b^Ty + \gamma = y^TAx + b^Tx + b^Ty + \gamma\,.
\end{aligned}$$

□

Beispiel 8.4.17. Wir bestimmen mittels Satz 8.4.16 die Tangente an die Ellipse $9\xi_1^2 + 16\xi_2^2 = 144$ im Punkt $y = (3.2, 1.8)^T$. Vergleiche die Lösung desselben Problems in Beispiel 8.4.15.

Lösung: Die Gleichung ist

$$\begin{aligned} 0 &= \begin{pmatrix}3.2\\1.8\end{pmatrix}^T \begin{pmatrix}9 & 0\\0 & 16\end{pmatrix} x - 144\,, \\ 0 &= 28.8\xi_1 + 28.8\xi_2 + 144\,. \end{aligned}$$

Kürzt man durch 28.8, erhält man $\xi_1 + \xi_2 = 5$. □

Definition 8.4.18. Sei $A = A^T \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$ und Q die Quadrik $Q = \{x \in \mathbb{R}^n \mid x^T Ax + 2b^T x + \gamma = 0\}$. Sei $y \in \mathbb{R}^n$ beliebig (nicht unbedingt auf Q).
Die Polare von y bezüglich der Quadrik Q ist die Hyperebene

$$\{x \in \mathbb{R}^n \mid y^T Ax + b^T x + (b^T y + \gamma) = 0\}\,, \text{ d.i.}$$

$$\{x \in \mathbb{R}^n \mid \begin{pmatrix}y\\1\end{pmatrix}^T \begin{pmatrix}A & b\\b^T & \gamma\end{pmatrix} \begin{pmatrix}x\\1\end{pmatrix} = 0\}\,.$$

(Liegt $y \in Q$, so erhält man also nach Satz 8.4.16 als Polare die Tangentialhyperebene.)

Satz 8.4.19. *Sei $Q \subset \mathbb{R}^n$ eine Quadrik und $y \in \mathbb{R}^n$, sei $z \in Q$. Dann sind äquivalent:*

(1) *Die Tangentialhyperebene an Q im Punkt z geht durch den Punkt y.*
(2) *Der Punkt z liegt auf der Polaren von y bezüglich Q.*

Beweis. Wir beschreiben Q durch die Gleichung

$$\begin{pmatrix}x\\1\end{pmatrix}^T \begin{pmatrix}A & b\\b^T & \gamma\end{pmatrix} \begin{pmatrix}x\\1\end{pmatrix} = 0\,.$$

Die Tangentialhyperebene T an Q im Punkt z hat dann nach Satz 8.4.16 die Gleichung

$$\begin{pmatrix}z\\1\end{pmatrix}^T \begin{pmatrix}A & b\\b^T & \gamma\end{pmatrix} \begin{pmatrix}x\\1\end{pmatrix} = 0\,.$$

Die Polare P von y bezüglich Q hat nach Definition 8.4.18 die Gleichung

$$\begin{pmatrix}y\\1\end{pmatrix}^T \begin{pmatrix}A & b\\b^T & \gamma\end{pmatrix} \begin{pmatrix}x\\1\end{pmatrix} = 0\,.$$

Also gilt

$$\begin{aligned} y \in T \quad &\iff \quad \begin{pmatrix}z\\1\end{pmatrix}^T \begin{pmatrix}A & b\\b^T & \gamma\end{pmatrix} \begin{pmatrix}y\\1\end{pmatrix} = 0 \\ &\iff \quad \begin{pmatrix}y\\1\end{pmatrix}^T \begin{pmatrix}A & b\\b^T & \gamma\end{pmatrix} \begin{pmatrix}z\\1\end{pmatrix} = 0 \iff z \in P\,. \end{aligned}$$

□

Bemerkung 8.4.20. Will man die Tangenten an eine Quadrik bestimmen, welche durch einen Punkt y gehen, schneidet man die Quadrik mit der Polaren von y. Durch jeden der Schnittpunkte geht eine Tangente. Es bildet sich ein Kegel von Tangenten, mit dem Scheitel y.

Beispiel 8.4.21. Bestimme die beiden Tangenten an die Ellipse

$$x^T \begin{pmatrix}1 & 0\\0 & 4\end{pmatrix} x + 2\begin{pmatrix}2\\4\end{pmatrix}^T x - 17 = 0$$

welche durch den Punkt $y = \begin{pmatrix}-2 & 2.125\end{pmatrix}^T$ gehen.

Lösung: Die Gleichung der Quadrik ist

$$\begin{pmatrix} x \\ 1 \end{pmatrix}^T \begin{pmatrix} 1 & 0 & 2 \\ 0 & 4 & 4 \\ 2 & 4 & -17 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = 0\,.$$

Die Polare von $\begin{pmatrix} -2 & 2.125 \end{pmatrix}^T$ ist

$$0 = \begin{pmatrix} -2 \\ 2.125 \\ 1 \end{pmatrix}^T \begin{pmatrix} 1 & 0 & 2 \\ 0 & 4 & 4 \\ 2 & 4 & -17 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 12.5 \\ -12.5 \end{pmatrix}^T \begin{pmatrix} x \\ 1 \end{pmatrix}$$

das ist die Gerade $\xi_2 = 1$[28]. In Parameterform ist diese Gerade

$$\begin{aligned} x &= \eta \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ \begin{pmatrix} x \\ 1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta \\ 1 \end{pmatrix}. \end{aligned}$$

Der Schnitt mit der Ellipse hat die Gleichung

$$\begin{aligned} 0 &= \begin{pmatrix} \eta \\ 1 \end{pmatrix}^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 4 & 4 \\ 2 & 4 & -17 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \eta \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \eta \\ 1 \end{pmatrix}^T \begin{pmatrix} 1 & 2 \\ 2 & -5 \end{pmatrix} \begin{pmatrix} \eta \\ 1 \end{pmatrix} \\ &= \eta^2 + 4\eta - 5\,. \end{aligned}$$

Wir erhalten $\eta = -2 \pm 3$, und damit sind die Schnittpunkte der Polaren mit der Ellipse

$$s_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad s_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} - 5 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -5 \\ 1 \end{pmatrix}.$$

Die Verbindungsgeraden von s_1 bzw. s_2 mit $\begin{pmatrix} -2 & 2.125 \end{pmatrix}^T$ sind die gesuchten Tangenten, also (in Parameterform)

$$x = \begin{pmatrix} -2 \\ 2.125 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ -1.125 \end{pmatrix} \quad \text{bzw. } x = \begin{pmatrix} -2 \\ 2.125 \end{pmatrix} + \lambda \begin{pmatrix} -3 \\ -1.125 \end{pmatrix}.$$

□

[28]Wir stellen uns nun etwas stur an, man könnte einfach $\xi_2 = 1$ in die Ellipsengleichung einsetzen und wäre gleich fertig. Aber wir wollen noch einmal zeigen, wie man eine Quadrik mit einem affinen Unterraum schneidet

Das Buch ist fertig geschrieben, Schreiber und Zeichner freuen sich auf das verdiente Festessen. Aber: “Pessime mus, saepius me provocas at iram. Ut te Deus perdat.” (Elendige Maus, immerfort provozierst du mich zur Wut! Dass dich Gott verdamme!)[29]

[29]Aus einer Handschrift, Böhmen ca. 1140. Entnommen aus: Der Künstler als Kunstwerk, Selbstporträts vom Mittelalter bis zur Gegenwart, U. Pfisterer und V. von Rosen, Hrsg., Reclam 2005.