Ergänzungsskriptum

zur Lehrveranstaltung

Lineare Algebra für PhysikerInnen

WS 2022/23

Franz Embacher

Fakultät für Mathematik | Fakultät für Physik der Universität Wien
E-mail: franz.embacher@univie.ac.at
WWW: http://homepage.univie.ac.at/franz.embacher/

Der Vorlesung wird hauptsächlich das Lehrbuch *Klaus Jänich: Lineare Algebra (Springer, 11. Auflage 2008)*, eBook der Universitätsbibliothek unter http://ubdata.univie.ac.at/AC06432777, zugrunde gelegt. Das vorliegende Ergänzungsskriptum ist eine geringfügig überarbeitete Version der Ergänzungsskripten früherer Semester. Es enthält Anmerkungen zum Buch, informiert über Ausdrucks- und Schreibweisen, die alternativ zu jenen des Buches verwendet werden können, und es enthält einige inhaltliche Erweiterungen, die zum Vorlesungsstoff gehören, aber nicht im Buch enthalten sind.

Die Tests und Übungen im Buch sind sehr zu empfehlen! Insbesondere die Tests (Multiple-Choice-Fragen) können zu einem schnellen Check, ob die Inhalte verstanden wurden, genutzt werden. Die richtigen Antworten sowie Hinweise, die Sie beachten sollten, wenn Ihre Antworten nicht stimmen, finden Sie am Ende des Buches. Wir werden in der Vorlesung einige Abschnitte des Buchs überspringen. Finden Sie selbst heraus, welche Testfragen sich auf die Abschnitte beziehen, die nicht übersprungen wurden!

Zu Beginn gleich drei Tipps: *Arbeiten Sie intensiv mit den Texten*, d.h. mit dem oben angegebenen Lehrbuch, mit dem Ergänzungsskriptum, das Sie gerade lesen und, wenn Sie das Bedürfnis danach haben, mit der auf meiner Homepage unter https://homepage.univie.ac.at/franz.embacher/Lehre/Lineare_Algebra_fuer_PhysikerInnen/LfP_ws2022.html angegebenen Alternativliteratur. Es wird sehr sinnvoll zu sein, die Textstellen, die sich auf das Thema, das gerade behandelt wird, *im Voraus und im Nachhinein* zu lesen. Was bedeutet „lesen" in diesem Zusammenhang? Versuchen Sie, zu viel Information wie möglich aus den Texten zu beziehen, denken Sie stets mit, machen Sie sich Notizen, vollziehen Sie das eine oder andere im Detail auf dem Papier nach. Sie dürfen die Worte und Formeln durchaus auf die Goldwaage legen: Lesen Sie genau, lesen Sie nicht oberflächlich, nehmen Sie sich genug Zeit dafür! Tipp Nummer zwei: Nehmen Sie die Übungsaufgaben ernst! Mit ihrer Hilfe können Sie nicht nur den ÜbungsleiterInnen, sondern (wichtiger!) auch sich selbst beweisen, dass Sie die Inhalte verstanden haben und auf konkrete Fälle anwenden können. Tipp Nummer drei: Wenn Sie die Zeit dazu haben, sehen Sie sich durchaus auch im Buch und im Ergänzungsskriptum die Inhalte des jeweils *nächsten* Vorlesungstermins an! Dann fällt das Mitkommen leichter, und Sie können die eine oder andere Frage gleich in die Vorlesung mitbringen.

1 Mengen und Abbildungen

1.1 Mengen

- **Seite 2**: Das im Buch verwendete Zeichen für die leere Menge ist $\emptyset$. Aufgund der Verwechslungsgefahr mit ϕ und Φ (vor allem in handgeschriebenen Texten) bevorzuge ich persönlich stattdessen das Symbol $\{\}$.

- **Seiten 2 und 4**: Neben dem im Buch verwendenen Symbol $\subset$ für „Teilmenge" ist auch das Symbol $\subseteq$ gebräuchlich. Um eigens auszudrücken, dass A eine *echte* Teilmenge von B ist (also $A \subseteq B$ und $A \neq B$), kann $A \subsetneqq B$ geschrieben werden.

- **Seite 4**: Die hier besprochenen Skizzen zur Veranschaulichung von Mengenoperationen heißen *Venn-Diagramme*.

- **Seite 5**: Statt Durchschnitt, Vereinigung und Differenz kann man auch *Durchschnittsmenge*, *Vereinigungsmenge* und *Differenzmenge* sagen. Ist $B \subseteq A$, so nennt man $A \backslash B$ auch die *Komplementärmenge* von B bezüglich A. Ein wichtiges Beispiel ist $\mathbb{R} \backslash \{0\}$, die Menge aller von 0 verschiedenen reellen Zahlen. Sie wird oft auch mit dem Kürzel $\mathbb{R}^*$ bezeichnet.

- **Seite 5**: Die Elemente a und b, aus denen ein (geordnetes) Paar (a, b) besteht, werden die *Komponenten* oder *Koordinaten* (oder schlicht *Eintragungen*) des Paares genannt. Aber Achtung: a und b sind nicht „die Elemente des Paares", da das Paar (a, b) nicht das Gleiche ist wie die Menge $\{a, b\}$! Analoges gilt für Tripel und n-tupel (die man übrigens auch n-Tupel schreiben kann).

- **Seite 6** (oben): Der Doppelpunkt kann auch zur Definition von Schreibweisen verwendet werden. Etwa so, für Mengen A und B:

  $$A \subsetneqq B \;\; :\Leftrightarrow \;\; A \subseteq B \text{ und } A \neq B. \tag{1}$$

  Damit wird die Schreibweise auf der linken Seite erklärt.

- **Seite 7**: Zur Schreibweise von n-Tupeln möchte ich gleich zu Beginn klarstellen: Sie sind völlig frei, ein Zahlentripel in der Form

  $$(a, b, c) \qquad \text{oder} \qquad (a \;\; b \;\; c) \qquad \text{oder} \qquad \begin{pmatrix} a \\ b \\ c \end{pmatrix} \tag{2}$$

  anzuschreiben, um ein Element von $\mathbb{R}^3$ zu spezifizieren. Wenn es Ihnen Spaß macht, können Sie auch ein Element von $\mathbb{R}^{15}$ angeben, indem Sie die fünfzehn Komponenten herzförmig anordnen – wichtig ist nur, dass von jeder Komponente klar ist, welche Platznummer sie im Tupel hat. So gesehen ist ein Tupel einfach eine *Liste*, unabhägig davon, wie man sie anschreibt. In der Physik wird oft die Spaltenform bevorzugt.

  Aber Achtung: Wir werden eine *Rechentechnik* kennenlernen, die der übersichtlichen Beschreibung und Manipulation linearer Abbildungen dient: die Matrizenrechnung. Wann immer wir Matrizenrechnung betreiben, *müssen* die dabei vorkommenden Elemente des

$\mathbb{R}^n$ in der vorgesehenen Weise als Spaltenvektor oder als Zeilenvektor angeschrieben werden. Das ist aus rein praktischen Gründen nötig, denn sonst würde die Matrizenrechnung nicht funktionieren.

Ein „Vektor aus $\mathbb{R}^3$" ist also weder „ein Spaltenvektor" noch „ein Zeilenvektor", sondern ein Zahlentripel, bestehend aus einer ersten, einer zweiten und einer dritten Zahl. Daran ändert sich nichts, wenn im Rahmen der Matrizenrechnung die Komponenten des Tripels meistens untereinander, manchmal auch nebeneinander geschrieben werden. Analoges gilt auch für die n-Tupel aus komplexen Zahlen, die den $\mathbb{C}^n$ bilden und später dazukommen werden.

1.2 Abbildungen

- **Seite 8**: Statt *Abbildung* kann ganz allgemein auch *Funktion* gesagt werden.
- **Seite 9**: Id_M heißt auch die *identische Abbildung* auf M.
- **Seite 11** (unterhalb des ersten Kastens): Hier wird gesagt, dass die Sprechweise, f sei eine Abbildung *auf* Y nur für surjektive Abbildungen verwendet wird. In der Praxis ist (spricht, schreibt) man hier etwas salopper. Sagt jemand, f sei eine „Abbildung auf Z", so kann damit auch gemeint sein, dass f *auf* Z definiert ist, dass also f eine Abbildung von Z in eine andere (oder in dieselbe) Menge ist. An dem zweimaligen „in" im vorigen Satz erkennt man auch, dass eine Abbildung *nach* einer Menge auch als Abbildung *in* diese Menge bezeichnet werden kann. Interpretieren Sie also in diese kleinen Wörtchen *nach*/*auf*/*in* nicht zu viel hinein! Wenn Sie eigens ausdrücken wollen, dass f surjektiv ist, sagen (oder schreiben) Sie es am besten dazu!
- **Seite 11** (Definition unten): Die Kurzschreibweise gf anstelle der genaueren Variante $g \circ f$ für eine *Verkettung* (also für das *Hintereinanderausführen*) zweier Abbildungen (und dementsprechend f^2 für $f \circ f$) ist in der linearen Algebra für *lineare* Abbildungen (Operatoren) üblich (was das genau ist, kommt später), aber ansonsten mit Vorsicht zu genießen, da eine Verwechslungsgefahr mit dem (etwa in der Analysis oft auftretenden) *punktweisen Produkt* von Funktionen $g \cdot f : x \mapsto g(x)f(x)$ besteht!
- **Seite 14**: Die Einschränkung von f auf die Menge A wird im Buch mit $f|A$ bezeichnet. Eine andere, häufig anzutreffende Schreibweise dafür ist $f|_A$.

2 Vektorräume

2.1 Reelle Vektorräume

- **Seite 22** (Definition): Die im Buch verwendeten Ausdrücke *skalare Multiplikation* und *Skalarmultiplikation* dürfen nicht mit dem später auftretenden *Skalarprodukt* verwechselt werden! Eine eindeutigere Bezeichnung wäre *Multiplikation mit einem Skalar* oder schlicht das *Bilden von* (reellen, später auch komplexen) *Vielfachen*.

- **Seite 23**: Axiom (1) heißt *Assoziativgesetz* (der Addition), Axiom (2) heißt *Kommutativgesetz* (der Addition). (Hinweis am Rande: Die Bezeichnung *Nullvektor*, siehe Axiom (3), bekommt in der Relativitätstheorie eine andere Bedeutung.)

  Nun, nachdem ein zentraler Begriff der linearen Algebra, der des reellen Vektorraums, offiziell definiert ist, versuchen Sie zu resümieren, was damit eigentlich gesagt ist! Was kann man mit (oder „in") einem reellen Vektorraum machen? Summen und (reelle) Vielfache bilden! Womit? Mit den Elementen der Menge V, die wir kurz auch als *Vektoren* bezeichnen. Und zwar auf der Basis gewisser Regeln, die in den Axiomen (1) – (8) festgelegt sind. Sind beispielsweise u und v Elemente eines reellen Vektorraums, so macht es Sinn, $2u + 3v$ zu bilden. Eine derartige „Kombination" von Elementen eines Vektorraums heißt *Linearkombination* – ein Begriff, der etwas später auch im Buch auftritt. Damit können wir pointiert formulieren: Der Sinn eines Vektorraums besteht darin, uns das Bilden von Linearkombinationen auf einer sicheren Basis zu erlauben. Mehr ist zunächst nicht vorgesehen. Man spricht auch von einer *linearen Struktur*, die dadurch festgelegt wird, und bisweilen findet man statt des Ausdrucks *Vektorraum* die (etwas ältere) Bezeichnung *linearer Raum*. Um diese linearen Strukturen wird es im Folgenden gehen.

  Einer der einfachsten reellen Vektorräume ist der $\mathbb{R}^2$. Um sich ein „Bild" von ihm zu machen, zieht man die mit zwei aufeinander normal stehenden Achsen ausgestattete „Zeichenebene", auch „Anschauungsebene" genannt, entweder in Gedanken oder in Form eines Blatts Papier heran. Das Zahlenpaar $(x_1, x_2) \in \mathbb{R}^2$ kann man sich dann entweder als Punkt mit den Koordinaten x_1 und x_2 oder als Pfeil vom Ursprung $(0, 0)$ zu diesem Punkt (d.h. als *Ortsvektor*) vorstellen und zeichnen. Sowohl die Punkt-Deutung als auch die Pfeil-Deutung reeller Zahlenpaare ist nützlich, wie Abbildung 1 illustriert.

  Mit dem $\mathbb{R}^3$, dem Modell des mit drei aufeinander normal stehenden Achsen ausgestatteten „Anschauungsraumes" funktioniert das analog, nur eben mehr in Gedanken als auf dem Papier. Diese Deutungen und Visualisierungen von Zahlenpaaren und Zahlentripeln helfen auch, sich unter den Sachverhalten, denen wir in Vektorräumen ganz allgemein begegnen werden, etwas einigermaßen Konkretes vorzustellen. Dabei sollte man aber – solange es um Vektorräume (ohne zusätzliche Strukturen) geht – versuchen, von Eigenschaften abzusehen, von denen in den Vektorraum-Axiomen nicht die Rede ist (wie etwa dem Abstand zweier Punkte, der Länge eines Pfeils oder dem Winkel zwischen zwei Pfeilen).

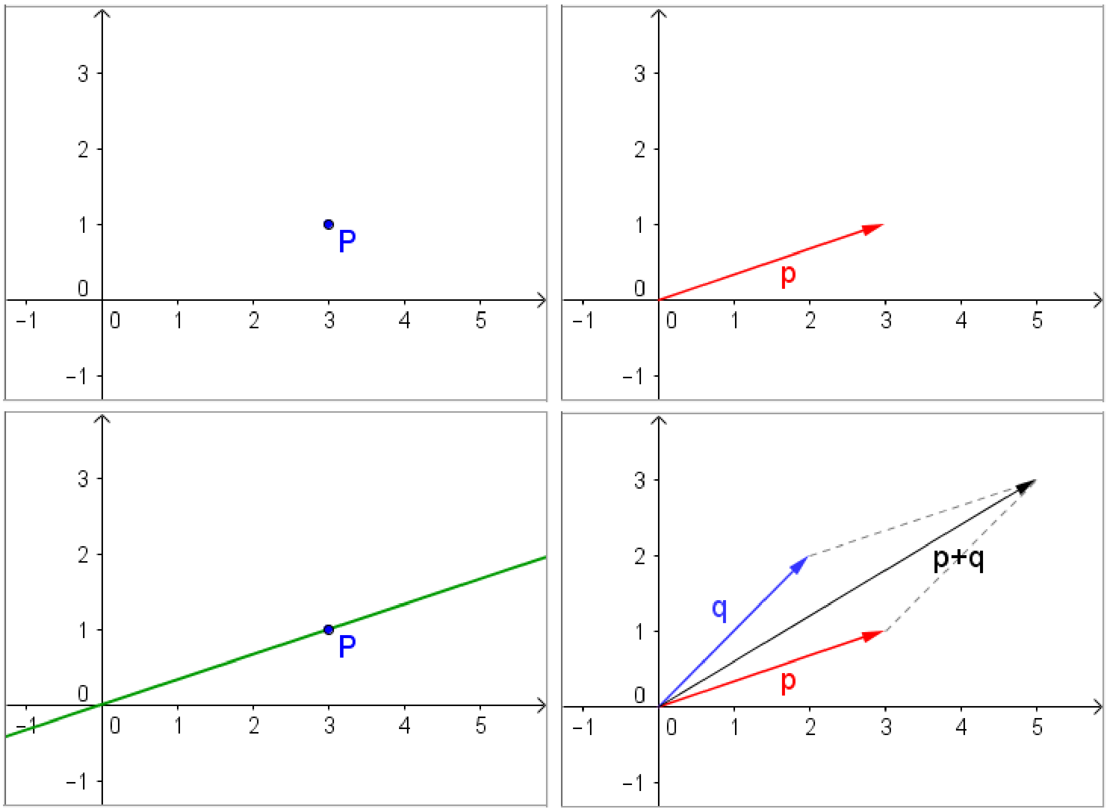

Abbildung 1: Die Deutung reeller Zahlenpaare als Punkte – links oben ist das Zahlenpaar $(3, 1)$ als Punkt gedeutet – ist nützlich etwa zur Charakterisierung von Teilmengen des $\mathbb{R}^2$. So stellt beispielsweise die grüne Gerade (links unten) die Menge aller Vielfachen von $(3, 1)$ dar. Die Deutung reeller Zahlenpaare als Pfeile – rechts oben ist das Zahlenpaar $(3, 1)$ als Pfeil gedeutet – ist nützlich zur Visualisierung der Vektorraum-Operationen. Als Beispiel ist rechts unten die Visualisierung der Addition $(3, 1)+(2, 2) = (5, 3)$ dargestellt.

2.2 Komplexe Zahlen und komplexe Vektorräume

- **Seite 28**: Die Ebene, die hier skizziert ist, also der $\mathbb{R}^2$, der ja als Menge mit $\mathbb{C}$ übereinstimmt, heißt *Gaußsche Zahlenebene* (oder einfach *komplexe Ebene*).

- **Seite 30**: Ist V ein Vektorraum über $\mathbb{K}$ (oder kurz: ein $\mathbb{K}$-Vektorraum), so wird $\mathbb{K}$ auch als *Grundkörper* bezeichnet. Das Wort *Skalar* bezeichnet in der linearen Algebra ein Element des Grundkörpers. Da wir uns auf die Grundkörper $\mathbb{R}$ und $\mathbb{C}$ beschränken, ist ein Skalar im Buch und in diesem Skriptum – je nachdem – eine reelle oder eine komplexe Zahl.

2.3 Untervektorräume

- **Seite 31**: Neben dem Wort *Untervektorraum* ist auch die etwas weniger genaue Bezeichnung *Teilraum* (oder *Unterraum*) üblich. Sie wird allerdings auch in anderen Zusammenhängen verwendet. Ihre genaue Bedeutung hängt davon ab, von welcher Struktur gerade die Rede ist. So ist ein Teilraum eines Vektorraums V eine Teilmenge von V, die (hinsichtlich der gleichen Operationen „Addition" und „Multiplikation mit einem Skalar", wie sie in V definiert sind) selbst ein Vektorraum ist. In analoger Weise haben auch andere mathematische Strukturen (wie euklidische Vektorräume, unitäre Vektorräume, affine Räume, Hilberträume, Banachräume, metrische Räume, topologische Räume,...) ihre „Teilräume", deren Strukturen sie „erben".

2.4 Test

- Arbeiten Sie den Test durch!

2.5 Körper

- Obwohl dies „ein Abschnitt für Mathematiker" ist, lesen Sie ihn! Für den Fall, dass Sie Lust auf eine kleine Herausforderung haben: Wenn Sie versuchen wollen, sich einen wirklich komplizierten Vektorraum vorzustellen, dann nehmen Sie diesen: $\mathbb{R}$ als Vektorraum über $\mathbb{Q}$!

2.6 Was sind Vektoren?

- **Seiten 38 – 41**: Hier, in diesem „Abschnitt für Physiker", wird zunächst eine zusätzliche Struktur angerissen, die später im Buch (Kapitel 8) genauer besprochen wird. Wir werden das in der Vorlesung genauso machen, vor allem um herauszustreichen, dass ein Vektorraum ohne zusätzliche Struktur „nicht weiß", was ein Skalarprodukt ist und daher auch das Konzept der Länge eines Vektors „nicht kennt".

- **Seite 41**: In der Definition am Beginn der Seite tritt zu ersten Mal im Buch das Wort *linear* in seiner „technischen" Bedeutung für Abbildungen auf, die Ihnen vielleicht nicht vertraut ist. Keine Sorge – er wird später präzisiert werden! Hier nur so viel: Eine Abbildung ist *linear*, wenn es keinen Unterschied macht, ob man zuerst Summen und Vielfache bildet und danach die Abbildung anwendet oder ob man zuerst die Abbildung anwendet und mit ihren Bildern die entsprechenden Summen und Vielfache bildet. Man sagt auch: Eine lineare Abbildung „respektiert" (oder „erhält") die lineare Struktur (d.h. die Vektorraum-Struktur). Hinsichtlich der Bezeichnung „linear" besteht leider eine gewisse Uneinheitlichkeit: Während Sie in Ihrem Mathematikunterricht wahrscheinlich gelernt haben, dass eine „lineare Funktion" $f : \mathbb{R} \to \mathbb{R}$ von der Form $f(x) = k\,x + d$ ist (mit $k, d \in \mathbb{R}$), gilt eine solche Abbildung für die lineare Algebra nur dann als „linear", wenn $d = 0$ ist. Denn nur dann gilt

$$f(x + y) = f(x) + f(y) \quad \text{für alle } x, y \in \mathbb{R} \tag{3}$$

(d.h. f respektiert das Bilden von Summen) und

$$f(\lambda x) = \lambda f(x) \quad \text{für alle } \lambda, x \in \mathbb{R} \tag{4}$$

(d.h. f respektiert das Bilden von reellen Vielfachen). Dabei spielt $\mathbb{R}$ eine Doppelrolle: als Grundkörper und als Vektorraum. Hier haben wir also den einfachsten nichttrivialen[1]

[1] Der einfachste, wenngleich triviale, reelle Vektorraum ist $\{0\}$ über $\mathbb{R}$.

reellen Vektorraum, den es gibt: $\mathbb{R}$ (oder $\mathbb{R}^1$, wenn Ihnen das lieber ist) als Vektorraum über $\mathbb{R}$!

- **Seiten 42 – 50**: Diese Seiten mögen zu Beginn schwer verdaulich zu sein. In der Vorlesung werden wir diesen Teil des Abschnitts 2.6 überspringen, da er nicht zum Stoff der Vorlesung, soweit Sie ihn in den Übungen anwenden und bei der Prüfung beherrschen sollen, gehört. Aber er betrifft ein (gerade für PhysikerInnen) wichtiges Thema, das Sie nur in wenigen anderen Lehrbüchern finden werden. Lesen Sie ihn daher trotzdem! Kehren Sie vielleicht auch später, gegen Ende der Vorlesung oder danach, zu diesem Abschnitt zurück!

2.7 Komplexe Zahlen vor 400 Jahren

- Keine Anmerkungen zu diesem (sehr lesenswerten) Abschnitt.

---

3 Dimensionen

3.1 Lineare Unabhängigkeit

- **Seite 56**: Die *lineare Hülle* wird auch *Erzeugnis* genannt. Neben dem im Buch verwendeten Symbol L sind auch die Abkürzungen span und $\langle \dots \rangle$ üblich. Im Buch wird die lineare Hülle von Tupeln, also von endlich vielen Vektoren gebildet. Der Begriff der linearen Hülle lässt sich auf offensichtliche Weise verallgemeinern, so dass er auf *beliebige Teilmengen* von V angewandt werden kann. Dabei ist aber darauf zu achten, dass eine Linearkombination immer nur mit *endlich* vielen Vektoren gebildet werden kann.

- **Seite 58** (Definition): Eine *Basis* ist im Buch ein n-Tupel von Vektoren. Nach einer anderen Konvention ist eine Basis eine *Menge* von Vektoren. Die im Buch verwendete Konvention hat Vorteile – etwa den, dass die Vertauschung zweier Basisvektoren eine klare Bedeutung hat: Dadurch entsteht eine *andere* Basis. Die Vektoren, die eine Basis bilden, werden mitunter auch als „Basiselemente“ bezeichnet, selbst wenn eine Basis ein n-Tupel und keine Menge ist. Insofern ist die Bezeichnung „Basisvektoren“ korrekter.

- **Seite 58** (Bemerkung 2): Die Zahlen $\lambda_1, \dots, \lambda_n$ heißen *Entwicklungskoeffizienten* (kurz *Koeffizienten*) oder *Komponenten* von v bezüglich der Basis $(v_1, \dots, v_n)$. Die „Entwicklung eines Vektors in eine Basis“ wird oft auch mit dem Summensymbol angeschrieben:

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n \equiv \sum_{j=1}^{n} \lambda_j v_j \, . \tag{5}$$

  Wichtig ist, dass bei festgehaltener Basis die Entwicklungskoeffizienten eines gegebenen Vektors eindeutig bestimmt sind. Umgekehrt können Entwicklungskoeffizienten $\lambda_1, \dots, \lambda_n$ beliebig vorgegeben werden und legen gemäß (5) einen eindeutig bestimmten Vektor v fest.

- **Seite 59**: Die *kanonische Basis* $(e_1, \ldots, e_n)$ des $\mathbb{K}^n$ wird auch als *Standardbasis* bezeichnet.

3.2 Der Dimensionsbegriff

- **Seite 61** (Definition): Die Dimension eines Vektorraums V wird im Buch mit $\dim V$ bezeichnet, und es wird beispielsweise gesagt, dass $\dim \mathbb{K}^n = n$ gilt. Mit gefällt die Schreibweise mit verpflichtender Klammer, also $\dim(V)$ und $\dim(\mathbb{K}^n) = n$, besser. Analoges gilt auch für später auftretende Bezeichnungen wie Bild, Kern, Rang (rg) und Determinante $(\det)$.
- **Seite 62** (Definition): Die lineare Algebra beschäftigt sich (abgesehen von wenigen ganz allgemein geltenden Sachverhalten) mit endlichdimensionalen Vektorräumen, während sich das mathematische Teilgebiet der *Funktionalanalysis* der Untersuchung unendlichdimensionaler Vektorräume widmet.

3.3 Test

- Arbeiten Sie den Test durch!

3.4 Beweis des Basisergänzungssatzes und des Austauschlemmas

- Diesen „Abschnitt für Mathematiker“ werden wir nur kurz streifen, aber nicht im Detail nachvollziehen – eine Maßnahme, die dazu dienen soll, jenen TeilnehmerInnen, die mit längeren formalen Begründungen ihre Probleme haben, nicht gleich zu Beginn der Vorlesung die Motivation zu nehmen. Das bedeutet also: In der Vorlesung werden *nicht ausnahmslos alle* allgemeinen Aussagen, die wir formulieren, streng bewiesen. Es bedeutet *nicht*, dass Beweise für das Verständnis keinen Wert hätten! Aber in diesem konkreten Fall verhält es sich so, dass Sie die Beweise leichter verstehen, wenn Sie sich erst einmal ein paar Monate an die lineare Algebra gewöhnt haben und diesen Abschnitt bei Interesse später lesen. Als Ersatz versuchen Sie, sich vorzustellen, was diese beiden Sätze für den $\mathbb{R}^3$ besagen!

3.5 Das Vektorprodukt

- In diesem „Abschnitt für Physiker“ werden einige nicht ganz einfache Begriffe und Bezeichnungen verwendet, die im Buch zuvor im Abschnitt 2.6 eingeführt wurden. Wir werden stattdessen in der Vorlesung als „Light-Version“ das *Vektorprodukt* (auch *vektorielles Produkt* oder *Kreuzprodukt* genannt) im $\mathbb{R}^3$ besprechen, basierend auf der Formel auf Seite 74 unten und im Hinblick auf die auch im Buch besprochene geometrische Charakterisierung und die Formel (9) auf Seite 75. Versuchen Sie aber trotzdem, diesen Abschnitt im Buch zu lesen!
- Ergänzung: **Das „Epsilon-Symbol“**
  Im Zusammenhang mit dem Vektorprodukt ist ein numerisches Objekt nützlich, das

als *Epsilon-Symbol* oder auch als „*Epsilon-Tensor*" bezeichnet wird. Für alle Werte $j, k, l \in \{1, 2, 3\}$ definieren wir ε_{jkl} folgendermaßen:

- $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$ (also wann immer die Indizes in „zyklischer" Reihenfolge $1 \to 2 \to 3 \to 1$ stehen).
- $\varepsilon_{321} = \varepsilon_{213} = \varepsilon_{132} = -1$ (also wann immer die Indizes in „antizyklischer" Reihenfolge $3 \to 2 \to 1 \to 3$ stehen).
- $\varepsilon_{jkl} = 0$ in allen anderen Fällen (also wann immer mindestens zwei Indizes gleich sind).

(Hinter den ersten beiden Festlegungen steckt eine Einteilung der *Permutationen* der Zahlen 1 bis 3 in „gerade" und „ungerade", doch mehr davon im Kapitel 6.) Das Epsilon-Symbol ist *total antisymmetrisch*, d.h. bei Vertauschung zweier Indizes wechselt es das Vorzeichen (etwa für die Vertauschung des zweiten mit dem dritten Index gilt $\varepsilon_{jlk} = -\varepsilon_{jkl}$), und es ändert seinen Wert nicht, wenn die Indizes zyklisch verschoben werden: $\varepsilon_{jkl} = \varepsilon_{klj} = \varepsilon_{ljk}$. Mit Hilfe dieses Objekts können nun die Komponenten des Vektorprodukts $u = x \times y$ in der (insbesondere in der Physik üblichen) Form

$$u_j = \sum_{k,l=1}^{3} \varepsilon_{jkl}\, x_k\, y_l \tag{6}$$

angeschrieben und berechnet werden. Probieren wir es mit der ersten Komponente aus:

$$u_1 = \sum_{k,l=1}^{3} \varepsilon_{1kl}\, x_k\, y_l = \underbrace{\varepsilon_{123}}_{1}\, x_2\, y_3 + \underbrace{\varepsilon_{132}}_{-1}\, x_3\, y_2 = x_2\, y_3 - x_3\, y_2\,. \tag{7}$$

In der Summe über k und l war der Wert 1 nicht mehr zu berücksichtigen, da das Epsilon-Symbol mit zwei gleichen Indizes ohnehin verschwindet, und Gleiches gilt für jene Summanden, für die $k = l$ ist. Somit bleiben nur mehr zwei Summanden übrig: jener mit $(k, l) = (2, 3)$ und jener mit $(k, l) = (3, 2)$. Mit (6) kann auch leicht überprüft werden, dass u (bezüglich des Standard-Skalarprodukts im $\mathbb{R}^3$) zu x orthogonal ist:

$$\langle x, u\rangle = \sum_{j=1}^{3} x_j\, u_j = \sum_{j,k,l=1}^{3} \underbrace{\varepsilon_{jkl}}_{\text{in } jk \text{ antisymmetrisch}} \; \Big| \; \underbrace{x_j\, x_k}_{\text{in } jk \text{ symmetrisch}} \; y_l = 0\,. \tag{8}$$

Man kann es auch so sehen: Mit der Ersetzung $\varepsilon_{jkl} = -\varepsilon_{kjl}$ wird diese Summe gleich minus sie selbst, ist daher gleich 0.

Das Konzept des Vektorprodukts kann auf beliebige dreidimensionale euklidische Vektorräume ausgedehnt werden, wobei allerdings zuvor eine zusätzliche Struktur benötigt wird: Man muss eine (beliebige) Basis für *rechtshändig* erklären (wodurch sich ein *orientierter euklidischer Vektorraum* ergibt). Danach ist jede Basis entweder rechtshändig oder linkshändig. (Genau genommen muss man das im $\mathbb{R}^3$ auch, aber hier nehmen wir aufgrund der üblichen bildlichen Darstellungsweise mehr oder weniger automatisch an, dass die Standardbasis rechtshändig ist. Machen wir es also ganz offiziell: Hiermit legen wir fest, dass die Standardbasis des $\mathbb{R}^3$ rechtshändig ist.)

- Ergänzung: **Die Einsteinsche Summenkonvention**
  An dieser Stelle ist es angebracht, eine Konvention vorzustellen, die in der Physik oft hilft, Rechnungen übersichtlich zu halten. Sie lautet: Kommt in einem Produkt von indizierten Größen (also Größen, die einen Index oder mehrere Indizes tragen) ein Index *zweimal* vor, so ist über diesen Index die Summe zu bilden, ohne dass dafür eigens ein Summensymbol angeschrieben wird. Wird diese Konvention vereinbart, so sehen die Ausdrücke für ein Skalarprodukt $\langle x, y\rangle$ und für die Komponenten des Vektorprodukts $u = x \times y$ so aus:

$$\begin{aligned} \langle x, y\rangle &= x_j\, y_j &(9)\\ u_j &= \varepsilon_{jkl}\, x_k\, y_l\,. &(10) \end{aligned}$$

  Neben dem Epsilon-Symbol wird häufig das *Kronecker-Symbol* („*Kronecker-Delta*") benutzt, das im Buch erst später (auf Seite 92) auftritt. Für alle Werte $j, k \in \{1, 2, 3\}$ definieren wir δ_{jk} folgendermaßen:

$$\delta_{jk} = \begin{cases} 1 & \text{wenn } \; j = k \\ 0 & \text{wenn } \; j \neq k. \end{cases} \tag{11}$$

  Mit der Summenkonvention und den zusätzlichen Beziehungen

$$\varepsilon_{jkl}\, \varepsilon_{jmn} = \delta_{km}\, \delta_{ln} - \delta_{kn}\, \delta_{lm} \quad \text{und} \quad \varepsilon_{jkl}\, \varepsilon_{jkn} = 2\, \delta_{ln} \quad \text{sowie} \quad \delta_{jj} = 3 \tag{12}$$

  können in relativ übersichtlicher Weise vereinfachte Ausdrücke für Größen, die aus Skalar- und Vektorprodukten aufgebaut sind, erhalten werden. Beispiel:

$$\begin{aligned} \|x \times y\|^2 &= \langle x \times y, x \times y\rangle = \varepsilon_{jkl}\, x_k\, y_l\, \varepsilon_{jmn}\, x_m\, y_n = \\ &= (\delta_{km}\, \delta_{ln} - \delta_{kn}\, \delta_{lm})\; x_k\, y_l\, x_m\, y_n = (x_k\, x_k)\,(y_l\, y_l) - (x_k\, y_k)\,(x_l\, y_l) = \\ &= \|x\|^2\, \|y\|^2 - \langle x, y\rangle^2\,. \end{aligned} \tag{13}$$

  In der Relativitätstheorie werden Sie auf eine Variante der Summenkonvention stoßen, bei der immer einer der doppelt vorkommenden Indizes „oben" und einer „unten" stehen muss (Beispiel: $u_\mu\, v^\mu$ steht für $\sum_{\mu=0}^{3} u_\mu\, v^\mu$), aber warum das so ist, geht über den Stoff einer Einführung in die lineare Algebra hinaus.

3.6 Der „Steinitzsche Austauschsatz"

- Diesen Abschnitt werden wir in der Vorlesung überspringen, er ist aber für historisch Interessierte lesenswert.

---

4 Lineare Abbildungen

4.1 Lineare Abbildungen

- **Seiten 81/82**: Eine lineare Abbildung wird auch als *linearer Operator* (oder, wenn die Linearität aus dem Zusammenhang klar ist, kurz als *Operator*) bezeichnet. Um eigens auszudrücken, dass sich die Linearität auf den Grundkörper $\mathbb{K}$ bezieht, kann auch der Terminus $\mathbb{K}$-*linear* verwendet werden. Mit dem im Buch verwendeten Begriff *Homomorphismus* verhält es sich ähnlich wie mit dem Terminus *Teilraum*: Ein Homomorphismus ist ganz allgemein eine Abbildung, die eine bestimmte Struktur respektiert. Eine lineare Abbildung könnte daher genauer auch als *Vektorraum-Homomorphismus* bezeichnet werden.

  Die beiden Forderungen $f(x+y) = f(x) + f(y)$ und $f(\lambda x) = \lambda f(x)$ für alle $x, y \in V$ und $\lambda \in \mathbb{K}$ können auch in der Form

  $$f(x + \lambda y) = f(x) + \lambda f(y) \qquad \text{für alle } x, y \in V \text{ und } \lambda \in \mathbb{K} \tag{14}$$

  zusammengezogen werden. Das überträgt sich auf beliebige Linearkombinationen: Haben wir etwa r Skalare $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ und r Vektoren $x_1, \ldots, x_r \in V$, so gilt

  $$f\left(\sum_{j=1}^{r} \lambda_j x_j\right) = \sum_{j=1}^{r} \lambda_j f(x_j). \tag{15}$$

  Wir sagen auch: „f vertauscht mit dem Bilden von Linearkombinationen“ oder „wir können (die Wirkung von) f in die Summe hineinziehen“. Machen Sie sich bitte vollends klar, was die Schreibweise mit dem Summensymbol in (15) besagt!

  Für das Bild und den Kern einer linearen Abbildung sind auch andere Bezeichnungen als die im Buch verwendeten üblich, vor allem die englischen Varianten Im (*image*) für Bild und Ker (*kernel*) für Kern. Statt $\mathrm{Hom}(V, W)$ wird oft auch $L(V, W)$ geschrieben.

  Die Beweise, dass $\mathrm{Bild}(f)$ und $\mathrm{Kern}(f)$ Untervektorräume sind, werden im Buch nicht vorgeführt und seien hiermit nachgetragen:

  - Zunächst gilt für jede lineare Abbildung f, dass $f(0) = 0$ ist. Beweis: $f(0+0) = f(0) + f(0)$, daher $f(0) = f(0) + f(0)$ und daher $f(0) = 0$. (Die erste 0 ist der Nullvektor von V, die zweite 0 ist der Nullvektor von W.) Daher sind $\mathrm{Bild}(f)$ und $\mathrm{Kern}(f)$ nicht leer: $\mathrm{Bild}(f)$ enthält den Nullvektor von W, $\mathrm{Kern}(f)$ enthält den Nullvektor von V. Damit ist das erste Kriterium für Untervektorräume ($\neq \{\}$) für das Bild und den Kern von f erfüllt.
  - Sind $x, y \in \mathrm{Bild}(f)$, so gibt es Vektoren $u, v \in V$ mit $f(u) = x$ und $f(v) = y$. Wegen $f(u+v) = f(u)+f(v) = x+y$ ist auch $x+y$ Element von $\mathrm{Bild}(f)$. Wegen $f(\lambda u) = \lambda f(u) = \lambda x$ für beliebige $\lambda \in \mathbb{K}$ ist auch λx Element von $\mathrm{Bild}(f)$. Daher ist $\mathrm{Bild}(f)$ ein Untervektorraum von W.
  - Sind $u, v \in \mathrm{Kern}(f)$, so gilt $f(u+v) = f(u) + f(v) = 0 + 0 = 0$ und $f(\lambda u) = \lambda f(u) = \lambda \cdot 0 = 0$, und daher sind auch $u + v$ und λu Elemente von $\mathrm{Kern}(f)$. Folglich ist $\mathrm{Kern}(f)$ ein Untervektorraum von V.

- **Seite 82** (Bemerkung 1): Zur Schreibweise von Isomorphismen: Im Buch wird des Öfteren von einem „Isomorphismus $V \cong W$" oder von einem „Isomorphismus $f : V \cong W$" gesprochen. Damit ist stets ein Isomorphismus $V \to W$ bzw. $f : V \to W$ gemeint. Das kann man auch in der Form $V \xrightarrow{\cong} W$ bzw. $f : V \xrightarrow{\cong} W$ anschreiben, wodurch der Zusatz, dass es sich um einen Isomorphismus handelt, wegfallen kann.

- **Seite 84** (Bemerkung 2): Albrecht Beutelspacher nennt die w_i in seinem Buch *Lineare Algebra* „Wunschvektoren" – eine schöne Bezeichnung, da wir uns die Bildvektoren der Basiselemente frei wünschen dürfen.

- **Seite 86**: Notiz 4 ist eine der wichtigsten Aussagen der linearen Algebra. Hier ein Beispiel: Sei $\mathbb{K} = \mathbb{R}$ und $\mathcal{P}_2$ gleich der Menge aller Polynomfunktionen $p : \mathbb{R} \to \mathbb{R}$ vom Grad ≤ 2. Mit den für Funktionen üblichen Definitionen

$$(p+q)(t) = p(t) + q(t) \qquad \text{(Summe zweier Funktionen)} \tag{16}$$

und

$$(\lambda p)(t) = \lambda p(t) \;\; \text{für } \lambda, t \in \mathbb{R} \qquad \text{(reelles Vielfaches einer Funktion)} \tag{17}$$

versehen ist $\mathcal{P}_2$ ein dreidimensionaler reeller Vektorraum. Schreiben wir ein allgemeines Element $p \in \mathcal{P}_2$ in der Form $p(t) = a_0 + a_1 t + a_2 t^2$ an, so ist

$$\begin{aligned} &\Phi : \mathbb{R}^3 \to \mathcal{P}_2 \\ &\Phi(a_0, a_1, a_2) = \text{die Polynomfunktion } t \mapsto a_0 + a_1 t + a_2 t^2 \end{aligned} \tag{18}$$

ein Isomorphismus. (Beweis: Übungen!) Um zwei Polynome in $\mathcal{P}_2$ zu addieren oder das λ-fache eines Polynoms zu bilden, können die entsprechenden Operationen genauso gut mit den aus den Koeffizienten gebilden Vektoren (a_0, a_1, a_2) in $\mathbb{R}^3$ ausgeführt und die Ergebnisse mit Φ in $\mathcal{P}_2$ rückübersetzt werden. Unter einem Gesichtspunkt, für den nur die lineare Struktur relevant ist (d.h. betrachtet durch die „Vektorraum-Brille") besteht zwischen $\mathcal{P}_2$ (bestehend aus Polynomfunktionen) und $\mathbb{R}^3$ (bestehend aus Zahlentripeln) kein besonderer Unterschied!

  Um an diesem sehr einfachen Beispiel zu illustrieren, wie ein konkreter Sachverhalt auf sein „lineares Substrat" reduziert werden kann, bemerken wir, dass die Ableitung p' jeder Polynomfunktion $p \in \mathcal{P}_2$ wieder ein Element von $\mathcal{P}_2$ ist. Die Zuordnung $p \mapsto p'$ ist sogar eine *lineare* Abbildung $\mathcal{P}_2 \to \mathcal{P}_2$ (Beweis: Übungen). Wegen

$$p(t) = a_0 + a_1 t + a_2 t^2 \;\; \Rightarrow \;\; p'(t) = a_1 + 2a_2 t \tag{19}$$

kann man die Ableitung von p auch ermitteln, indem man im $\mathbb{R}^3$ die lineare Operation

$$\begin{aligned} &D : \mathbb{R}^3 \to \mathbb{R}^3 \\ &D(a_0, a_1, a_2) = (a_1, 2a_2, 0) \end{aligned} \tag{20}$$

ausführt und das Ergebnis $(a_1, 2a_2, 0)$ mit Hilfe von Φ in $\mathcal{P}_2$ rückübersetzt! Die Abbildung D im $\mathbb{R}^3$ ist hinsichtlich der linearen Struktur dem Differenzieren in $\mathcal{P}_2$ „gleichwertig", sie „spiegelt" sie. Und vor allem: Sie sieht von allen sonstigen individuellen Eigenschaften der Polynomfunktionen (wie Nullstellen, Monotonieeigenschaften,...) ab, reduziert das Ableiten eines Polynoms (das eigentlich zum Geschäft der Analysis gehört) gewissermaßen auf seinen „buchhalterischen" Kern und macht es damit zu einem Thema der linearen Algebra.

Konstruktionen wie diese werden wir später noch genauer diskutieren.

- **Seite 87**: Für den Rang einer linearen Abbildung wird statt rg auch oft rank geschrieben.

  Die Beziehung $\dim(\mathrm{Kern}(f)) + \mathrm{rg}(f) = n$ (in Worten: „die Summe der Dimensionen von Kern und Bild einer linearen Abbildung $f : V \to W$ ist stets gleich der Dimension von V") ist bemerkenswert: Um die Dimension des Bildes von f zu bestimmen, muss man (neben der Dimension von V) nur die Dimension des Kerns von f kennen – das Bild von f muss dafür gar nicht untersucht werden!

  Um die Struktur des im Buch gegebenen Beweises besser zu verstehen hier der **Beweis für den Spezialfall** $\dim(\mathrm{Kern}(f)) = 2$ und $n = 4$:
  Sei (v_1, v_2) eine Basis von $\mathrm{Kern}(f)$. Wir ergänzen sie zu einer Basis (v_1, v_2, v_3, v_4) von V und setzen $w_1 = f(v_3)$ und $w_2 = f(v_4)$. Jedes $v \in V$ kann dann in der Form $v = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \lambda_4 v_4$ geschrieben werden. Damit wird

$$\begin{aligned} f(v) &= f(\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \lambda_4 v_4) = \\ &= \lambda_1 \underbrace{f(v_1)}_{0} + \lambda_2 \underbrace{f(v_2)}_{0} + \lambda_3 \underbrace{f(v_3)}_{w_1} + \lambda_4 \underbrace{f(v_4)}_{w_2} = \\ &= \lambda_3 w_1 + \lambda_4 w_2. \end{aligned} \tag{21}$$

  Somit ist jedes Bildelement eine Linearkombination von w_1 und w_2, was impliziert, dass $\mathrm{Bild}(f) = L(w_1, w_2)$ ist. Zuletzt wird gezeigt, dass (w_1, w_2) linear unabhängig und daher eine Basis von $\mathrm{Bild}(f)$ ist: Seien also $\alpha_1, \alpha_2 \in \mathbb{K}$ mit

$$\alpha_1 w_1 + \alpha_2 w_2 = 0. \tag{22}$$

  Das ist gleichbedeutend mit $\alpha_1 f(v_3) + \alpha_2 f(v_4) = 0$ oder $f(\alpha_1 v_3 + \alpha_2 v_4) = 0$. Somit ist $\alpha_1 v_3 + \alpha_2 v_4 \in \mathrm{Kern}(f)$ und kann, da (v_1, v_2) eine Basis von $\mathrm{Kern}(f)$ ist, in der Form

$$\alpha_1 v_3 + \alpha_2 v_4 = \lambda_1 v_1 + \lambda_2 v_2 \tag{23}$$

  geschrieben werden. Da aber (v_1, v_2, v_3, v_4) eine Basis von V und damit linear unabhängig ist, folgt $\alpha_1 = \alpha_2 = \lambda_1 = \lambda_2 = 0$. Damit wurde gezeigt:

$$\text{Aus } \alpha_1 w_1 + \alpha_2 w_2 = 0 \text{ folgt } \alpha_1 = \alpha_2 = 0, \tag{24}$$

  d.h. (w_1, w_2) ist linear unabhängig. Ergo $\mathrm{rg}(f) = \dim(\mathrm{Bild}(f)) = 2$, womit alles bewiesen ist. Abbildung 2 zeigt eine schematische Skizze des Sachverhalts.

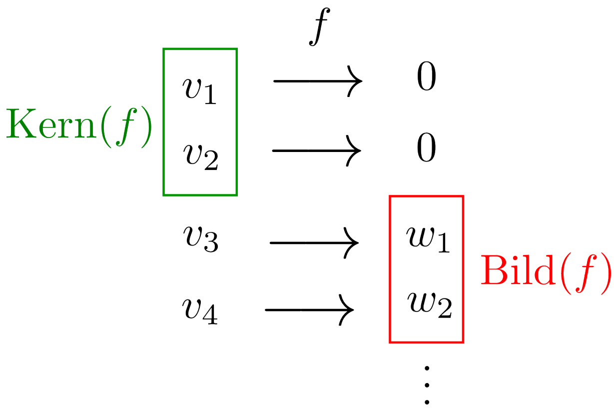

Abbildung 2: Schematische Skizze zum Beweis, dass aus $\dim(\mathrm{Kern}(f)) = 2$ und $n = 4$ folgt: $\mathrm{rg}(f) = 2$. Die Punkte deuten an, dass der Zielvektorraum W Elemente enthalten kann, die nicht von f getroffen werden. Die Dimension von W muss *mindestens* 2 sein, kann aber durchaus größer sein.

4.2 Matrizen

- **Seite 88**: Die *Koeffizienten* einer Matrix werden auch ihre *Komponenten* oder *Eintragungen* genannt.

- **Seite 89** (Definition): Im Buch werden für eine Matrix A und für ihre Koeffizienten $a_{11}, a_{12}, \ldots$ einander entsprechende, aber unterschiedliche Symbole verwendet. Man kann die Koeffizienten von A auch als $A_{11}, A_{12}, \ldots$ bezeichnen, ähnlich wie die Komponenten eines n-Tupels x oft als $x_1, x_2 \ldots$ geschrieben werden. Aber Achtung: Später im Buch (Kapitel 6) werden die Symbole $A_{11}, A_{12}, \ldots$ in einer anderen Bedeutung verwendet, weshalb ich mich in den meisten Fällen der Notation des Buches anschließe.

- **Seite 92**: Da das Symbol i für die imaginäre Einheit reserviert ist, sollte es eigentlich nicht zur Bezeichnung von Komponenten oder als Summationsindex verwendet werden. Statt δ_{ij} und a_{ij} schreibe ich lieber δ_{jk} und a_{jk}. Um Ihnen aber zusätzliche Übersetzungsarbeit zu ersparen, beuge ich mich der Notation des Buches und werde in der Vorlesung dort, wo es auch im Buch so steht, zähneknirschend das i als Index verwenden.

- **Seiten 91/94**: Auf diesen beiden Seiten steht jeweils eine wichtige Information, die wir deutlich herausheben wollen:

  – Seite 91, unterhalb des Kastens: Es gilt $M(m \times n, \mathbb{K}) \cong \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$, d.h. die $m \times n$-Matrizen können (in „natürlicher“ Weise) als die linearen Abbildungen $\mathbb{K}^n \to \mathbb{K}^m$ aufgefasst werden, und umgekehrt kann der Vektorraum aller linearen Abbildungen $\mathbb{K}^n \to \mathbb{K}^m$ (in „natürlicher“ Weise) mit dem Vektorraum aller $m \times n$-Matrizen identifiziert werden.

  – Seite 94: Ist $\dim(V) = n$ und $\dim(W) = m$, so ist $\mathrm{Hom}(V, W) \cong M(m \times n, \mathbb{K})$, d.h. der Vektorraum aller linearen Abbildungen $V \to W$ kann mit dem Vektorraum aller $m \times n$-Matrizen identifiziert werden. Diese Identifizierung ist allerdings nicht „natürlich“, sondern benötigt die Wahl je einer Basis von V und W. Sind die

Basen gewählt, so ist der Isomorphismus $\mathrm{Hom}(V,W) \xrightarrow{\cong} M(m \times n, \mathbb{K})$ gerade die Zuordnung $f \mapsto A$, die jeder linearen Abbildung f eine Matrix („ihre" Matrix bezüglich der gewählten Basen) A zuweist.

Wir merken uns also:

$$\mathrm{Hom}(V,W) \overset{\text{benötigt Basiswahl}}{\cong} \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m) \overset{\text{natürlich}}{\cong} M(m \times n, \mathbb{K}). \tag{25}$$

Auf Seite 111, Notiz 1, wird nachgetragen, dass die Dimension dieser Räume mn ist. Aber auch ohne jetzt vorzublättern sollte das klar sein!

4.3 Test

- Arbeiten Sie den Test durch!

4.4 Quotientenvektorräume

- Wir werden diesen Abschnitt anhand der *Äquivalenzrelation*

$$x \sim y \ \text{(„}x\text{ ist äquivalent zu }y\text{“)} \ :\Leftrightarrow \ x - y \in U \tag{26}$$

(für einen gegebenen Untervektorraum U von V und für beliebige $x, y \in V$) aufziehen. Für alle $x, y, z \in V$ erfüllt sie:

  - $x \sim x$ (Reflexivität).
  - Aus $x \sim y$ folgt $y \sim x$ (Symmetrie).
  - Aus $x \sim y$ und $y \sim z$ folgt $x \sim z$ (Transitivität).

(Jede Relation auf einer Menge, die diese drei Bedingungen erfüllt, heißt Äquivalenzrelation. Ein einfaches Beispiel für eine Äquivalenzrelation auf der Menge der in einem Hörsaal anwesenden Personen ist die Relation „ist im selben Jahr geboren wie".)

Für jedes $x \in V$ heißt die Menge $K(x)$ aller zu x äquivalenten Elemente von V eine *Nebenklasse*, und die (mit einer geeigneten Vektorraumstruktur versehene) Menge aller Nebenklassen ist der *Quotientenvektorraum* (kurz *Quotientenraum*) V/U (ausgesprochen meist als „V modulo U"). Die Menge $K(x)$ wird im Buch mit $x + U$ bezeichnet.

Statt Quotienten(vektor)raum ist manchmal auch die Bezeichnung *Faktorraum* zu lesen. Vergegenwärtigen Sie sich, dass ein Quotient auch ein Faktor ist: Im Rahmen der natürlichen Zahlen ist $\frac{12}{3}$ ist ein Quotient, und mit ihm wird ein Faktor von 12 (nämlich 4) angegeben. Nachdem dieser Abschnitt in der Vorlesung behandelt wurde, denken Sie darüber nach, in welchem Sinn V/U ein „Quotient" bzw. ein „Faktor" ist!

- **Seite 100** (Lemma): Ist $U = \mathrm{Kern}(f)$, so ist φ sogar injektiv. Daher gilt für jede lineare Abbildung $f : V \to W$ der sogenannte *Homomorphiesatz*[2]

$$V/\mathrm{Kern}(f) \cong \mathrm{Bild}(f). \tag{27}$$

[2] Der Isomorphismus $\varphi : V/\mathrm{Kern}(f) \to \mathrm{Bild}(f)$, dessen Existenz der Homomorphiesatz besagt, ist gegeben durch $\varphi : v + \mathrm{Kern}(f) \mapsto f(v)$.

Ohne also das Bild von f genauer untersuchen zu müssen, weiß man, dass es in natürlicher Weise „strukturgleich“ zu $V/\mathrm{Kern}(f)$ ist. Das ist einerseits erstaunlich, aber andererseits können wir uns leicht überlegen, dass die Linearität von f auf eine recht transparente Weise die Beziehung zwischen Bild und Kern stiftet:

$$\begin{aligned} x \text{ und } y \text{ haben das gleiche Bild } &\Leftrightarrow\ f(x) = f(y)\ \Leftrightarrow \\ &\Leftrightarrow\ f(x-y) = 0\ \Leftrightarrow\ x-y \in \mathrm{Kern}(f). \end{aligned} \tag{28}$$

Das ergibt noch einmal eine schöne Vorstellung vom Begriff des Quotientenvektorraums: Ist f eine besonders interessante Eigenschaft der Elemente von V, und möchte man wissen, welche Werte sie annehmen kann (also wie $\mathrm{Bild}(f)$ beschaffen ist), so erklärt man Elemente von V, die die gleiche Eigenschaft haben ($f(x) = f(y)$), als äquivalent und gruppiert äquivalente Elemente in Klassen. Dann gehört zu jedem Element von $\mathrm{Bild}(f)$ klarerweise genau eine solche Klasse, und die Menge aller Bildvektoren ($\mathrm{Bild}(f)$) entspricht der Menge aller Klassen ($V/\mathrm{Kern}(f)$).

Aus dem Homomorphiesatz folgt sofort (wenn man es nicht ohnehin schon wüsste), dass $\dim(V) - \dim(\mathrm{Kern}(f)) = \dim(\mathrm{Bild}(f))$ gilt. (Auf Seite 87, im Kasten, steht die gleiche Beziehung, wobei die Abkürzungen $\dim(V) = n$ und $\dim(\mathrm{Bild}(f)) = \mathrm{rg}(f)$ verwendet wurden.)

4.5 Drehungen und Spiegelungen des $\mathbb{R}^2$

- Keine Anmerkungen zu diesem Abschnitt.

---

5 Matrizenrechnung

5.1 Multiplikation

- **Seite 114** (Bemerkung 1): Eine *quadratische* Matrix ist eine $n \times n$-Matrix für ein $n \in \mathbb{N}\backslash\{0\}$. Überrascht?

- **Seite 114/115**: Hier wird die *Inverse* einer Matrix A (die zu A *inverse Matrix* A^{-1}) eingeführt. Obwohl aus mathematischer Sicht alles Nötige gesagt wird und auch die Beziehungen $AB = E$ und $BA = E$ mit $B = A^{-1}$ vorkommen, möchte ich letztere hier noch einmal offiziell in der Form

$$A\,A^{-1} = A^{-1}A = E \tag{29}$$

anschreiben. In dieser Form wird die Beziehung zwischen einer invertierbaren Matrix und ihrer Inversen meist benötigt. Beispiel: Gilt $Ax = b$ für $x, b \in \mathbb{K}^n$ und $A \in M(n \times n, \mathbb{K})$, und ist A invertierbar, so multiplizieren wir beide Seiten der Beziehung $Ax = b$ (von links) mit A^{-1} und erhalten mit

$$\underbrace{\underbrace{A^{-1}A}_{E}\,x}_{x} = A^{-1}b \tag{30}$$

sofort eine Lösungsformel für ein lineares $n \times n$-Gleichungssystem mit invertierbarer Koeffizientenmatrix! Lineare Gleichungssysteme sind das Thema von Kapitel 7.

Eine wichtige Rechenregel für das Invertieren von Matrizen sei noch hervorgehoben:

$$(AB)^{-1} = B^{-1}A^{-1}. \tag{31}$$

5.2 Rang einer Matrix

- Keine Anmerkungen zu diesem Abschnitt.

5.3 Elementare Umformungen

- Keine Anmerkungen zu diesem Abschnitt.

5.4 Test

- Arbeiten Sie den Test durch!

5.5 Wie invertiert man eine Matrix?

- Die einzige Anmerkung zu diesem Abschnitt: Nutzen Sie bei aufwändigen Rechnungen (wie der Berechnung der Inversen einer großen Matrix oder der Lösung eines umfangreichen linearen Gleichungssystems) geeignete elektronische Hilfsmittel! Machen Sie sich – möglichst bereits zu Beginn Ihres Studiums – mit einem Computeralgebra-System wie *Mathematica* vertraut! Die Syntax ist meist schnell gelernt, und nachdem Sie die ersten Schritte hinter sich gebracht haben, können Sie ein solches Werkzeug jederzeit für mathematische Aufgaben, die sich Ihnen stellen (auch zur Kontrolle der Berechnungen, die Sie auf dem Papier gemacht haben) einsetzen.

5.6 Mehr über Drehungen und Spiegelungen

- Keine Anmerkungen zu diesem Abschnitt.

5.7 Historische Notiz

- Lesenswert!

Ergänzung 1: Lineare Abbildungen und ihre Matrizen

Die Matrix einer linearen Abbildung

Ein für die Physik wichtiges Thema wollen wir nun etwas deutlicher herausarbeiten, als dies im Buch gemacht wird. Dafür kehren wir zunächst zu den Seiten 93/94 zurück, auf denen die Matrix einer linearen Abbildung definiert wird. Für das Folgende ist es nützlich, eigene Symbole für Basen zu verwenden: Ist V ein Vektorraum über $\mathbb{K}$, und ist $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V, so bezeichnen wir den *kanonischen Basisisomorphismus* $\mathbb{K}^n \to V$ kurz mit $\Phi_{\mathcal{B}}$. $\Phi_{\mathcal{B}}$

ordnet jedem n-Tupel $(\lambda_1, \ldots, \lambda_n)$ von Elementen aus $\mathbb{K}$ den Vektor $\lambda_1 v_1 + \ldots \lambda_n v_n \in V$ zu. Umgekehrt ordnet $\Phi_{\mathcal{B}}^{-1}$ jedem Vektor v seine Entwicklungskoeffizienten $(\lambda_1, \ldots, \lambda_n)$ bezüglich der Basis $\mathcal{B}$ zu.

Sind V und W Vektorräume über $\mathbb{K}$, ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V und $\mathcal{C} = (w_1, \ldots, w_m)$ eine Basis von W, und ist schließlich $f : V \to W$ eine lineare Abbildung, so geben wir der zugehörigen (im Buch A genannten) linearen Abbildung $\mathbb{K}^n \to \mathbb{K}^m$ eine eigene Bezeichnung:

$$[f]_{\mathcal{C},\mathcal{B}} := \Phi_{\mathcal{C}}^{-1} \circ f \circ \Phi_{\mathcal{B}}. \tag{32}$$

Als lineare Abbildung $\mathbb{K}^n \to \mathbb{K}^m$ kann sie auch als $m \times n$-Matrix über $\mathbb{K}$, d.h. als Element von $M(m \times n, \mathbb{K})$ angesehen werden und heißt die *Matrix* (oder *Matrixdarstellung*) *von* f bezüglich der Basen $\mathcal{B}$ und $\mathcal{C}$. Das kommutative Diagramm, das die Wirkung von $[f]_{\mathcal{C},\mathcal{B}}$ beschreibt, sieht so aus:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \Phi_{\mathcal{B}} \uparrow \cong & & \cong \uparrow \Phi_{\mathcal{C}} \\ \mathbb{K}^n & \xrightarrow{[f]_{\mathcal{C},\mathcal{B}}} & \mathbb{K}^m \end{array}$$

Es ist ist genau dasselbe wie im Buch auf Seite 94, nur jetzt mit unseren neuen, praktischen Bezeichnungen.

Es ist dies eine wichtige und (gerade für die Physik) extrem nützliche Konstruktion. Versuchen Sie daher, sie wirklich zu verstehen! Sehen wir uns an, wozu sie gut ist: Um die Wirkung von f auf ein Element $x \in V$ zu beschreiben (oder zu berechnen), kann folgendermaßen vorgegangen werden:

- Wir beginnen mit einem $x \in V$.
- Anstatt sofort f darauf wirken zu lassen, „übersetzen“ wir x zunächst einmal mit Hilfe der zu $\Phi_{\mathcal{B}}$ inversen Abbildung in ein n-Tupel von Skalaren:

  $$x \in V \quad \overset{\Phi_{\mathcal{B}}^{-1}}{\longmapsto} \quad \Phi_{\mathcal{B}}^{-1}(x) \in \mathbb{K}^n. \tag{33}$$

  (Im obigen kommutativen Diagramm folgen wir dem linken Pfeil in die Gegenrichtung, also von oben nach unten.)
- Auf das so erhaltene n-Tupel wirkt die Matrix von f:

  $$\Phi_{\mathcal{B}}^{-1}(x) \in \mathbb{K}^n \quad \overset{[f]_{\mathcal{C},\mathcal{B}}}{\longmapsto} \quad [f]_{\mathcal{C},\mathcal{B}}\, \Phi_{\mathcal{B}}^{-1}(x) \in \mathbb{K}^m. \tag{34}$$

  (Im obigen kommutativen Diagramm folgen wir dem unteren Pfeil.)
- Nun müssen wir das Ergebnis mit Hilfe von $\Phi_{\mathcal{C}}$ noch in ein Element von W „rückübersetzen“:

  $$[f]_{\mathcal{C},\mathcal{B}}\, \Phi_{\mathcal{B}}^{-1}(x) \in \mathbb{K}^m \quad \overset{\Phi_{\mathcal{C}}}{\longmapsto} \quad \Phi_{\mathcal{C}}\left([f]_{\mathcal{C},\mathcal{B}}\, \Phi_{\mathcal{B}}^{-1}(x)\right) \in W. \tag{35}$$

  (Im obigen kommutativen Diagramm folgen wir dem rechten Pfeil.)

Damit haben wir die Wirkung von f im Umweg über $\mathbb{K}^n$ und $\mathbb{K}^m$ beschrieben:

$$f(x) = \Phi_\mathcal{C}\left([f]_{\mathcal{C},\mathcal{B}}\, \Phi_\mathcal{B}^{-1}(x)\right) \quad \text{für alle}\, x \in V \tag{36}$$

oder kurz[3]

$$f = \Phi_\mathcal{C} \circ [f]_{\mathcal{C},\mathcal{B}} \circ \Phi_\mathcal{B}^{-1}. \tag{37}$$

Um die Wirkung von f (beispielsweise einem ganz komplizierten linearen Operator von einem ganz schwer zu verstehenden Vektorraum V in einen ebenso schwer zu verstehenden Vektorraum W) zu studieren, kann die Wirkung der Matrix von f auf n-Tupel von Zahlen (also eine Operation, an der grundsätzlich nichts „ganz schwer zu verstehen" ist) untersucht werden. Die beiden Φ's dienen dann lediglich der Hin- und Zurückrechnung zwischen Tupeln und Elementen von V bzw. W. Hinsichtlich der linearen Struktur (betrachtet durch die „Vektorraum-Brille") besteht also zwischen f und $[f]_{\mathcal{C},\mathcal{B}}$ kein besonderer Unterschied! Viele Eigenschaften von f, auf die wir im Laufe dieser Vorlesung zu sprechen kommen werden, sind für beide gleich. Da eine Matrix bloß ein „Zahlenschema" ist, wurde damit von allen individuellen Eigenschaften der Elemente von V und W sowie von allen individuellen Eigenschaften und Finessen der Abbildung f abgesehen und die Wirkung von f letztlich auf eine reine Haushalterei mit Zahlen (auf seinen „buchhalterischen" Kern) reduziert.

Ist $V = W$ (die in Anwendungen am häufigsten auftretende Situation), so ist es meist zweckmäßig, nur eine einzige Basis zu verwenden, also $\mathcal{C} = \mathcal{B}$ zu setzen. In diesem Fall wollen wir die Matrix von f einfach als $[f]_\mathcal{B}$ bezeichnen: die Matrix(darstellung) von f bezüglich der Basis $\mathcal{B}$. Es gilt dann

$$f = \Phi_\mathcal{B} \circ [f]_\mathcal{B} \circ \Phi_\mathcal{B}^{-1} \qquad \text{mit} \qquad [f]_\mathcal{B} = \Phi_\mathcal{B}^{-1} \circ f \circ \Phi_\mathcal{B}, \tag{38}$$

und die ganze Logik des „Umwegs" über den $\mathbb{R}^n$ kann nun wie in (33) – (36) mit $W = V$, $m = n$, $\mathcal{C} = \mathcal{B}$ (und der Bezeichnung $[f]_\mathcal{B}$ statt $[f]_{\mathcal{B},\mathcal{B}}$) noch einmal durchexerziert werden.

Beispiel:
Als Beispiel nehmen wir uns den weiter oben bereits eingeführten Vektorraum $\mathcal{P}_2$ aller reellen Polynomfunktionen vom Grad ≤ 2 vor. In der in (18) definierten linearen Abbildung Φ erkennen wir den kanonischen Basisisomorphismus $\Phi_\mathcal{B}$ bezüglich der Basis

$$\mathcal{B} = \left(t \mapsto 1, t \mapsto t, t \mapsto t^2\right) \tag{39}$$

von $\mathcal{P}_2$. Als lineare Abbildung $f : \mathcal{P}_2 \to \mathcal{P}_2$ nehmen wir das Bilden der Ableitung:

$$f(p) = p'. \tag{40}$$

Die dem Bilden der Ableitung entsprechende Abbildung

$$\begin{aligned} &\mathbb{R}^3 \to \mathbb{R}^3 \\ &(a_0, a_1, a_2) \mapsto (a_1, 2a_2, 0) \end{aligned} \tag{41}$$

[3] In der linearen Algebra wird das Zeichen $\circ$ für die Verkettung oft weggelassen, womit beispielsweise (37) auch in der Form

$$f = \Phi_\mathcal{C}\, [f]_{\mathcal{C},\mathcal{B}}\, \Phi_\mathcal{B}^{-1}$$

geschrieben werden kann.

haben wir bereits in (20) hingeschrieben (und dort D genannt). Ohne großen Aufwand erkennen wir ihre Matrix, die wir mit unserer neuen Bezeichnung als

$$[f]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \tag{42}$$

hinschreiben. Kurzer Check:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \\ 0 \end{pmatrix}. \tag{43}$$

Um uns die vorliegende Situation möglichst genau vor Augen zu halten, schreiben wir den kanonischen Basisisomorphismus noch einmal an,

$$\begin{aligned} &\Phi_{\mathcal{B}} : \mathbb{R}^3 \to \mathcal{P}_2 \\ &\Phi_{\mathcal{B}}(a_0, a_1, a_2) = \text{die Polynomfunktion } t \mapsto a_0 + a_1 t + a_2 t^2 , \end{aligned} \tag{44}$$

und notieren auch dessen Inverse:

$$\begin{aligned} &\Phi_{\mathcal{B}}^{-1} : \mathcal{P}_2 \to \mathbb{R}^3 \\ &\Phi_{\mathcal{B}}^{-1}\big(\text{die Polynomfunktion } t \mapsto a_0 + a_1 t + a_2 t^2\big) = (a_0, a_1, a_2). \end{aligned} \tag{45}$$

Mit (40) – (45) können nun die in (38) angegebenen Beziehungen zwischen f und seiner Matrix $[f]_{\mathcal{B}}$ anhand unseres Beispiels im Detail nachvollzogen werden. Für Berechnungen mit der Matrizenmultiplikation sind die hier auftretenden Zahlentripel (a_0, a_1, a_2) und $(a_1, 2a_2, 0)$ als Spaltenvektoren

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} a_1 \\ 2a_2 \\ 0 \end{pmatrix} \tag{46}$$

anzuschreiben. Erinnern Sie sich bitte an die oben zur Seite 7 des Buches gemachte Anmerkung!

Im Hinblick auf physikalische Anwendungen sehen wir uns jetzt an, wie eine lineare Abbildung f „in Komponenten“ beschrieben wird. „In Komponenten“ (oder „in Komponentenschreibweise“) bedeutet, dass Vektoren und lineare Abbildungen jeweils auf eine Basis bezogen und die uns interessierenden Beziehungen durch die entsprechenden Koeffizienten (= Komponenten) ausgedrückt werden. Wir wollen uns dabei auf den Fall $V = W$ beschränken. Sei also $f : V \to V$ eine lineare Abbildung, und sei $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V. Schreiben wir die Entwicklung eines Vektors $x \in V$ bezüglich der Basis $\mathcal{B}$ in der Form $x = \sum_{j=1}^{n} \lambda_j v_j$ an, so ergibt sich die Wirkung der Abbildung f aufgrund ihrer Linearität zu

$$f(x) = f\left(\sum_{j=1}^{n} \lambda_j v_j\right) = \sum_{j=1}^{n} \lambda_j f(v_j). \tag{47}$$

Nun kann jedes Element $f(v_j) \in V$ wieder in die Basis $\mathcal{B}$ entwickelt werden, was wir in der Form

$$f(v_j) = \sum_{k=1}^{n} v_k T_{kj} \tag{48}$$

anschreiben. Hier treten also ganz automatisch die Koeffizienten einer $n \times n$-Matrix T auf. Mit

$$f(x) = \sum_{j=1}^{n} \lambda_j f(v_j) = \sum_{j,k=1}^{n} \lambda_j v_k T_{kj} \equiv \sum_{k=1}^{n} \left(\sum_{j=1}^{n} T_{kj} \lambda_j \right) v_k \tag{49}$$

können wir die Entwicklungskoeffizienten von $f(x)$ ablesen und die Wirkung von $[f]_\mathcal{B}$ nachvollziehen:

$$(\lambda_1, \dots \lambda_n) \overset{\Phi_\mathcal{B}}{\longmapsto} x \overset{f}{\longmapsto} f(x) \overset{\Phi_\mathcal{B}^{-1}}{\longmapsto} \left(\sum_{j=1}^{n} T_{1j} \lambda_j, \dots, \sum_{j=1}^{n} T_{nj} \lambda_j \right). \tag{50}$$

Diese Zuordnung bescheibt aber unter dem Strich genau die Wirkung der Matrix T auf $\mathbb{K}^n$. Daher sind die Matrizen $[f]_\mathcal{B}$ und T identisch! In Komponenten:

$$([f]_\mathcal{B})_{jk} = T_{jk}\,. \tag{51}$$

Formel (48), die uns die Komponenten von $[f]_\mathcal{B}$ liefert, steht auch im Buch, allerdings viel weiter hinten, auf Seite 215. Wir schreiben Sie nochmal offiziell in der Form

$$f(v_j) = \sum_{k=1}^{n} v_k \left([f]_\mathcal{B}\right)_{kj} \tag{52}$$

an.

Um also die Koeffizienten der Matrix einer linearen Abbildung $f : V \to V$ bezüglich einer Basis $\mathcal{B}$ von V in der Praxis zu ermitteln, haben Sie zwei Möglichkeiten:

- Sie entwickeln die Bilder der Basisvektoren, also die $f(v_j)$, in die Basis $\mathcal{B}$ und nutzen (52), um die Komponenten der Matrix von f abzulesen.
- Bei einfachen Beispielen kann es auch leichter gehen: Wie bereits oben anhand eines Beispiels vorgeführt ($V = \mathcal{P}_2$, $\mathcal{B} = (t \mapsto 1, t \mapsto t, t \mapsto t^2)$ und $f =$ „Bilden der Ableitung“) ist es manchmal möglich, ein „allgemeines“ Element von V anzuschreiben und den kanonischen Basisisomorphismus ohne großen Aufwand direkt zu identifizieren. Die Wirkung der Matrix von f ergibt sich dann unmittelbar aus ihrer Definition $[f]_\mathcal{B} := \Phi_\mathcal{B}^{-1} \circ f \circ \Phi_\mathcal{B}$. Halten Sie sich in diesem Fall immer ein kommutatives Diagramm wie

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ \Phi_\mathcal{B} \big\uparrow \cong & & \cong \big\uparrow \Phi_\mathcal{B} \\ \mathbb{K}^n & \xrightarrow{[f]_\mathcal{B}} & \mathbb{K}^n \end{array}$$

vor Augen! Es kann helfen, langwierige Rechnungen zu vermeiden! Sie können es auch mit konkreten Elementen aufzeichnen, wie etwa anhand des Beispiels mit $\mathcal{P}_2$ und $f =$ „Bilden der Ableitung“:

$$\begin{array}{ccc} \text{Polynom } t \mapsto a_0 + a_1 t + a_2 t^2 & \xmapsto{f} & \text{Polynom } t \mapsto a_1 + 2a_2 t \\ \Phi_{\mathcal{B}} \big\uparrow \cong & & \cong \big\uparrow \Phi_{\mathcal{B}} \\ (a_0, a_1, a_2) & \xmapsto{[f]_{\mathcal{B}}} & (a_1, 2a_2, 0) \end{array}$$

Sie können die Elemente des $\mathbb{K}^n$ natürlich auch gleich in Spaltenform anschreiben, hier wieder anhand des letzten Beispiels:

$$\begin{array}{ccc} \text{Polynom } t \mapsto a_0 + a_1 t + a_2 t^2 & \xmapsto{f} & \text{Polynom } t \mapsto a_1 + 2a_2 t \\ \Phi_{\mathcal{B}} \big\uparrow \cong & & \cong \big\uparrow \Phi_{\mathcal{B}} \\ \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} & \xmapsto{[f]_{\mathcal{B}}} & \begin{pmatrix} a_1 \\ 2a_2 \\ 0 \end{pmatrix} \end{array}$$

Wenn es Ihnen jetzt noch immer schwer fällt, die gesuchte Matrix $[f]_{\mathcal{B}}$ hinzuschreiben, setzen Sie für $\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$ die Vektoren der Standardbasis ein – dann erhalten Sie direkt die Spalten von $[f]_{\mathcal{B}}$.

Die Matrix eines Elements von V und die in den $\mathbb{K}^n$ gespiegelte Wirkung von f

Die Beziehung zwischen der Wirkung von f und der Wirkung von $[f]_{\mathcal{B}}$ kann auch in einer anderen, sehr suggestiven Schreibweise dargestellt werden: Für jedes $x \in V$ führen wir die Bezeichnung $[x]_{\mathcal{B}} = \Phi_{\mathcal{B}}^{-1}(x) \in \mathbb{K}^n$ ein. Dieser Vektor besteht also aus den Entwicklungskoeffizienten von x bezüglich der Basis $\mathcal{B}$. Wir können ihn als „Matrixdarstellung von x bezüglich der Basis $\mathcal{B}$“ ansehen und schreiben ihn am besten gleich in Spaltenform an:

$$[x]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}. \tag{53}$$

Analog ist dann auch $[f(x)]_{\mathcal{B}}$ definiert, womit die gänzlich „in den $\mathbb{K}^n$ gespiegelte“ Wirkung der Abbildung f die Form

$$[f(x)]_{\mathcal{B}} = [f]_{\mathcal{B}}\, [x]_{\mathcal{B}} \tag{54}$$

annimmt („Matrix von $f(x)$ = Matrix von f mal Matrix von x“). Wenn Sie diese Beziehung verstanden haben (und auch ihre ästhetische Ausstrahlung genießen können), liegt Ihnen schon ein großer Teil der linearen Algebra zu Füßen.

Die Matrix einer Verkettung linearer Abbildungen
Doch damit nicht genug: Ist eine zweite lineare Abbildung $g : V \to V$ gegeben, so können wir mit ihr (bezüglich derselben Basis $\mathcal{B}$) dasselbe machen wie mit f, womit sogleich folgt

$$[g(f(x))]_{\mathcal{B}} = [g]_{\mathcal{B}}\,[f(x)]_{\mathcal{B}} = [g]_{\mathcal{B}}\,[f]_{\mathcal{B}}\,[x]_{\mathcal{B}} \qquad \text{für beliebige } [x]_{\mathcal{B}} \in \mathbb{K}^n. \tag{55}$$

Daher gilt[4]

$$[g \circ f]_{\mathcal{B}} = [g]_{\mathcal{B}}\,[f]_{\mathcal{B}}, \tag{56}$$

d.h. die Matrix einer Verkettung (Hintereinanderausführung) linearer Abbildungen ist gleich dem Produkt der entsprechenden Matrizen. Zusammen mit den (aus der Linearität von $\Phi_{\mathcal{B}}$ folgenden) Beziehungen

$$[f + g]_{\mathcal{B}} = [f]_{\mathcal{B}} + [g]_{\mathcal{B}} \tag{57}$$

$$[\lambda f]_{\mathcal{B}} = \lambda\,[f]_{\mathcal{B}} \qquad \text{für } \lambda \in \mathbb{K} \tag{58}$$

und den selbstverständlichen Tatsachen

$$[\mathrm{Id}_V]_{\mathcal{B}} = E \qquad \text{und} \qquad [0]_{\mathcal{B}} = 0 \tag{59}$$

hat (56) in der Anwendung weitreichende Konsequenzen, da beliebige Kombinationen von Additionen, Bilden von Vielfachen und Hintereinanderausführungen von linearen Abbildungen $V \to V$ ersatzweise auch mit den entsprechenden Matrizen bezüglich einer festgehaltenen Basis durchgeführt werden können. Beispielsweise:

- Für drei lineare Abbildungen $f, g, h : V \to V$ gilt die Beziehung $h = 5\,\mathrm{Id}_V - f + 2\,g - 3\,g^2$ (wobei g^2 als $g \circ g$ zu verstehen ist) genau dann, wenn für ihre Matrizen $[h]_{\mathcal{B}} = 5\,\mathrm{E} - [f]_{\mathcal{B}} + 2\,[g]_{\mathcal{B}} - 3\,[g]_{\mathcal{B}}^2$ gilt.
- Für eine lineare Abbildung $f : V \to V$ und einen Vektor $x \in V$ gilt genau dann $f(x) = -3\,x$, wenn $[f]_{\mathcal{B}}\,[x]_{\mathcal{B}} = -3\,[x]_{\mathcal{B}}$ gilt. Um etwa alle Vektoren $x \in V$ zu finden, die $f(x) = -3\,x$ erfüllen, kann anstelle einer Berechnung in V ebensogut nach allen $\xi \in \mathbb{K}^n$ gesucht werden, die $[f]_{\mathcal{B}}\,\xi = -3\,\xi$ erfüllen.

Insgesamt führt unsere Analyse also zu dem Schluss, dass praktisch alles, was man mit linearen Abbildungen in einem n-dimensionalen $\mathbb{K}$-Vektorraum hinsichtlich der „linearen Struktur" machen kann, ebensogut im $\mathbb{K}^n$ gemacht werden kann.

Beispiel:
Wir greifen das zuvor diskutierte Beispiel mit dem Vektorraum $\mathcal{P}_2$ und der linearen Abbildung $f =$ „Bilden der Ableitung" wieder auf und bilden höhere Ableitungen, aber nicht direkt durch Hintereinander-Ausführen von f, sondern durch Hintereinander-Ausführen der Matrix von f. Mit (42) und der allgemeinen Regel

[4] Oder, mit der Kurzschreibweise $g\,f$ für $g \circ f$,

$$[g\,f]_{\mathcal{B}} = [g]_{\mathcal{B}}\,[f]_{\mathcal{B}}.$$

(56) mit $g = f$ berechnen wir die Matrix für das Bilden der zweiten Ableitung (mit der Kurzschreibweise f^2 für $f \circ f$)

$$[f^2]_\mathcal{B} = [f]_\mathcal{B}{}^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{60}$$

und die Matrix für das Bilden der dritten Ableitung

$$[f^3]_\mathcal{B} = [f]_\mathcal{B}{}^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{61}$$

Die in (60) und (61) erhaltenen Matrizen sollten uns nicht wirklich überraschen. (Warum?) Klarerweise verschwinden auch alle höheren Potenz der Matrix von f. Auch das sollte uns nicht überraschen. (Warum?) Sobald die Matrix von f (die wir uns durch einmaliges Differenzieren eines allgemeinen Polynoms vom Grad ≤ 2 verschafft haben) und der kanonische Basisisomorphismus bekannt waren, konnten wir die Wirkungen aller höheren Ableitungen in $\mathcal{P}_2$ durch bloße Matrizenmultiplikationen ermitteln! (Natürlich ist das ein sehr einfaches Beispiel – wer tut sich soviel Arbeit an, um ein Polynom vom Grad ≤ 2 einmal, zweimal und dreimal zu differenzieren? Aber wenn die Sachlage komplizierter wird – und in der Quantenmechanik wird sie komplizierter –, kann die Übersetzung in die Matrizensprache eine große Hilfe sein.)

Eine analoge Analyse kann bei Bedarf auch für lineare Abbildungen $f : V \to W$ mit verschiedenen Vektorräumen V und W über dem gleichen Grundkörper $\mathbb{K}$ durchgeführt werden. Anstelle von (54) heißt es dann

$$[f(x)]_\mathcal{C} = [f]_{\mathcal{C},\mathcal{B}}\,[x]_\mathcal{B}\,, \tag{62}$$

d.h. man muss nun berücksichtigen, dass die Basis $\mathcal{B}$ für V und die Basis $\mathcal{C}$ für W zuständig ist. Verkettungen mehrerer linearer Abbildungen sind nur dann möglich, wenn die Dimensionen der beteiligten Vektorräume das zulassen, und das ist klarerweise genau dann der Fall, wenn die entsprechenden Matrizen multipliziert werden können.

Transformationsverhalten unter Basiswechsel

Einen Sachverhalt, der im Buch in einem späteren Kapitel eher nebenbei erwähnt wird, wollen wir hier noch eigens hervorheben: Sei V ein Vektorraum über $\mathbb{K}$ und $f : V \to V$ eine lineare Abbildung. Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine Basis von V und $\Phi_\mathcal{B} : \mathbb{K}^n \to V$ der kanonische Basisisomorphismus, so ist, wie wir bereits wissen, die Matrix von f bezüglich der Basis $\mathcal{B}$ durch

$$[f]_\mathcal{B} = \Phi_\mathcal{B}^{-1} \circ f \circ \Phi_\mathcal{B} \tag{63}$$

gegeben. Ist nun $\mathcal{B}' = (v'_1, \ldots, v'_n)$ eine *andere* Basis von V und $\Phi_{\mathcal{B}'} : \mathbb{K}^n \to V$ der zugehörige kanonische Basisisomorphismus, daher

$$[f]_{\mathcal{B}'} = \Phi_{\mathcal{B}'}^{-1} \circ f \circ \Phi_{\mathcal{B}'} \tag{64}$$

die Matrix von f bezüglich dieser anderen Basis, so stellt sich die Frage, wie die Matrizen $[f]_{\mathcal{B}}$ und $[f]_{\mathcal{B}'}$ zusammenhängen bzw. ineinander umgerechnet werden können, d.h. mit anderen Worten, wie ihr *Transformationsverhalten unter Basiswechsel* aussieht. Mit

$$f = \Phi_{\mathcal{B}} \circ [f]_{\mathcal{B}} \circ \Phi_{\mathcal{B}}^{-1} \tag{65}$$

erhalten wir aus (64)

$$[f]_{\mathcal{B}'} = \Phi_{\mathcal{B}'}^{-1} \circ \Phi_{\mathcal{B}} \circ [f]_{\mathcal{B}} \circ \Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{B}'} = \left(\Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{B}'}\right)^{-1} \circ [f]_{\mathcal{B}} \circ \Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{B}'}, \tag{66}$$

was mit der Abkürzung $S = \Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{B}'}$ in der Form

$$[f]_{\mathcal{B}'} = S^{-1}[f]_{\mathcal{B}}\, S \tag{67}$$

geschrieben werden kann. Da S und S^{-1} lineare Abbildungen $\mathbb{K}^n \to \mathbb{K}^n$ sind und daher als $n \times n$-Matrizen über $\mathbb{K}$ angesehen werden können, handelt es sich bei der rechten Seite dieser Formel um ein Produkt dreier Matrizen.

Dieses Ergebnis nehmen wir zum Anlass, um ganz allgemein zu definieren[5]: Zwei $n\times n$-Matrizen A und B heißen (zueinander) *ähnlich*, wenn es eine invertierbare $n \times n$-Matrix S gibt, so dass $B = S^{-1}A\,S$ gilt. Die Ähnlichkeit von Matrizen ist eine Äquivalenzrelation: Die Reflexivität ist evident (mit $S = E$), die Symmetrie folgt daraus, dass $B = S^{-1}A\,S$ auch in der Form $A = S\,B\,S^{-1} \equiv (S^{-1})^{-1}B\,S^{-1}$ geschrieben werden kann, und warum die Transitivität gegeben ist, zeigen Sie in den Übungen.

Zwei Matrizen derselben linearen Abbildung f bezüglich unterschiedlicher Basen sind also zueinander ähnlich. Jede Eigenschaft, die zueinander ähnliche Matrizen stets teilen, ist damit automatisch eine Eigenschaft der linearen Abbildung f. Ein Beispiel ist die *Spur*: Die Spur einer $n \times n$-Matrix $A = (a_{jk})$ ist die Summe ihrer Diagonalelemente:

$$\mathrm{Tr}(A) = \sum_{j=1}^{n} a_{jj}\,, \tag{68}$$

wobei Tr für *trace* steht[6]. Eine wichtige Rechenregel für die Spur ist

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA) \tag{69}$$

für beliebige $n\times n$-Matrizen A, B (Übungen!), womit es ganz leicht ist, zu zeigen, dass ähnliche Matrizen stets die gleiche Spur haben: Mit $B = S^{-1}A\,S$ schließen wir

$$\mathrm{Tr}(B) = \mathrm{Tr}(S^{-1}A\,S) = \mathrm{Tr}(A\,S\,S^{-1}) = \mathrm{Tr}(AE) = \mathrm{Tr}(A). \tag{70}$$

Folglich ist somit auch definiert, was die Spur $\mathrm{Tr}(f)$ einer linearen Abbildung $f : V \to V$ ist (obwohl eine lineare Abbildung an sich gar keine „Diagonalelemente“ besitzt). Gleiches gilt

[5] Diese Definition steht auch im Buch, allerdings erst auf Seite 233.

[6] Manchmal wird für die Spur auch Sp geschrieben – dieses Symbol reservieren wir uns aber für das Spektrum.

für den bereits besprochenen *Rang* und für die *Determinante*, die in Kapitel 6 drankommt. Eigenschaften wie diese werden (*unter Basiswechsel*) *invariant* genannt.

Wir kehren nun zur Beziehung (67) zurück. Welche Bedeutung hat die Matrix S, und wie kann man ihre Koeffizienten in der Praxis berechnen? Dazu entwickeln wir die Vektoren der Basis $\mathcal{B}'$ in die Basis $\mathcal{B}$ und *vice versa*:

$$v'_k = \sum_{j=1}^{n} v_j R_{jk} \qquad \text{daher} \qquad v_k = \sum_{j=1}^{n} v'_j (R^{-1})_{jk}\,, \tag{71}$$

wobei R eine intervierbare $n\times n$-Matrix über $\mathbb{K}$ ist. Die Entwicklungskoeffizienten eines Vektors $x \in V$ bezüglich der beiden Basen

$$x = \sum_{k=1}^{n} \lambda_k v_k = \sum_{k=1}^{n} \lambda'_k v'_k \tag{72}$$

werden daher mit Hilfe der Formeln

$$\lambda'_j = \sum_{k=1}^{n} (R^{-1})_{jk} \lambda_k \qquad \text{und} \qquad \lambda_j = \sum_{k=1}^{n} R_{jk} \lambda'_k \tag{73}$$

ineinander umgerechnet. Da nun die Wirkung der Matrix S, die ja als $\Phi_{\mathcal{B}}^{-1} \circ \Phi_{\mathcal{B}'}$ definiert war, (schematisch) so aussieht:

$$(\lambda'_1, \ldots \lambda'_n) \stackrel{\Phi_{\mathcal{B}'}}{\longmapsto} x \stackrel{\Phi_{\mathcal{B}}^{-1}}{\longmapsto} (\lambda_1, \ldots \lambda_n), \tag{74}$$

erkennen wir durch Vergleich mit der zweiten Beziehung in (73), dass die Matrizen S und R übereinstimmen! Wir erhalten also als Ergebnis, dass S genau jene Matrix ist, die in den Umrechnungsformeln der Basen bzw. der Vektorkomponenten jeweils in der Form

$$v'_k = \sum_{j=1}^{n} v_j S_{jk} \qquad \text{bzw.} \qquad \lambda_j = \sum_{k=1}^{n} S_{jk} \lambda'_k \tag{75}$$

auftritt. In vielen praktischen Fällen ist ein Basiswechsel direkt in der Form $v'_k = \sum\limits_{j=1}^{n} v_j c_{jk}$ angegeben. In einem solchen Fall wissen wir sofort, dass $S_{jk} = c_{jk}$ gilt.

Eine analoge Analyse kann auch für das Verhalten der Matrix einer linearen Abbildung $f : V \to W$ bezüglich zweier Basen $\mathcal{B}$ und $\mathcal{C}$ unter einem Wechsel *beider* Basen durchgeführt werden.

Beispiel:
Um diese theoretischen Überlegungen anhand eines konkreten Beispiels vorzuführen, betrachten wir jene lineare Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$, die durch die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \tag{76}$$

definiert ist. Bezüglich der Standardbasis $\mathcal{B} = (e_1, e_2)$ des $\mathbb{R}^2$ stimmt die Matrix von f natürlich mit A überein: $[f]_{\mathcal{B}} = A$. Die Vektoren e_k $(k = 1, 2)$ spielen die Rolle der v_k in den zuvor angestellten allgemeinen Überlegungen. Nun gehen wir zu einer anderen Basis des $\mathbb{R}^2$ über: Sei $\mathcal{B}' = (v_1', v_2')$ mit

$$v_1' = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{und} \qquad v_2' = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{77}$$

Die Matrix von f bezüglich dieser neuen Basis, $[f]_{\mathcal{B}'}$, wird mittels (67) gewonnen, wobei die Matrix S durch (75) bestimmt ist. Wir benutzen insbesondere die erste Beziehung von (75). In Worten ausgedrückt besagt sie, dass die Komponenten der Matrix S die Entwicklungskoeffizienten der neuen Basisvektoren sind, wenn diese in die alte Basis entwickelt werden. In unserem Beispiel bedeutet das:

$$v_1' = e_1 S_{11} + e_2 S_{21} \tag{78}$$

$$v_2' = e_1 S_{12} + e_2 S_{22}. \tag{79}$$

Wir entwickeln die neuen Basisvektoren in die alte Basis:

$$v_1' = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot 1 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot 1 = e_1 \cdot 1 + e_2 \cdot 1 \tag{80}$$

$$v_2' = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot 1 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot (-1) = e_1 \cdot 1 + e_2 \cdot (-1). \tag{81}$$

Durch Vergleich mit (78) – (79) lesen wir die Komponenten von S ab: $S_{11} = 1$, $S_{21} = 1$, $S_{12} = 1$ und $S_{22} = -1$. Wir erhalten also die Matrix

$$S = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{und daraus ihre Inverse} \quad S^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}. \tag{82}$$

Damit können wir (67) anwenden:

$$\begin{aligned} [f]_{\mathcal{B}'} &= S^{-1} [f]_{\mathcal{B}} \, S = S^{-1} A S = \\ &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}, \end{aligned} \tag{83}$$

wie Sie leicht nachrechnen können. Damit ist die Matrix von f bezüglich der Basis $\mathcal{B}'$ gefunden.

Nachbemerkung:
Die Matrix $[f]_{\mathcal{B}'}$ ist eine *Diagonalmatrix*, was bei Berechnungen sehr hilfreich sein kann – ein Aspekt, der später bei der Besprechung von Eigenwerten, Eigenvektoren und Diagonalisierbarkeit wieder aufgegriffen werden wird. Wir wollen hier nur erwähnen, dass die neuen (wie sich nun herausstellt geschickt gewählten) Basisvektoren die schöne Eigenschaft haben, dass sie von f in Vielfache ihrer selbst übergeführt werden: Mit (54) und den offensichtlichen Beziehungen[7]

$$[v_1']_{\mathcal{B}'} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{und} \qquad [v_2']_{\mathcal{B}'} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{84}$$

[7] Ist Ihnen klar, warum diese Beziehungen wirklich *offensichtlich* sind?

folgt

$$[f(v_1')]_{\mathcal{B}'} = [f]_{\mathcal{B}'}[v_1']_{\mathcal{B}'} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix} = [3\,v_1']_{\mathcal{B}'} \tag{85}$$

und

$$[f(v_2')]_{\mathcal{B}'} = [f]_{\mathcal{B}'}[v_2']_{\mathcal{B}'} = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} = [-v_2']_{\mathcal{B}'}\,. \tag{86}$$

Da gemäß (85) $f(v_1')$ und $3\,v_1'$ die gleichen Entwicklungskoeffizienten bezüglich der Basis $\mathcal{B}'$ haben, stimmen sie überein, und das gleiche gilt gemäß (86) für $f(v_2')$ und $-v_2'$. Es folgt also

$$f(v_1') = 3\,v_1' \qquad \text{und} \qquad f(v_2') = -v_2'\,. \tag{87}$$

Diese Beziehungen können auch in der alten Basis (d.h. unter Verwendung der ursprünglichen Definition von f, ohne dass dabei überhaupt von Basen gesprochen werden muss) nachgerechnet werden. Die entsprechenden Rechnungen sehen so aus:

$$A\,v_1' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3\,v_1' \tag{88}$$

und

$$A\,v_2' = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -\begin{pmatrix} 1 \\ -1 \end{pmatrix} = -v_2'. \tag{89}$$

Die Tatsache, dass es zwei Vektoren gibt, die von f auf Vielfache ihrer selbst abgebildet werden, ist unabhängig von der verwendeten Basis, aber sie kommt klarerweise in der neuen Basis transparenter zum Ausdruck als in der alten. Versuchen Sie, die Wirkung von A und den soeben besprochenen Sachverhalt in einer Zeichnung zu illustrieren!

Einige Ergänzungen

Dualraum

Ist V ein Vektorraum über $\mathbb{K}$, so nennen wir eine lineare Abbildung $f: V \to \mathbb{K}$ ein *lineares Funktional* auf V (oder eine *Linearform* auf V). Dabei ist $\mathbb{K}$ als (eindimensionaler) Vektorraum über sich selbst aufgefasst.

Beispiel:

$$\begin{aligned} &f: \mathbb{R}^3 \to \mathbb{R} \\ &f(x_1, x_2, x_3) = x_1 - 2x_2 + 3x_3\,. \end{aligned} \tag{90}$$

$\mathrm{Kern}(f)$ ist die Menge aller reellen Zahlentripel (x_1, x_2, x_3), für die

$$x_1 - 2x_2 + 3x_3 = 0 \tag{91}$$

gilt, also – geometrisch ausgedrückt – eine Ebene im $\mathbb{R}^3$ durch den Ursprung. So gesehen hatten Sie also bereits im Mathematikunterricht mit dem Kern linearer Funktionale zu tun! $\mathrm{Bild}(f)$ ist natürlich ganz $\mathbb{R}$.

Die Menge aller linearen Funktionale auf V heißt *Dualraum* von V und wird mit dem Symbol V^* bezeichnet. Da $V^* = \mathrm{Hom}(V, \mathbb{K})$, ist V^* ebenfalls ein Vektorraum über $\mathbb{K}$.

Satz: Sei V ein endlichdimensionaler Vektorraum über $\mathbb{K}$. Dann gilt $\dim(V^*) = \dim(V)$.

> Beweis: Mit $\dim(V) = n$ ist $V^* = \mathrm{Hom}(V, \mathbb{K}) \cong M(1 \times n, \mathbb{K})$ (siehe die obige Anmerkung zu den Seiten 91/94) und somit $\dim(V^*) = \dim(M(1 \times n, \mathbb{K})) = n$.

Der Identifizierung $V^* \cong M(1 \times n, \mathbb{K})$ muss die Wahl einer Basis $\mathcal{B}$ von V vorausgegangen sein[8]. Die Matrix $[f]_\mathcal{B}$ eines linearen Funktionals $f \in V^*$ bezüglich $\mathcal{B}$ ist als Element von $M(1 \times n)$ eine *einzeilige Matrix* (kurz: ein *Zeilenvektor*).

Beispiel: Die Abbildung

$$\begin{aligned} &g : \mathcal{P}_2 \to \mathbb{R} \\ &g : \big(\text{Polynomfunktion } t \mapsto a_0 + a_1 t + a_2 t^2\big) \ \longmapsto \ a_0 - 2a_1 + 3a_2 \end{aligned} \tag{92}$$

ist ein lineares Funktional auf $\mathcal{P}_2$. Ihre Matrix bezüglich der Basis

$$\mathcal{B} = \big(t \mapsto 1, t \mapsto t, t \mapsto t^2\big) \tag{93}$$

von $\mathcal{P}_2$ ist $(1 \ \ -2 \ \ 3) \in M(1 \times 3, \mathbb{R})$. Durch die Brille dieser Basis betrachtet, entspricht (92) der Wirkung

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \ \longmapsto \ (1 \ \ -2 \ \ 3) \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = a_0 - 2a_1 + 3a_2 \tag{94}$$

(beachten Sie: 1×3-Matrix mal 3×1-Matrix $= 1 \times 1$-Matrix $=$ Zahl), die ihrerseits mit (90) identisch ist! Lassen Sie sich von den unterschiedlichen Bezeichnungen und Schreibweisen nicht verwirren! (94) beschreibt tatsächlich *dieselbe* Abbildung $\mathbb{R}^3 \to \mathbb{R}$ wie (90)!

Ein **zweites Beispiel** für ein lineares Funktional auf $\mathcal{P}_2$:

$$\begin{aligned} &h : \mathcal{P}_2 \to \mathbb{R} \\ &h : p \ \longmapsto \ \int_0^1 p(t)\, dt. \end{aligned} \tag{95}$$

Durch die Brille einer Basis betrachtet, wird ein lineares Funktional also durch einen Zeilenvektor dargestellt, und umgekehrt definiert jeder solche Zeilenvektor (genau) ein lineares Funktional.

Da V und V^* die gleiche Dimension besitzen, sind sie zueinander isomorph. Unter den (unendlich vielen) Isomorphismen $V \to V^*$ haben manche in Bezug auf eine vorab gewählte Basis von V eine besondere Struktur: Jede Basis $\mathcal{B} = (v_1, \dots v_n)$ von V gibt Anlass zu einer Basis

[8] Wenn Sie jetzt ganz pingelig sein wollen, können Sie fragen, welche Basis man im Zielraum $\mathbb{K}$ wählt. Dreimal dürfen Sie raten! Richtig: die Zahl 1. Sie entspricht der „Standardbasis“ in $\mathbb{K}$.

$\mathcal{C} = (w_1, \ldots, w_n)$ von V^* (genannt die zu $\mathcal{B}$ *duale Basis*), indem für jedes j die Wirkung von w_j auf jeden Basisvektor v_k von V durch $w_j(v_k) = \delta_{jk}$ festgelegt wird. Damit wird

$$\begin{aligned} w_j &: V \to \mathbb{K} \\ w_j &\left(\sum_{k=1}^{n} \lambda_k\, v_k\right) = \lambda_j. \end{aligned} \tag{96}$$

In Worten: w_j ordnet jedem Vektor $x \in V$ seinen j-ten Entwicklungskoeffizienten λ_j bezüglich der Basis $\mathcal{B}$ zu. Die lineare Abbildung $V \to V^*$, die jedem Basisvektor v_j das lineare Funktional w_j zuordnet, ist ein Isomorphismus. (In den Übungen werden Sie sich das alles genauer ansehen!) Im allgemeinen Fall ist keiner der so konstruierten Isomorphismen vor allen anderen ausgezeichnet.

Der Dualraum des Dualraums von V, mit V^{**} bezeichnet, heißt *Bidualraum* (oder *Doppeldualraum*) von V. Er besitzt natürlich die gleiche Dimension wie V und V^*, und daher gibt es ebenfalls unendlich viele Isomorphismen $V \to V^{**}$. Aber einer von diesen ist in besonderer Weise ausgezeichnet – er erlaubt eine „natürliche Identifizierung" von V^{**} mit V. Dieser „natürliche" (oder „kanonische") Isomorphismus $\Psi : V \to V^{**}$ wird folgendermaßen konstruiert: Für jedes $x \in V$ ist $\Psi(x) \in V^{**}$, also seinerseits ein lineares Funktional auf V^*, d.h. eine lineare Abbildung $\Psi(x) : V^* \to \mathbb{K}$. Deren Wirkung wird nun durch

$$\Psi(x) : f \in V^* \mapsto f(x) \in \mathbb{K} \tag{97}$$

festgelegt. Schön, nicht? Im Sinne der Identifizierung von V^{**} mit V wird das sogar manchmal in der Form $x(f) := f(x)$ hingeschrieben! Ein Vektor $x \in V$, interpretiert als Element von V^{**}, wirkt auf ein $f \in V^*$ als „Auswertung an der Stelle x". Keine Basis oder sonstigen frei wählbaren Objekte wurden dabei benötigt. In den Übungen können Sie den Beweis versuchen, dass es sich dabei tatsächlich um einen Isomorphismus handelt. Er ist an sich nicht aufwändig, aber aufgrund der aufeinander aufgebauten Räume V, $V^* =$ Menge der linearen Funktionale auf V, $V^{**} =$ Menge der linearen Funktionale auf $V^* =$ Menge der linearen Funktionale auf der Menge der linearen Funktionale auf V ein bisschen abstrakt.

Was ist der Dualraum von $\mathbb{K}^n$? Natürlich wieder $\mathbb{K}^n$, zumindest in einem gewissen Sinn: Werden die Elemente von $V = \mathbb{K}^n$ als Spaltenvektoren angeschrieben, also als Matrizen in $M(n\times 1, \mathbb{K})$, so kann der Dualraum V^* mit $M(1 \times n, \mathbb{K})$ identifiziert, seine Elemente daher in Zeilenform angeschrieben werden. Schematisch, mit $x \in V$ und $f \in V^*$, sieht diese Schreibweise dann so aus:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad f = (f_1 \ \ldots \ f_n), \quad f(x) = f\,x = (f_1 \ \ldots \ f_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{j=1}^{n} f_j\, x_j\,. \tag{98}$$

Die Elemente des Bidualraums von V schreibt man dann wieder in Spaltenform an. So gesehen reduziert sich der kanonische Isomorphismus $V \to V^{**}$ für $V = \mathbb{K}^n$ auf die identische Abbildung!

Direkte Summe

Ist V ein Vektorraum über $\mathbb{K}$, und sind U und W Untervektorräume von V mit $U \cap W = \{0\}$, so nennen wir $U+W$, die Menge aller $u+w$ mit $u \in U$ und $w \in W$, die *direkte Summe* von U und W und schreiben dafür $U \oplus W$. Analog werden direkte Summen von mehr Untervektorräumen $U_1 \oplus \cdots \oplus U_r$ definiert.

Dieses Konzept kann auf Vektorräume U und W ausgedehnt werden, die nicht notwendigerweise als Untervektorräume eines Vektorraums V gegeben sind. $U \oplus W$ ist in diesem Fall definiert als das kartesische Produkt $U \times W$, ausgestattet mit der naheliegenden Vektorraumstruktur

$$(u,w) + (u',w') = (u+u', w+w') \qquad \text{für alle } u,u' \in U \text{ und } w,w' \in W \tag{99}$$

$$\lambda(u,w) = (\lambda u, \lambda w) \qquad \text{für alle } u \in U, w \in W \text{ und } \lambda \in \mathbb{K}. \tag{100}$$

Dass diese Definition mit der obigen als $U+W$ für Untervektorräume mit trivialem Durchschitt zusammenpasst, zeigen Sie in den Übungen. Eigenschaften:

- $\dim(U \oplus W) = \dim(U) + \dim(W)$
- Jedes $x \in U \oplus W$ kann *eindeutig* in eine Summe $u + w$ mit $u \in U$ und $w \in W$ „zerlegt" werden. (Wird $U \oplus W$ wie oben als $U \times W$ konstruiert, so wird $u \in U$ mit $(u, 0)$ identifiziert und $w \in W$ mit $(0, w)$.)

Beispiel: $\mathbb{R}^3 = xy$-Ebene $\oplus$ z-Achse.

Projektionen

Ist V ein Vektorraum über $\mathbb{K}$, so nennen wir eine lineare Abbildung $f : V \to V$ eine *Projektion*[9] oder einen *Projektionsoperator*, wenn

$$f \circ f = f \tag{101}$$

(Kurzform: $f^2 = f$) gilt. Überlegen Sie, warum man lineare Abbildungen dieses Typs „Projektionen" nennt! Eigenschaften:

- Für jede Projektion $f : V \to V$ gilt $V = \mathrm{Bild}(f) \oplus \mathrm{Kern}(f)$.
- $\mathrm{Bild}(f)$ besteht aus allen $x \in V$, für die $f(x) = x$ gilt.

(Beweis: Übungen.) Beispiel:

$$\begin{aligned} &f : \mathbb{R}^3 \to \mathbb{R}^3 \\ &f(x_1, x_2, x_3) = (x_1, x_2, 0) \end{aligned} \tag{102}$$

ist eine Projektion mit

$$\mathrm{Bild}(f) = \mathbb{R}^2 \times \{0\} = xy\text{-Ebene}, \tag{103}$$

$$\mathrm{Kern}(f) = \{0\} \times \{0\} \times \mathbb{R} = z\text{-Achse}. \tag{104}$$

[9] Achtung: Das Wort „Projektion" wird im Buch an mehreren Stellen in einer etwas anderen Bedeutung verwendet als hier.

In offensichtlich analoger Weise nennen wir eine quadratische Matrix A eine *Projektion(smatrix)*, wenn $A^2 = A$ gilt. Beispiel:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{105}$$

Eine lineare Abbildung $f : V \to V$ ist genau dann eine Projektion, wenn ihre Matrix $[f]_{\mathcal{B}}$ (bezüglich einer beliebigen Basis $\mathcal{B}$ und damit bezüglich jeder Basis) eine Projektionsmatrix ist. Ist $f : V \to V$ eine Projektion, dann ist auch $\mathrm{Id}_V - f$ eine Projektion. Ist A eine Projektionsmatrix, dann ist auch $\mathrm{E} - A$ eine Projektionsmatrix.

Nilpotente lineare Abildungen

Ist V ein Vektorraum über $\mathbb{K}$, so nennen wir eine lineare Abbildung $f : V \to V$ *nilpotent*, wenn eine „Potenz" von ihr verschwindet, d.h. wenn es ein $k \in \mathbb{N}\backslash\{0\}$ gibt, so dass

$$\underbrace{f \circ \cdots \circ f}_{k \text{ mal}} = 0 \tag{106}$$

ist. Analog nennen wir eine quadratische Matrix A nilpotent, wenn es ein $k \in \mathbb{N}\backslash\{0\}$ gibt, so dass $A^k = 0$ ist. Beispiel:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \tag{107}$$

(Überlegen Sie, was diese Matrix bewirkt, wenn sie, ausgehend von einem Vektor $x \in \mathbb{R}^3$, mehrere Male hintereinander angewandt wird!) Eine lineare Abbildung $f : V \to V$ ist genau dann nilpotent, wenn ihre Matrix $[f]_{\mathcal{B}}$ (bezüglich einer beliebigen Basis $\mathcal{B}$ und damit bezüglich jeder Basis) nilpotent ist. Nilpotente Matrizen spielen eine Rolle bei der Jordanschen Normalform, die später vorgestellt wird.

---

6 Die Determinante

6.1 Die Determinante

- Keine Anmerkungen zu diesem Abschnitt.

6.2 Berechnung von Determinanten

- Keine Anmerkungen zu diesem Abschnitt.

6.3 Die Determinante der transponierten Matrix

- Keine Anmerkungen zu diesem Abschnitt.

6.4 Eine Determinantenformel für die inverse Matrix

- Keine Anmerkungen zu diesem Abschnitt.

6.5 Determinante und Matrizenprodukt

- Keine Anmerkungen zu diesem Abschnitt.

6.6 Test

- Arbeiten Sie den Test durch!

6.7 Determinante eines Endomorphismus

- Keine Anmerkungen zu diesem Abschnitt.

6.8 Die Leibnizsche Formel

- Keine Anmerkungen zu diesem Abschnitt.

6.9 Historische Notiz

- Lesenswert!

Ergänzung 2: Determinante einer 3×3-Matrix und das „Epsilon-Symbol“

In physikalischen Anwendungen ist manchmal folgende Formel für die Determinante einer 3×3-Matrix nützlich:

$$\det(A) = \varepsilon_{jkl}\, a_{1j}\, a_{2k}\, a_{3l}\,, \tag{108}$$

wobei die Einsteinsche Summenkonvention benutzt wurde. Werden alle drei Summen (d.h. die Summen über j, k und l) ausgeführt, so ergibt sich eine explizite Formel, die sogenannte *Regel von Sarrus*:

$$\det(A) = a_{11}\, a_{22}\, a_{33} + a_{12}\, a_{23}\, a_{31} + a_{13}\, a_{21}\, a_{32} - a_{13}\, a_{22}\, a_{31} - a_{12}\, a_{21}\, a_{33} - a_{11}\, a_{23}\, a_{32}\,. \tag{109}$$

Jeder Summand entspricht einer nichtverschwindenden Komponente des Epsilon-Symbols.

Das Epsilon-Symbol kann auf beliebige Dimensionen verallgemeinert werden ($\varepsilon_{j_1 j_2 \dots j_n}$ ist total antisymmetrisch und $\varepsilon_{12\dots n} = 1$), und auch die obige Formel (108) für die Determinante überträgt sich zwanglos auf $n \times n$-Matrizen.

Ein Wort noch zur Determinante einer 3×3-Matrix: Wie im Buch auf Seite 72, Formel (5″), angegeben, ist die Determinante einer 3×3-Matrix das *Spatprodukt* ihrer Zeilenvektoren (und analog das Spatprodukt ihrer Spaltenvektoren). Sie ist gleich dem *orientierten Volumen* des von den Zeilen (oder Spalten) im $\mathbb{R}^3$ aufgespannten Parallelepipeds: Bilden die Zeilen (oder Spalten) eine rechtshändige Basis, so ist es positiv. Bilden sie eine linkshändige Basis, so ist es negativ. Bilden sie keine Basis, d.h. sind sie linear abhängig, so ist es 0.

7 Lineare Gleichungssysteme

7.1 Lineare Gleichungssysteme

- **Seite 159** (Bemerkung 1): Der Beweis beginnt mit den Worten „Die Spalten von A erzeugen $\operatorname{Bild} A$“, was uns bereits bekannt ist. Es gibt eine nützliche Form, das Gleichungssystem $Ax = b$ anzuschreiben, die diesen Sachverhalt noch einmal verdeutlicht und eine andere Sichtweise erlaubt, was eine Lösung eines solchen Gleichungssystems ist: Vergewissern Sie sich zuerst, dass

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \tag{110}$$

gilt! Die Aussage $Ax = b$ für ein gegebenes $b \in \mathbb{K}^m$ ist daher gleichbedeutend mit der Aussage, dass b eine Linearkombination der Spalten von A ist mit den x_j als Koeffizienten:

$$x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \tag{111}$$

Somit ist dieses Gleichungssystem genau dann lösbar, wenn b in der linearen Hülle der Spalten von A liegt. Damit ist auch Bemerkung 1 bewiesen.

- **Seite 160** (Bemerkung 2): Der hier formulierte Sachverhalt wird oft so ausgedrückt: Die *allgemeine* Lösung des linearen Gleichungssystems $Ax = b$ ist von der Form

$$x = x_0 + x_{\mathrm{hom}}\,, \tag{112}$$

wobei x_0 eine (und zwar *irgendeine*) *spezielle* Lösung ist und x_{hom} für die allgemeine Lösung des zugehörigen homogenen Gleichungssystems $Ax = 0$ steht[10].

Ist $b \neq 0$, und ist die Lösungsmenge des Gleichungssystems $Ax = b$ nicht leer, so ist $0 \notin \mathrm{Lös}(A, b)$, daher $\mathrm{Lös}(A, b)$ kein Untervektorraum von $\mathbb{K}^n$. Was aber dann? $\mathrm{Lös}(A, b)$ ist ein Element des Quotientenvektorraums $\mathbb{K}^n/\mathrm{Kern}(A)$, d.h. eine *Nebenklasse*, bildlich gesprochen eine um eine (beliebige) Lösung x_0 verschobene Version des Untervektorraums $\mathrm{Kern}(A)$! Spezialfälle sind:

  - Ist $\dim(\mathrm{Kern}(A)) = 1$, so handelt es sich um eine „Gerade“ (im $\mathbb{K}^n$), die am Ursprung vorbeizielt.
  - Ist $\dim(\mathrm{Kern}(A)) = n - 1$, so handelt es sich um eine „Hyperebene“ (im $\mathbb{K}^n$), auf der der Ursprung nicht liegt.

[10] Eine ganz ähnliche Formulierung wird Ihnen in anderen Lehrveranstaltungen im Zusammenhang mit linearen inhomogenen Differentialgleichungen begegnen.

Machen Sie sich in diesem Zusammenhang klar, dass die Parameterdarstellung der Geraden, die Sie aus dem Mathematikunterricht kennen, auf einer ganz analogen Situation beruht: x_0 spielt die Rolle eines Punktes auf der Geraden, $\mathrm{Kern}(A)$ spielt die Rolle der Menge aller Vielfachen eines Richtungsvektors.

7.2 Die Cramersche Regel

- Keine Anmerkungen zu diesem Abschnitt.

7.3 Der Gaußsche Algorithmus

- Keine Anmerkungen zu diesem Abschnitt.

7.4 Test

- Arbeiten Sie den Test durch!

7.5 Mehr über lineare Gleichungssysteme

- Keine Anmerkungen zu diesem Abschnitt.

7.6 Wiegen mit der Kamera

- Dieser Abschnitt wird in der Vorlesung übersprungen.

7.7 Historische Notiz

- Lesenswert!

---

8 Euklidische Vektorräume

8.1 Skalarprodukte

- Die einleitenden Definitionen dieses Abschnitts sollten Ihnen bereits vom Abschnitt 2.6 bekannt sein.

8.2 Orthogonale Vektoren

- **Seite 183**: Die Begriffe Orthonormalsystem und Orthonormalbasis werden oft als *ON-System* und *ON-Basis* abgekürzt.
- **Seite 184** (Lemma 3): Es folgt unmittelbar: Ist U ein Untervektorraum des endlichdimensionalen euklidischen Vektorraums V, so ist $V = U \oplus U^\perp$.

- **Seite 185/186**: Generell heißt jede lineare Abbildung $f : V \to V$, die eine Projektion ist, und für die jedes Element von $\mathrm{Bild}(f)$ orthogonal zu jedem Element von $\mathrm{Kern}(f)$ ist, eine *Orthogonalprojektion*. Wir wissen bereits, dass für Projektionen $V = \mathrm{Bild}(f) \oplus \mathrm{Kern}(f)$ gilt. Ist f eine Orthogonalprojektion, so handelt es sich bei $\mathrm{Bild}(f) \oplus \mathrm{Kern}(f)$ sogar um eine *orthogonale direkte Summe*, d.h. es gilt zusätzlich $\mathrm{Bild}(f) \perp \mathrm{Kern}(f)$. Ist $f : V \to V$ eine Orthogonalprojektion, dann ist auch $\mathrm{Id}_V - f$ eine Orthogonalprojektion.

  Ist U ein Untervektorraum von V, so gibt es genau eine Orthogonalprojektion f, deren Bild U ist. (Man nennt f dann salopp die Orthogonalprojektion *auf* U.) Mit Hilfe der (nach Lemma 3 auf Seite 184 eindeutigen) Zerlegung jedes Elements $v \in V$ in einen Anteil u in U („parallel zu U") und einen Anteil w orthogonal zu U

  $$v = \underbrace{u}_{\in U} + \underbrace{w}_{\perp U} \tag{113}$$

  kann die Wirkung von f durch $f(v) = u$ charakterisiert werden. u wird oft in der Form $v^{\|}$, w in der Form $v^{\perp}$ geschrieben, also

  $$v = \underbrace{v^{\|}}_{\in U} + \underbrace{v^{\perp}}_{\perp U} \qquad \text{und daher} \qquad f(v) = v^{\|}\,. \tag{114}$$

  Die Wirkung von f besteht also darin, den Anteil orthogonal zu U „abzuzwicken". Ist $(u_1, \ldots, u_r)$ eine Orthonormalbasis von U, so gilt nach Lemma 3

  $$v^{\|} = \sum_{j=1}^{r} \langle v, u_j \rangle u_j \tag{115}$$

  $$v^{\perp} = v - \sum_{j=1}^{r} \langle v, u_j \rangle u_j\,. \tag{116}$$

  Versuchen Sie, sich diese beiden – in Anwendungen sehr praktischen – Formeln zu merken! Eine andere Sichtweise auf den gleichen Sachverhalt ergibt sich, indem die Orthonormalbasis $(u_1, \ldots, u_r)$ von U zu einer Orthonormalbasis $(u_1, \ldots u_r, v_{r+1}, \ldots, v_n)$ von V vervollständigt wird. Dann ist

  $$v = \underbrace{\sum_{j=1}^{r} \langle v, u_j \rangle u_j}_{v^{\|} = f(v)} + \underbrace{\sum_{j=r+1}^{n} \langle v, v_j \rangle v_j}_{v^{\perp} = v - f(v)}\,. \tag{117}$$

  Klarerweise ist $(v_{r+1}, \ldots, v_n)$ eine Orthonormalbasis von $U^{\perp}$, und es gilt $\mathrm{Kern}(f) = U^{\perp}$. Wird anstelle von f die Orthogonalprojektion $g = \mathrm{Id}_V - f$ betrachtet, so vertauschen U und $U^{\perp}$ ihre Rollen, d.h. es ist $\mathrm{Bild}(g) = U^{\perp}$ und $\mathrm{Kern}(g) = U$.

  Achtung: Der Begriff „Projektion" wird nichtganz einheitlich verwendet! Im Buch wird die surjektive Abbildung $P_U : V \to U$ mit $U = \mathrm{Bild}(f)$, die jedes $v \in V$ in $f(v)$ abbildet, als „Projektion" bezeichnet. Beachten Sie den subtilen Unterschied zwischen f und P_U! Siehe dazu auch Fußnote 9.

8.3 Orthogonale Abbildungen

- Keine Anmerkungen zu diesem Abschnitt.

Ergänzung 3: Drehungen und Spiegelungen im $\mathbb{R}^3$

Wir erweitern unseren Horizont nun ein bisschen, indem wir ein paar Beispiele für Drehungen und Spiegelungen im $\mathbb{R}^3$ (in dem es ja echte zweidimensionale Untervektorräume gibt) betrachten. Dabei handelt es sich in allen Fällen um orthogonale Matrizen.

- Die Matrix
$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{118}$$
beschreibt eine Drehung um die z-Achse. Bei „Aufsicht" wird um den Winkel φ im Gegenuhrzeigersinn gedreht. Mit ein bisschen Übung und Gewöhnung erkennt man an der „Blockform" von A sofort, dass sowohl die xy-Ebene als auch die z-Achse *invariante Teilräume* sind.
- Die Matrix
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{119}$$
beschreibt eine Spiegelung an der xy-Ebene. Ihre Wirkung besteht darin, die dritte Komponente eines Vektors mit -1 zu multiplizieren.
- Die Matrix
$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{120}$$
beschreibt – ja, was würden Sie sagen?
- Und die Matrix
$$A = \begin{pmatrix} \cos\varphi & 0 & -\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & \cos\varphi \end{pmatrix} \tag{121}$$
beschreibt – was? Erkennen Sie die invarianten Teilräume?
- Unser letztes Beispiel: Die Matrix
$$A = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{122}$$
beschreibt eine Spiegelung an jener Ebene durch den Ursprung, die normal zum Vektor $u = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ liegt. Wenden wir zur Probe A auf u und auf die beiden zu u orthogonalen

Vektoren $v = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ und $w = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ (die gemeinsam mit u eine Basis bilden) an, so ergibt sich $Au = -u$, $Av = v$ und $Aw = w$. Für jeden Vektor $x = au + bv + cw$ mit $a, b, c \in \mathbb{R}$ gilt daher $Ax = -au + bv + cw$.

8.4 Gruppen

- Keine Anmerkungen zu diesem Abschnitt.

8.5 Test

- Arbeiten Sie den Test durch!

Ergänzung 4: Dualraum und Skalarprodukt

Wir haben bereits erwähnt, dass der Dualraum V^* eines endlichdimensionalen Vektorraums V zwar die gleiche Dimension wie V hat, aber dass es keinen ausgezeichneten Isomorphismus $V \to V^*$ gibt. Nun, angesichts des Skalarprodukts als zusätzlicher Struktur, ändert sich dies:

Ist V ein euklidischer Vektorraum, so gibt es einen ausgezeichneten („natürlichen" oder „kanonischen") Isomorphismus $\Omega : V \to V^*$, der wie folgt definiert ist: Für jedes $y \in V$ wirkt $\Omega(y) \in V^*$ so:

$$\Omega(y) : x \in V \mapsto \langle y, x \rangle \in \mathbb{R}. \tag{123}$$

Das wird manchmal auch in der Form $\Omega(y) = \langle y, \cdot \rangle$ oder $\Omega : y \mapsto \langle y, \cdot \rangle$ angeschrieben. (Die Zuordnungsvorschrift der Abbildung $\Omega(y)$ lautet also in Worten „bilde das Skalarprodukt mit y".) Mittels Ω wird jedes $y \in V$ mit der zugehörigen Abbildung $\langle y, \cdot \rangle$ identifiziert. Damit kann der Dualraum V^* selbst in natürlicher Weise zu einem euklidischen Vektorraum gemacht werden, indem Ω zu einer orthogonalen Abbildung erklärt wird, d.h. indem für $u, v \in V^*$

$$\langle u, v \rangle := \langle \Omega^{-1}(u), \Omega^{-1}(v) \rangle \tag{124}$$

definiert wird.

Für $V = \mathbb{R}^n$ mit dem Standard-Skalarprodukt bedeutet das im Rahmen des Matrixformalismus, dass Ω Spaltenvektoren in Zeilenvektoren verwandelt. Vergleichen Sie die Schreibweise für lineare Funktionale auf V, wenn kein Skalarprodukt verwendet wird

$$f : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \longmapsto \quad (f_1 \ \dots \ f_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{j=1}^{n} f_j \, x_j \tag{125}$$

(„Zeile mal Spalte" im $\mathbb{R}^n$; die Zeile entspricht einem Element f des Dualraums V^*, die Spalte entspricht einem Element x von V) mit der Schreibweise, die durch das Standard-Skalarprodukt ermöglicht wird:

$$\Omega(y) : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto \langle \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \rangle = \sum_{j=1}^{n} y_j\, x_j \tag{126}$$

(„Spalte „mal" Spalte" im Sinn des Skalarprodukts, mit zwei Elementen x, y von V)! Der Isomorphismus Ω kann also im Rahmen des Matrixformalismus als das „Umkippen" eines Spaltenvektors gedeutet werden. Als Haarspalterei erscheint das nur, wenn man das Standard-Skalarprodukt als einzig mögliches Skalarprodukt ansieht – was es aber nicht ist!

---

9 Eigenwerte

9.1 Eigenwerte und Eigenvektoren

- **Seite 197**: Die Menge aller Eigenwerte einer linearen Abbildung f eines endlichdimensionalen Vektorraums in sich (d.h. eines Endomorphismus) heißt dessen *Spektrum*, oft mit $\mathrm{Sp}(f)$ oder $\sigma(f)$ bezeichnet. Die spätere Bezeichnung „Spektraldarstellung" leitet sich davon ab[11].

  Aufgrund der uns bereits bekannten Eigenschaften der Verwandtschaft von f mit seiner (bezüglich einer Basis $\mathcal{B}$ genommenen) Matrix $[f]_{\mathcal{B}}$ sollte klar sein: f und $[f]_{\mathcal{B}}$ haben die gleichen Eigenwerte, und die Eigenvektoren von f und $[f]_{\mathcal{B}}$ werden durch $\Phi_{\mathcal{B}}$ bzw. $\Phi_{\mathcal{B}}^{-1}$ ineinander übergeführt. Eigenwerte und Eigenvektoren eines linearen Operators (sowie die noch zu besprechende geometrische Vielfachheit der Eigenwerte und das charakteristische Polyom) können wahlweise auch anhand der Matrix von f ermittelt werden.

- **Seite 198** (Definition): Nachdem definiert wurde, wann eine lineare Abbildung $f : V \to V$ *diagonalisierbar* ist, sollte auch klar sein, wann eine quadratische Matrix diagonalisierbar ist, denn eine $n \times n$-Matrix stellt eine lineare Abbildung $\mathbb{K}^n \to \mathbb{K}^n$ dar. Wenn Sie jetzt fragen, was die Matrix dieser linearen Abbildung, also was „die Matrix einer Matrix A" ist, so lautet die Antwort: Sie ist gleich sie selbst (also A), *sofern* die Standardbasis des $\mathbb{K}^n$ zugrundegelegt wird. Aber bezüglich einer *anderen* Basis des $\mathbb{K}^n$ wird „die Matrix von A" mitunter ganz anders aussehen. Gelingt es Ihnen, eine Basis zu finden, bezüglich der diese „Matrix von A" eine Diagonalmatrix ist, so haben Sie A „diagonalisiert". Wir sind einer solchen Situation bereits früher beim Thema Transformationsverhalten unter Basiswechsel begegnet, siehe (76) – (89).

  Daher ist eine lineare Abbildung $f : V \to V$ genau dann diagonalisierbar, wenn ihre Matrix $[f]_{\mathcal{B}}$ (bezüglich einer beliebigen Basis $\mathcal{B}$ und damit bezüglich jeder Basis) diagonalisierbar ist. Da wir die Transformationseigenschaften der Matrix einer linearen Abbildung schon kennen, gilt: Eine (reelle/komplexe) $n \times n$-Matrix A ist genau dann (über $\mathbb{R}$/über $\mathbb{C}$) diagonalisierbar, wenn es eine (reelle/komplexe) invertierbare $n \times n$-Matrix S gibt, für die $S^{-1} A\, S$ eine Diagonalmatrix ist.

[11] Im Unendlichdimensionalen wird die Menge aller Eigenwerte als *Punktspektrum* bezeichnet.

- **Seite 199** (Definition): Der Eigenraum zum Eigenwert λ besteht aus allen Eigenvektoren zum Eigenwert λ *und* dem Nullvektor (der ja kein Eigenvektor sein kann)!

  Ist der Eigenraum zum Eigenwert λ mehrdimensional (also seine geometrische Vielfachheit größer als 1), so nennt man den Eigenwert λ *entartet*. Ein für die Physik wichtiges Konzept, dem Sie in der Quantenmechanik wieder begegnen werden!

9.2 Das charakteristische Polynom

- **Seite 202**: Das auf dieser Seite gezeigte kommutative Diagramm sollte Ihnen nun schon bestens vertraut sein. Die Matrix A nennen wir in diesem Skriptum auch $[f]_{\mathcal{B}}$, wobei $\mathcal{B}$ jene Basis von V ist, auf die sich der kanonische Basisisomorphismus $\Phi_{\mathcal{B}}$ (im Diagramm einfach Φ genannt) bezieht.

- **Seite 203/204**: Aus dem Fundamentalsatz der Algebra folgt, dass jedes komplexe Polynom vom Grad $n \geq 1$ in Linearfaktoren zerlegt werden kann:

$$P(z) = c\,(z - z_1)(z - z_2)\dots(z - z_n) \qquad \text{mit} \;\; c \neq 0, \tag{127}$$

  wobei nicht alle z_j voneinander verschieden sein müssen. Ist P das charakteristische Polynom einer linearen Abbildung f eines *komplexen* Vektorraums in sich oder einer *komplexen* quadratischen Matrix A, so sind die z_j die Eigenwerte. Jeder Eigenwert kommt mindestens so oft vor, wie es seiner geometrischen Vielfachheit entspricht, unter Umständen aber auch öfter. (Man spricht dann auch von der *algebraischen Vielfachheit* eines Eigenwerts – sie ist stets größer-gleich seiner geometrischen Vielfachheit). Es gilt:

  - Die Determinante von f oder A ist gleich dem Produkt der Eigenwerte, wobei jeder Eigenwert so oft gezählt wird, wie er in der Zerlegung des charakteristischen Polynoms in Linearfaktoren auftritt (also entsprechend seiner algebraischen Vielfachheit).
  - Die Spur von f oder A ist gleich der Summe der Eigenwerte, wobei jeder Eigenwert so oft gezählt wird, wie er in der Zerlegung des charakteristischen Polynoms in Linearfaktoren auftritt (also entsprechend seiner algebraischen Vielfachheit).

  Für lineare Abbildungen in *reellen* Vektorräumen und für Matrizen, die als lineare Abbildungen $\mathbb{R}^n \to \mathbb{R}^n$ aufgefasst werden, also für $\mathbb{K} = \mathbb{R}$, gelten diese zwei Aussagen nur dann, wenn das charakteristische Polynom über $\mathbb{R}$ in Linearfaktoren zerlegt werden kann. Ist das nicht der Fall, so gibt es echt komplexe Eigenwerte, die im Reellen überhaupt nicht auftreten (Beispiel: $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, daher $P_A(\lambda) = \lambda^2 + 1$.) Aus diesem Grund werden Matrizen mit reellen Koeffizienten oft als komplexe Matrizen aufgefasst, d.h. über $\mathbb{C}$ betrachtet. (Über $\mathbb{C}$ faktorisiert λ^2+1 in $(\lambda+i)(\lambda-i)$, daher besitzt die obige Matrix A, wenn $\mathbb{K} = \mathbb{C}$ zugrundegelegt wird, die Eigenwerte $\pm i$.) Wenn Sie eine Übungsaufgabe der Form „Berechne die Eigenwerte der Matrix A!“ (ohne genauere Angaben über den Grundkörper) vor sich haben, so berechnen sie am besten *alle* (komplexen) Eigenwerte, auch wenn die Matrix nur reelle Koeffizienten hat!

- Interessanter Nachtrag zu diesem Abschnitt: Ist P_f das charakteristische Polynom von f, so gilt $P_f(f) = 0$ (*Satz von Cayley-Hamilton*). Beispiel: Ist das charakteristische Polynom durch $P_f(\lambda) = 4 - 5\lambda + \lambda^2$ gegeben, so gilt $4\,\mathrm{Id} - 5f + f \circ f = 0$. Analoges gilt für das charakteristische Polynom einer quadratischen Matrix A: Ist das charakteristische Polynom durch $P_A(\lambda) = 4 - 5\lambda + \lambda^2$ gegeben, so gilt $4E - 5A + A^2 = 0$. Für diagonalisierbare Operatoren bzw. Matrizen ist der Satz leicht zu zeigen (wie?), aber er gilt ganz allgemein.

9.3 Test

- Arbeiten Sie den Test durch!

9.4 Polynome

- Dieser Abschnitt wird in der Vorlesung übersprungen.

---

10 Die Hauptachsen-Transformation

10.1 Selbstadjungierte Endomorphismen

- **Seite 213** (Definition): Nachtrag, da jetzt der Begriff der Selbstadjungiertheit zur Verfügung steht: Ist V ein euklidischer Vektorraum und $f : V \to V$ eine Projektion, so ist f genau dann eine Orthogonalprojektion, wenn f selbstadjungiert ist. Die Orthogonalprojektionen sind also genau die selbstadjungierten Projektionen.

- **Seite 215**: Die Formel $f(v_j) = \sum_{i=1}^{n} a_{ij} v_i$ sollte Ihnen bekannt vorkommen: Es handelt sich genau um (48) bzw. (52)!

10.2 Symmetrische Matrizen

- Keine Anmerkungen zu diesem Abschnitt.

10.3 Die Hauptachsen-Transformation für selbstadjungierte Endomorphismen

- **Seite 219**: Das im ersten Korollar auftretende Produkt $P^{-1}AP$ sollte Ihnen bereits vom allgemeinen Transformationsverhalten (67) der Matrix einer linearen Abbbildung unter Basiswechsel bekannt sein. In diesem allgemeinen Zusammenhang haben wir die Matrix, die hier P heißt, S genannt.

  Zu der im zweiten Korollar angegebenen Form der Spektraldarstellung selbstadjungierter Operatoren wäre noch hervorzuheben, dass die Orthogonalprojektionen P_k, die als

Projektionen natürlich ${P_k}^2 = P_k$ erfüllen (wir lassen nun konsequent das Zeichen $\circ$ für die Vekettung weg), untereinander durch die Beziehungen $P_k P_l = 0$ für $k \neq l$ verknüpft sind. (Eine weitere Beziehung wird weiter unten, in (130), nachgeliefert.) Das hat nun eine faszinierende und nützliche Konsequenz: Führen wir f zweimal hintereinander aus und benutzen ${P_k}^2 = P_k$ sowie $P_k P_l = 0$ für $k \neq l$, so ergibt sich

$$f^2 = \sum_{k,l=1}^{r} \lambda_k \, \lambda_l \, P_k P_l = \sum_{k=1}^{r} {\lambda_k}^2 \, P_k \, , \tag{128}$$

was sich von der Spektraldarstellung von f nur dadurch unterscheidet, dass die Eigenwerte λ_k durch ${\lambda_k}^2$ ersetzt worden sind. Das Gleiche können Sie mit jeder Potenz machen, und ganz allgemein gilt für jede Polynomfunktion F

$$F(f) = \sum_{k=1}^{r} F(\lambda_k) \, P_k \, . \tag{129}$$

Lässt sich das auf weitere Typen von Funktionen verallgemeinern? Man könnte ausprobieren, ob für F eine Taylorreihe gewählt werden kann (das wäre im Sinn der Analysis eine naheliegende Verallgemeinerung des Polynombegriffs), aber um uns die damit verbundenen Umstände zu ersparen, erheben wir (129) zur *Definition* von $F(f)$ für *beliebige* Funktionen $F : \mathrm{Sp}(f) \to \mathbb{R}$. Beachten Sie, dass F tatsächlich nur auf die Eigenwerte angewandt werden muss! Beispiele:

- Sind alle $\lambda_k \neq 0$, so können wir $F(\lambda) = \dfrac{1}{\lambda}$ wählen. Überrascht, dass die so definierte lineare Abbildung $\dfrac{1}{f}$ gleich der Inversen f^{-1} ist? (Übungen!)
- Mit $F(\lambda) = 1$ ergibt sich die sogenannte *Vollständigkeitsrelation*

$$\sum_{k=1}^{r} P_k = \mathrm{Id}_V \, . \tag{130}$$

Sie stellt sicher, dass in der Spektraldarstellung von f auch wirklich *alle* Eigenräume berücksichtigt sind und es daher eine Basis von V gibt, die nur aus Eigenvektoren von f besteht (das Kriterium der Diagonalisierbarkeit). Man kann (130) aber auch ganz anders zeigen – am einfachsten bezüglich einer Basis, in der f diagonal ist. Das machen Sie in den Übungen. Hier nur ein dreidimensionales Beispiel: Die Eigenwerte von f seien $\lambda_1 = -1$ (mit geometrischer Vielfachheit 1) und $\lambda_2 = 3$ (mit geometrischer Vielfachheit 2). Bezüglich einer geeigneten Orthonormalbasis $\mathcal{B}$ ist

$$[f]_{\mathcal{B}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = -1 \, \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{[P_1]_{\mathcal{B}}} + 3 \, \underbrace{\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{[P_2]_{\mathcal{B}}} . \tag{131}$$

Klarerweise folgt $[P_1]_{\mathcal{B}} + [P_2]_{\mathcal{B}} = E_3$, daher $P_1 + P_2 = \mathrm{Id}_V$.

- In der Physik besonders bedeutsam:

$$e^{af} = \sum_{k=1}^{r} e^{a\,\lambda_k}\, P_k\,. \tag{132}$$

Ganz analog kann vorgegangen werden, wenn beispielsweise (nur so zum Spaß) $\sin(A)$ für eine reelle symmetrische 3×3-Matrix A berechnet werden soll. Ist A eine Diagonalmatrix, so wird $F(A)$ in schönem Einklang mit dem bisher Gesagten einfach durch Anwendung von F auf die Diagonalelemente von A (die ja die Eigenwerte sind) ermittelt.

- **Seite 220/221** (Rezept): Hier wird eine Matrix P, für die $P^{-1}AP$ diagonal ist, und die zusätzlich orthogonal ist, eine *Hauptachsen-Transformation* für A genannt. Die Forderung, dass P eine orthogonale Matrix sein soll, rührt daher, dass die zugehörige Basis aus Eigenvektoren eine Orthonormalbasis sein soll. Um eine solche Matrix P zu finden, gehen Sie wie angegeben vor! Einige Anmerkungen dazu:

  - In Schritt 2 können Sie natürlich auch jedes andere Verfahren anwenden, um den Eigenraum zu ermitteln. Verstehen Sie die Erwähnung des Gaußschen Verfahrens lediglich als Vorschlag!
  - Schritt 3 dient dazu, die Orthonormalität der gesuchten Basis sicherzustellen. Ist der Eigenraum eindimensional, so besteht die ganze Arbeit darin, einen Eigenvektor zu normieren. Ist der Eigenraum mehrdimensional, so müssen Sie *irgendeine* Orthonormalbasis für ihn ermitteln, wie auch immer Sie das anstellen! Verstehen Sie daher auch die Erwähnung des Schmidtschen Orthonormalisierungsverfahrens nur als Vorschlag!
  - Hat man eine solche Basis einmal ermittelt, so drückt Schritt 4 nichts anderes aus als unser früheres (mit der Bezeichnung S statt P formuliertes) Ergebnis (75)!

Betrachten wir dazu ein konkretes

Beispiel:
Um eine Hauptachsentransformation für

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \tag{133}$$

zu finden, werden zuerst die Eigenwerte berechnet. Ergebnis: 2 und 7. Danach werden normierte Eigenvektoren zu den beiden Eigenwerten bestimmt. Ergebnisse (bis auf Vorzeichen eindeutig):

$$\text{EV zum EW 2} = \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}, \quad \text{EV zum EW 7} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix}. \tag{134}$$

Da die Eigenwerte voneinander verschieden sind, sind die Eigenräume eindimensional. Daher ist automatisch sichergestellt, dass die Eigenvektoren zueinander orthogonal sind (was sich mit den obigen Vektoren leicht verifizieren lässt). Da sie auch normiert sind, bilden sie eine Orthonormalbasis $\mathcal{B}$ von $\mathbb{R}^2$. Die Matrix P wird nun mit diesen beiden Vektoren als Spalten zusammengesetzt:

$$P = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \equiv \frac{1}{\sqrt{5}} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}. \tag{135}$$

P, die gesuchte Hauptachsentransformation, ist automatisch eine orthogonale Matrix. Wie leicht nachzurechnen ist, gilt

$$P^{-1}AP = \begin{pmatrix} 2 & 0 \\ 0 & 7 \end{pmatrix}. \tag{136}$$

Letzteres ist klarerweise $[A]_{\mathcal{B}}$.

- Ergänzung: **Diagonalisieren einer beliebigen diagonalisierbaren Matrix**
Wir wollen an dieser Stelle noch kurz auf das allgemeine Diagonalisierungsverfahren für eine diagonalisierbare (nicht notwendigerweise reelle symmetrische) Matrix A eingehen: Sei A eine über $\mathbb{K}$ diagonalisierbare $n \times n$-Matrix. Ist $\mathcal{B} = (v_1, \dots v_n)$ eine Basis von $\mathbb{K}^n$ aus Eigenvektoren von A zu den Eigenwerten $\lambda_1, \dots, \lambda_n$ (also $Av_k = \lambda_k v_k$), so setzen wir die Matrix S aus den als Spalten angeschriebenen Eigenvektoren zusammen, also $S_{jk} = (v_k)_j$. Aufgrund der Basiseigenschaft der v_k ist S automatisch invertierbar. Damit ist

$$(AS)_{jk} = \sum_{l=1}^{n} A_{jl} \underbrace{S_{lk}}_{(v_k)_l} = (Av_k)_j = \lambda_k (v_k)_j = \lambda_k S_{jk} \tag{137}$$

und daher

$$\begin{aligned}(S^{-1}AS)_{rk} &= \sum_{j=1}^{n} (S^{-1})_{rj} (AS)_{jk} = \sum_{j=1}^{n} (S^{-1})_{rj} \lambda_k S_{jk} = \\ &= \lambda_k \sum_{j=1}^{n} (S^{-1})_{rj} S_{jk} = \lambda_k \delta_{rk}\,. \end{aligned} \tag{138}$$

Somit ist $S^{-1}AS$ eine Diagonalmatrix mit den Eigenwerten von A auf der Hauptdiagonale:

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} \equiv [A]_{\mathcal{B}}. \tag{139}$$

Gleichzeitig beschreibt die Matrix S (wie früher beim Transformationsverhalten der Matrix einer linearen Abbildung unter Basiswechsel besprochen) den Übergang von der Standardbasis des $\mathbb{K}^n$ zu der aus den Eigenvektoren von A gebildeten Basis $\mathcal{B}$.

10.4 Test

- Arbeiten Sie den Test durch!

11 Klassifikation von Matrizen

11.1 Was heißt „Klassifizieren"?

- Dieser Abschnitt wird in der Vorlesung übersprungen.

11.2 Der Rangsatz

- Dieser Abschnitt wird in der Vorlesung übersprungen.

11.3 Die Jordansche Normalform

- **Seite 234**: Aus der Existenz der Jordanschen Normalform geht hervor, dass jede komplexe $n \times n$-Matrix die Summe aus einer diagonalisierbaren Matrix und einer nilpotenten Matrix ist. Weiters gilt, dass die geometrische Vielfachheit eines Eigenwerts gleich der Anzahl der zu diesem Eigenwert gehörenden Jordan-Kästchen ist.

11.4 Nochmals die Hauptachsentransformation

- Keine Anmerkungen zu diesem Abschnitt.

11.5 Der Sylvestersche Trägheitssatz

- Dieser (im Prinzip für PhysikerInnen relevante und lesenwerte) Abschnitt muss in der Vorlesung leider ebenfalls übersprungen werden.

11.6 Test

- Arbeiten Sie den Test durch!

Ergänzung 5: Unitäre Vektorräume

Wir kommen nun noch zu einigen Inhalten, die zum Vorlesungsstoff gehören, aber nicht im Buch enthalten sind. Wie Ihnen vielleicht aufgefallen ist, werden manche der zentralen Themen des Buchs (Skalarprodukt, Orthogonalität, Drehungen/Spiegelungen und Hauptachsen-Transformation) nur für reelle Vektorräume vorgestellt. In der Physik (insbesondere in der Quantentheorie) werden die Verallgemeinerungen dieser Begriffe für komplexe Vektorräume benötigt. Diesen wenden wir uns nun zu.

Komplexe Zahlen

Von den komplexen Zahlen werden wir neben dem bereits im Buch Gesagten nur wenig benötigen:

- Ist $z = x + iy \in \mathbb{C}$ (mit $x, y \in \mathbb{R}$), so ist $\overline{z} := x - iy$ die zu z *komplex konjugierte* Zahl.
- Der *(Absolut)Betrag* einer komplexen Zahl $z = x + iy$ ist definiert durch $|z| := \sqrt{x^2 + y^2} \geq 0$. Es gilt $\overline{z}\, z = |z|^2$.
- An zwei Stellen wird „e hoch einer imaginären Zahl“ vorkommen. Die komplexe Exponentialfunktion (allgemein e^z für $z \in \mathbb{C}$) erfüllt die gleichen Rechenregeln wie die reelle Exponentialfunktion, und für ein imaginäres z, d.h. $z = i\,y$ mit $y \in \mathbb{R}$, ist sie durch die schöne (*Eulersche*) Formel $e^{iy} = \cos y + i\, \sin y$ gegeben. Die komplexen Zahlen dieser Form sind genau jene, deren Betrag gleich 1 ist, d.h. die in der Gaußschen Zahlenebene auf dem Einheitskreis liegen.

Unitäre Vektorräume

Definition: Sei V ein komplexer Vektorraum. Ein *inneres Produkt*[12] in V ist eine Abbildung

$$\begin{aligned} &V \times V \to \mathbb{C} \\ &(x, y) \mapsto \langle x, y \rangle \end{aligned} \tag{140}$$

mit den folgenden Eigenschaften:

(i) Für jedes $x \in V$ ist die Abbildung $\langle\,\cdot\,, x\rangle : V \to \mathbb{C}$, $v \mapsto \langle v, x\rangle$ *konjugiert linear*, d.h. $\langle v + u, x\rangle = \langle v, x\rangle + \langle u, x\rangle$ und $\langle \lambda v, x\rangle = \overline{\lambda}\,\langle v, x\rangle$ für alle $v, u \in V$ und $\lambda \in \mathbb{C}$.

Für jedes $x \in V$ ist die Abbildung $\langle x, \cdot\,\rangle : V \to \mathbb{C}$, $v \mapsto \langle x, v\rangle$ linear, d.h. $\langle x, v + u\rangle = \langle x, v\rangle + \langle x, u\rangle$ und $\langle x, \lambda v\rangle = \lambda\,\langle x, v\rangle$ für alle $v, u \in V$ und $\lambda \in \mathbb{C}$.

(ii) $\langle y, x\rangle = \overline{\langle x, y\rangle}$ für alle $x, y \in V$.

(iii) $\langle x, x\rangle > 0$ für alle $x \neq 0$.

[12] Der Ausdruck *Skalarprodukt* wird für komplexe Vektorräume eher selten verwendet.

Aufgrund von (ii) hätte in (i) auch *eine* der beiden Bedingungen genügt. Ein inneres Produkt ist linear in einer und konjugiert linear in der anderen Stelle, damit (iii) möglichst im Anklang an den reellen Fall zu realisieren ist (siehe $\mathbb{C}^n$ als Beispiel weiter unten). Der letzte Grund für diese Maßnahme liegt darin, dass das Quadrat einer komplexen Zahl nicht notwendigerweise reell ist (geschweige denn ≥ 0). Wir haben Linearität in der zweiten Stelle verlangt (wie in der Physik meist üblich). In manchen Lehrbüchern ist es umgekehrt, und es wird Linearität in der ersten Stelle gefordert.

Definition: Unter einem *unitären Vektorraum* versteht man ein Paar $(V, \langle\,\cdot\,,\,\cdot\,\rangle)$ bestehend aus einem komplexen Vektorraum V und einem inneren Produkt $\langle\,\cdot\,,\,\cdot\,\rangle$ auf V. Eine andere Bezeichnung ist *(komplexer) Vektorraum mit innerem Produkt*, wobei der Zusatz „komplex" weggelassen werden kann, wenn man die Konvention vereinbart, die Bezeichnung „inneres Produkt" nicht für ein Skalarprodukt auf einem reellen Vektorraum zu verwenden, oder kurz *unitärer Raum*. Eine weitere Bezeichnung ist *Prähilbertraum*. Ist V endlichdimensional, so kann man auch die Bezeichnung *Hilbertraum* verwenden[13].

Beispiele:

- $\mathbb{C}^n$ mit dem inneren Produkt[14]

$$\langle x, y\rangle = \sum_{j=1}^{n} \overline{x_j}\, y_j\,, \tag{141}$$

  wobei $x = (x_j) \in \mathbb{C}^n$ und $y = (y_j) \in \mathbb{C}^n$. Beachten Sie, dass ohne die komplexe Konjugation der x_j die Forderung $\langle x, x\rangle \geq 0$ nicht erzielt werden könnte, da $\langle x, x\rangle$ dann nicht einmal reell sein müsste. Mit der komplexen Konjugation ist das aber sichergestellt, denn

$$\langle x, x\rangle = \sum_{j=1}^{n} \overline{x_j}\, x_j \;\equiv\; \sum_{j=1}^{n} |x_j|^2 \geq 0\,. \tag{142}$$

  Aufgrund dieses wichtigen Unterschieds zum reellen Fall wurde gleich von vornherein verlangt, dass das innere Produkt nicht bilinear, sondern „*sesquilinear*" ist.

- $V =$ Vektorraum der stetigen Funktionen $[-1, 1] \to \mathbb{C}$ mit

$$\langle f, g\rangle = \int_{-1}^{1} \overline{f(x)}\, g(x)\, dx \tag{143}$$

  und zahlreiche Verallgemeinerungen mit anderen Integrationsbereichen und für andere Funktionstypen.

[13] Die in der Quantenphysik benötigten Hilberträume sind meist unendlichdimensional. Damit ein unendlichdimensionaler Prähilbertraum ein Hilbertraum ist, gehört noch eine zusätzliche Eigenschaft (die „*Vollständigkeit*") dazu, die außerhalb der Zuständigkeit der linearen Algebra und damit dieser Vorlesung liegt.

[14] Im Sinn von Fußnote 12 sollte man es „Standard-inneres Produkt" nennen, aber um diesen unschönen Terminus zu vermeiden, werde ich es im Folgenden einfach als „inneres Produkt in $\mathbb{C}^n$" bezeichnen.

Achtung: Für gewisse Funktionenklassen ist durch die obige Formel (143) *kein* inneres Produkt definiert. Ein solcher Fall tritt beispielsweise auf, wenn Funktionen zugelassen sind, die integrierbar sind, deren Werte aber nur auf einer endlichen Menge von Stellen $\neq 0$ sind. Integrale über derartige Funktionen verschwinden, und daher passiert es unangenehmerweise, dass $f \neq 0$, aber $\langle f, f \rangle = 0$ ist. In solchen (insbesondere für die Quantenmechanik relevanten) Fällen muss zuerst der zugrundeliegende Vektorraum umdefiniert werden. Die Elemente des reparierten Vektorraums sind dann nicht mehr Funktionen, sondern Klassen von „fast überall gleichen" Funktionen. Aber das geht meilenweit über unseren Stoff hinaus.

Im Folgenden werden wir nur mehr endlichdimensionale unitäre Vektorräume betrachten.

Einige Definitionen und Folgerungen:

- Die *Norm* von $x \in V$ ist definiert durch $\|x\| := \sqrt{\langle x, x \rangle} \geq 0$. Es gilt $\|x\| = 0 \;\Leftrightarrow\; x = 0$ und $\|\lambda x\| = |\lambda|\, \|x\|$ für alle $x \in V$ und $\lambda \in \mathbb{C}$. Weiters gelten die Cauchy-Schwarzsche Ungleichung (Seite 180) und die Dreiecksungleichung (Seite 181).

- Zwei Vektoren $x, y \in V$ heißen (zueinander) *orthogonal*, wenn $\langle x, y \rangle = 0$ gilt. Die Konzepte der Orthogonalität von Untervektorrräumen, des orthogonalen Komplements und der Orthonormalbasis werden in offensichtlicher Weise aus dem Reellen übernommen.

- Ist $(v_1, \ldots, v_n)$ eine Orthonormalbasis, so können die Entwicklungskoeffizienten c_j eines Vektors $x = \sum_{j=1}^{n} c_j\, v_j \in V$ mit Hilfe der einfachen Formel

  $$c_j = \langle v_j, x \rangle \tag{144}$$

  berechnet werden.

  Für eine Orthonormalbasis $(v_1, \ldots, v_n)$ folgt daher aus (144) die Identität

  $$\sum_{j=1}^{n} \langle v_j, x \rangle\, v_j = x \quad \text{für alle } x \in V \tag{145}$$

  (*Vollständigkeitsrelation*). Gilt umgekehrt diese Beziehung für ein n-Tupel $(v_1, \ldots, v_n)$ von Vektoren, so ist dieses eine Orthonormalbasis.

  Die zu (145) analoge Vollständigkeitsrelation in euklidischen Vektorräumen (d.h. im Reellen) wurde im Buch auf Seite 183 (Lemma 1) erwähnt (und dort „Entwicklungsformel" genannt).

- Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine Orthonormalbasis von V, so können die Komponenten der Matrix $[x]_{\mathcal{B}}$ eines Vektors $x \in V$ und der Matrix $[f]_{\mathcal{B}}$ einer linearen Abbildung $f : V \to V$ mit Hilfe der einfachen Formeln

  $$([x]_{\mathcal{B}})_j = \langle v_j, x \rangle \qquad \text{und} \qquad ([f]_{\mathcal{B}})_{jk} = \langle v_j, f(v_k) \rangle \tag{146}$$

  berechnet werden. Der euklidischen (reellen) Variante der zweiten Formel sind wir übrigens im Buch auf Seite 215 bereits begegnet. Die Formeln (144) und (146) illustrieren, dass es in unitären Vektorräumen *sehr* einfach ist, Vektoren in eine Orthonormalbasis zu entwickeln und die Matrizen linearer Abbildungen bezüglich einer Orthonormalbasis zu berechnen. Das wird in der Quantenmechanik ständig ausgenutzt.

Aber nicht nur das: Die Identifizierung zwischen Vorgängen in V mit den entsprechenden Vorgängen im $\mathbb{C}^n$ schließt nun auch das innere Produkt ein, *sofern* die Basis $\mathcal{B}$, auf die sich die Matrizen von Vektoren und linearen Abbildungen beziehen, eine Orthonormalbasis ist. Insbesondere gilt dann

$$\langle x, y\rangle = \langle [x]_{\mathcal{B}}, [y]_{\mathcal{B}}\rangle\,, \tag{147}$$

wobei $\langle\,\cdot\,,\,\cdot\,\rangle$ auf der linken Seite das innere Produkt in V bezeichnet und $\langle\,\cdot\,,\,\cdot\,\rangle$ auf der rechten Seite jenes in $\mathbb{C}^n$.

Die wichtige Beziehung (147) sagt uns, dass alle unitären Vektorräume der gleichen (endlichen) Dimension durch die Brille einer Orthonormalbasis betrachtet gleich aussehen. Für jedes n gibt es in diesem Sinn („bis auf Isomorphie") nur *einen* n-dimensionalen unitären Vektorraum!

- Das *Schmidtsche Orthonormalisierungsverfahren* in unitären Vektorräumen funktioniert wie in euklidischen Vektorräumen, wobei man nun aber bei den aufretenden inneren Produkten auf die richtige Reihenfolge der Vektoren achten muss. (Wir erinnern uns, dass $\langle x, y\rangle$ in einem unitären Vektorraum nicht unbedingt gleich $\langle y, x\rangle$ ist.) Sei also $(v_1,\dots,v_r)$ ein linear unabhängiges r-Tupel von Vektoren. Aus ihm gewinnen wir eine ON-Basis $(\widetilde{v}_1,\dots,\widetilde{v}_r)$ von $L(v_1,\dots,v_r)$ folgendermaßen:

$$\widetilde{v}_1 = \frac{v_1}{\|v_1\|} \tag{148}$$

$$\widetilde{v}_2 = \frac{v_2 - \langle\widetilde{v}_1, v_2\rangle\,\widetilde{v}_1}{\|v_2 - \langle\widetilde{v}_1, v_2\rangle\,\widetilde{v}_1\|} \quad \text{..... orthogonal zu } \widetilde{v}_1 \tag{149}$$

$$\vdots$$

$$\widetilde{v}_{k+1} = \frac{v_{k+1} - \sum_{j=1}^{k}\langle\widetilde{v}_j, v_{k+1}\rangle\,\widetilde{v}_j}{\|v_{k+1} - \sum_{j=1}^{k}\langle\widetilde{v}_j, v_{k+1}\rangle\,\widetilde{v}_j\|} \quad \text{..... orthogonal zu } \widetilde{v}_1,\dots,\widetilde{v}_k \tag{150}$$

$$\vdots$$

$$\widetilde{v}_r = \frac{v_r - \sum_{j=1}^{r-1}\langle\widetilde{v}_j, v_r\rangle\,\widetilde{v}_j}{\|v_r - \sum_{j=1}^{r-1}\langle\widetilde{v}_j, v_r\rangle\,\widetilde{v}_j\|} \quad \text{..... orthogonal zu } \widetilde{v}_1,\dots,\widetilde{v}_{r-1} \tag{151}$$

  Vergleichen Sie die Reihenfolge der Vektoren in den inneren Produkten mit jener, die für den euklidischen Fall im Buch auf Seite 185 angegeben ist, und zeigen Sie zur Übung, dass $\langle\widetilde{v}_1, \widetilde{v}_2\rangle = 0$ gilt!

- Analog zum reellen Fall kann der Dualraum V^* eines (endlichdimensionalen) unitären Vektorraums V in kanonischer (natürlicher) Weise mit V identifiziert werden, wobei allerdings die Zuordnung

$$\begin{aligned}\Omega : V &\to V^*\\ y &\mapsto \Omega(y) : x \in V \mapsto \langle y, x\rangle \in \mathbb{C}\end{aligned} \tag{152}$$

(die also, in Kurzschreibweise, $y \in V$ mit der Abbildung $\langle y, \cdot \rangle$ identifiziert) keine lineare, sondern eine konjugiert lineare Abbildung ist, denn es gilt zwar $\Omega(x+y) = \Omega(x) + \Omega(y)$, aber $\Omega(\lambda x) = \overline{\lambda}\,\Omega(x)$ für alle $x, y \in V$ und $\lambda \in \mathbb{C}$.

- Sind V und W (endlichdimensionale) unitäre Vektorräume, und ist $f : V \to W$ eine lineare Abbildung, so ist die zu f *adjungierte Abbildung* (der zu f *adjungierte Operator*) $f^\dagger : W \to V$ (eindeutig) definiert durch die Eigenschaft

$$\langle x, f(y) \rangle = \langle f^\dagger(x), y \rangle \quad \text{für alle } x \in W \text{ und } y \in V. \tag{153}$$

$f^\dagger$ wird ausgesprochen als „f Dagger“ (von *dagger* = Dolch). Analog ist die zu einer Matrix $A \in M(m \times n, \mathbb{C})$ *adjungierte Matrix* $A^\dagger \in M(n \times m, \mathbb{C})$ definiert, und es gilt

$$A^\dagger = \overline{A^t} = \overline{A}^{\,t}. \tag{154}$$

Sind in V und W Orthonormalbasen gewählt, so ist bezüglich dieser Basen die Matrix von $f^\dagger$ gleich der Adjungierten der Matrix von f (also $[f^\dagger]_{\mathcal{B},\mathcal{C}} = ([f]_{\mathcal{C},\mathcal{B}})^\dagger$). Für die Adjuktion von linearen Abbildungen gilt

$$(f+g)^\dagger = f^\dagger + g^\dagger \tag{155}$$
$$(\lambda f)^\dagger = \overline{\lambda}\, f^\dagger \qquad \text{für beliebige } \lambda \in \mathbb{C} \tag{156}$$
$$(f \circ g)^\dagger = g^\dagger \circ f^\dagger \tag{157}$$
$$(f^\dagger)^\dagger = f \tag{158}$$
$$\det(f^\dagger) = \overline{\det(f)} \tag{159}$$
$$\lambda \text{ ist Eigenwert von } f \Leftrightarrow \overline{\lambda} \text{ ist Eigenwert von } f^\dagger, \tag{160}$$

und entsprechende Beziehungen gelten für das Adjungieren von Matrizen, insbesondere $(AB)^\dagger = B^\dagger A^\dagger$ (was Sie in den Übungen beweisen).

- Mit Hilfe der Operation des Matrix-Adjungierens kann das innere Produkt in $\mathbb{C}^n$ sehr einfach dargestellt werden: Werden $x, y \in \mathbb{C}^n$ als komplexe Spaltenvektoren geschrieben, so gilt

$$\langle x, y \rangle = \sum_{j=1}^{n} \overline{x_j}\, y_j = x^\dagger y\,, \tag{161}$$

also „Zeile mal Spalte“.

- Seien V und W unitäre Vektorräume der gleichen (endlichen) Dimension. Eine lineare Abbildung $f : V \to W$ heißt *unitär* (oder *unitärer Operator*), wenn f ein Isomorphismus ist und

$$f^{-1} = f^\dagger \tag{162}$$

gilt. Unitäre Operatoren identifizieren unitäre Vektorräume der gleichen Dimension nicht nur hinsichtlich der linearen Struktur, sondern auch hinsichtlich des inneren Produkts, denn sie erfüllen

$$\langle x, f(y) \rangle = \langle f^{-1}(x), y \rangle \quad \text{für alle } x \in W \text{ und } y \in V, \tag{163}$$

was mit $z = f^{-1}(x)$ in

$$\langle f(z), f(y)\rangle = \langle z, y\rangle \quad \text{für alle } y, z \in V \tag{164}$$

umformuliert werden kann. Unitäre lineare Abbildungen auf unitären Vektorräumen spielen in etwa die gleiche Rolle wie orthogonale lineare Abbildungen in euklidischen Vektorräumen.

Eine quadratische Matrix A heißt unitär, wenn sie invertierbar ist und $A^{-1} = A^\dagger$ gilt. Die Matrix $[f]_{\mathcal{C},\mathcal{B}}$ eines unitären Operators $f : V \to W$ bezüglich orthonormaler Basen $\mathcal{B}$ und $\mathcal{C}$ ist unitär. Ein wichtiges Kriterium, das oft benutzt werden kann, um von einer gegebenen Matrix sogleich zu erkennen, ob sie unitär ist, ist dieses: Eine $n \times n$-Matrix A ist genau dann unitär, wenn ihre Spalten (oder Zeilen) eine Orthonormalbasis des $\mathbb{C}^n$ bilden.

- Mit dem Begriff der unitären Matrix können wir nun einen Nachtrag zum Transformationsverhalten der Matrix einer linearen Abbildung unter Basiswechsel machen. Wir beschränken uns dabei auf den Fall $W = V$. Sei also $f : V \to V$ eine beliebige lineare Abbildung, und seien $\mathcal{B}$ und $\mathcal{B}'$ Orthonormalbasen. Dann ist die Matrix S, die im Transformationsgesetz (67) auftritt, unitär. Wir können in diesem Fall statt (67) also auch

$$[f]_{\mathcal{B}'} = S^\dagger\, [f]_{\mathcal{B}}\, S \tag{165}$$

  schreiben, um zu unterstreichen, dass der Basiswechsel zwischen Orthonormalbasen erfolgt. Entsprechend nennen wir zwei $n\times n$-Matrizen A und B (zueinander) *unitär ähnlich*, wenn es eine unitäre $n \times n$-Matrix S gibt, so dass $B = S^\dagger A\, S$ gilt.

- Einige weitere wichtige Typen von linearen Abbildungen auf (endlichdimensionalen) unitären Vektorräumen und die entsprechenden Matrizen:

  ◦ Eine lineare Abbildung $f : V \to V$ heißt *selbstadjungiert* oder *hermitesch*[15] (meist, nicht ganz korrekt, als „hermitisch" ausgesprochen[16]), wenn

$$f^\dagger = f \tag{166}$$

  gilt. Entsprechend heißt eine quadratische Matrix A selbstadjungiert oder hermitesch, wenn $A^\dagger = A$ gilt. Die Matrix einer hermiteschen Abbildung bezüglich einer Orthonormalbasis ist hermitesch.

  ◦ Eine lineare Abbildung $f : V \to V$ heißt *antiselbstadjungiert* oder *antihermitesch*, wenn

$$f^\dagger = -f \tag{167}$$

  gilt. Entsprechend heißt eine quadratische Matrix A antiselbstadjungiert oder antihermitesch, wenn $A^\dagger = -A$ gilt. Die Matrix einer antihermiteschen Abbildung bezüglich einer Orthonormalbasis ist antihermitesch.

[15] Die Begriffe *selbstadjungiert* und *hermitesch* bedeuten im Endlichdimensionalen dasselbe. Im unendlichdimensionalen Kontext, der vor allem für die Quantentheorie wichtig ist, hat *selbstadjungiert* eine klare Bedeutung (deren Details über unseren Stoff hinausgehen), während der Begriff *hermitesch* von verschiedenen Autoren leicht unterschiedlich verwendet wird.

[16] Die Bezeichnung „hermitesch" leitet sich vom Namen des französischen Mathematikers Charles Hermite (1822 – 1901) ab.

- Eine lineare Abbildung $f : V \to V$ heißt *normal* (oder *normaler Operator*), wenn es eine Orthonormalbasis gibt, bezüglich der die Matrix von f diagonal ist (man sagt auch: wenn f „bezüglich einer Orthonormalbasis diagonalisierbar" ist). Wichtiges und extrem einfaches Kriterium: f ist genau dann normal, wenn

$$f \circ f^\dagger = f^\dagger \circ f \tag{168}$$

gilt, d.h. wenn f mit seiner Adjungierten *kommutiert* (vertauscht). Entsprechend heißt eine quadratische Matrix A normal, wenn $A\,A^\dagger = A^\dagger A$ gilt. Die Matrix eines normalen Operators bezüglich einer Orthonormalbasis ist normal.

Hier eine Zusammenfassung der wichtigsten im Zusammenhang mit unitären Vektorräumen auftretenden Matrizentypen und einige ihrer Eigenschaften:

Bezeichnung	**definierende Eigenschaft**	**Eigenwerte**	**ist auch...**
normal	$A\,A^\dagger = A^\dagger A$	komplex	diagonalisierbar
hermitesch	$A^\dagger = A$	reell ($\overline{\lambda} = \lambda$)	normal
reell und symmetrisch	$\overline{A} = A$ und $A^\dagger = A$	reell ($\overline{\lambda} = \lambda$)	hermitesch
antihermitesch	$A^\dagger = -A$	imaginär ($\overline{\lambda} = -\lambda$)	normal
unitär	$A^{-1} = A^\dagger$	$e^{i\varphi}$ für $\varphi \in \mathbb{R}$ ($\lambda^{-1} = \overline{\lambda}$)	normal
Orthogonalprojektion	$A^2 = A = A^\dagger$	$\in \{0, 1\}$ ($\lambda^2 = \lambda = \overline{\lambda}$)	hermitesch

Alle Matrizen der hier aufgelisteten Typen sind normal, d.h. bezüglich einer ON-Basis diagonalisierbar. Wie aus der Tabelle ersichtlich, bestimmt der Typ einer normalen Matrix, in welcher Teilmenge von $\mathbb{C}$ ihre Eigenwerte liegen. Der Vergleich der mittleren beiden Spalten zeigt eindrucksvoll, wie sich die Eigenschaften von A auf entsprechende Eigenschaften der Eigenwerte übertragen. Interessanterweise gilt, gewissermaßen als Umkehrung, dass die Eigenwerte einer normaler Matrix deren Typ bestimmen: Ist A normal, und sind alle Eigenwerte von A reell/imaginär/vom Betrag $1/\in \{0, 1\}$, so ist A hermitesch/antihermitesch/unitär/eine Orthogonalprojektion. (Das wird sich weiter unten, wenn wir auf die Spektraldarstellung normaler Operatoren zu sprechen kommen, ganz zwanglos zeigen lassen.)

Besonders wichtig für die Quantenphysik ist die Tatsache, dass hermitesche Matrizen nur reelle Eigenwerte haben. Hier ein kleiner Beweis: Ist λ ein Eigenwert von A, so existiert ein $v \neq 0$ mit $Av = \lambda v$. Daraus folgt nach Multiplikation von links mit $v^\dagger$

$$v^\dagger A v = \lambda\, v^\dagger v\,. \tag{169}$$

Beide Seiten dieser Beziehung sind komplexe Zahlen, die wir als 1×1-Matrizen auffassen, und deren adjungierte 1×1-Matrizen wir nun unter Ausnutzung der für das Adjungieren zuständigen Rechenregeln bilden. Die linke Seite von (169) wird dadurch in sich selbst übergeführt: $(v^\dagger A v)^\dagger = v^\dagger A^\dagger (v^\dagger)^\dagger = v^\dagger A v$, wobei die Hermitizität von A verwendet wurde. Die rechte Seite von (169) wird dadurch komplex konjugiert: $(\lambda v^\dagger v)^\dagger = \overline{\lambda}\, v^\dagger (v^\dagger)^\dagger = \overline{\lambda}\, v^\dagger v$. Daher haben wir

$$v^\dagger A v = \overline{\lambda}\, v^\dagger v\,, \tag{170}$$

woraus mit (169) folgt, dass $\overline{\lambda}\, v^\dagger v = \lambda\, v^\dagger v$ gilt. Wegen $v^\dagger v \neq 0$ (es gilt ja $v^\dagger v = \langle v, v\rangle = \|v\|^2$ bezüglich des inneren Produkts im $\mathbb{C}^n$) ergibt sich $\overline{\lambda} = \lambda$, was bedeutet, dass λ reell ist[17].

Wichtig ist auch der Umstand, dass die Eigenwerte einer unitären Matrix stets den Betrag 1 haben (also von der Form $e^{i\varphi}$ für $\varphi \in \mathbb{R}$ sind). Auch das lässt sich leicht zeigen: Ist A unitär, so gilt $A^\dagger A = E$. Ist nun λ ein Eigenwert von A, so existiert ein $v \neq 0$ mit $Av = \lambda v$. Daher gilt

$$(Av)^\dagger A v = v^\dagger \underbrace{A^\dagger A}_{E} v = v^\dagger v. \tag{171}$$

Die linke Seite ist aber gleich

$$(\lambda v)^\dagger \lambda v = \overline{\lambda}\lambda\, v^\dagger v = |\lambda|^2\, v^\dagger v, \tag{172}$$

daher gilt $|\lambda|^2\, v^\dagger v = v^\dagger v$, und wegen $v \neq 0$ folgt $|\lambda|^2 = 1$ und somit $|\lambda| = 1$.

- Zuletzt die *Spektraldarstellung normaler Operatoren*: Ist $f : V \to V$ ein normaler Operator, dann gilt

$$f = \sum_{k=1}^{r} \lambda_k\, P_k\,, \tag{173}$$

wobei $\lambda_1, \ldots, \lambda_r$ die verschiedenen Eigenwerte von f sind und $P_k : V \to V$ jeweils die Orthogonalprojektion auf den Eigenraum E_{λ_k} bezeichnet. Für diese Orthogonalprojektionen (die definitionsgemäß hermitesch sind, also ${P_k}^\dagger = P_k$ erfüllen) gilt

$$P_j\, P_k = \delta_{jk}\, P_k \tag{174}$$

(hier keine Summe über k, obwohl es zweimal vorkommt!) und die Vollständigkeitsrelation

$$\sum_{k=1}^{r} P_k = \mathrm{Id}_V\,. \tag{175}$$

Für manche Anwendungen ist es nötig, die P_k durch eine Orthonormalbasis von f auszudrücken und beliebige Vektoren $x \in V$ in diese Basis zu entwickeln. Das geht so:

[17] Wenn Sie mit der Zeit einen Blick für derartige Dinge entwickeln, können Sie diesen Beweis verkürzen: Die linke Seite von (169) ist reell, $v^\dagger v$ ist ebenfalls reell und außerdem $\neq 0$, und daher ist auch λ reell.

Wählt man in in jedem Eigenraum eine Orthonormalbasis, also im Eigenraum E_{λ_k} die Basis $(w_1^{(k)}, \ldots, w_{n_k}^{(k)})$, so gilt für jedes $x \in V$

$$P_k(x) = \sum_{j=1}^{n_k} \langle w_j^{(k)}, x \rangle \, w_j^{(k)} \, . \tag{176}$$

(Die Schreibweise $x \mapsto \langle w, x \rangle \, w$ für ein gegebenes $w \in V$ mit $\|w\| = 1$ ist nur eine clevere Art, eine Orthogonalprojektion auf den von w aufgespannten eindimensionalen Untervektorraum anzuschreiben – siehe Übungen.) Um P_k „an sich", also ohne Angabe eines Vektors x zu schreiben, kann (176) auch in der Form

$$P_k = \sum_{j=1}^{n_k} \langle w_j^{(k)}, \,\cdot\, \rangle \, w_j^{(k)} \, . \tag{177}$$

ausgedrückt werden[18]. Mit (175) erhalten wir daraus

$$x = \sum_{k=1}^{r} \sum_{j=1}^{n_k} \langle w_j^{(k)}, x \rangle \, w_j^{(k)} \qquad \text{für alle } x \in V. \tag{178}$$

Damit ist eine bequeme (in der Quantenmechanik oft benutzte) Formel für die Entwicklung eines beliebigen Vektors in eine Orthonormalbasis aus Eigenvektoren von f gefunden.

Wie im Zusammenhang mit der Spektraldarstellung selbstadjungierter Operatoren in euklidischen Vektorräumen bereits besprochen, kann auch hier $F(f)$ für eine beliebige Funktion $F : \mathrm{Sp}(f) \to \mathbb{C}$ gebildet werden. In der Quantenphysik besonders bedeutsam:

$$e^{iaf} = \sum_{k=1}^{r} e^{i\,a\,\lambda_k} \, P_k \tag{179}$$

für $a \in \mathbb{R}$ und ein hermitesches f.

Aus der Existenz der Spektraldarstellung normaler Operatoren folgt ein Zusammenhang, den wir bereits oben (unterhalb der Tabelle der Matrizentypen) für normale Matrizen erwähnt haben, und den wir nun allgemein für normale Operatoren formulieren: Ist f normal, und sind alle Eigenwerte von f reell/imaginär/vom Betrag $1/\in \{0, 1\}$, so ist f hermitesch/antihermitesch/unitär/eine Orthogonalprojektion. Sind beispielsweise alle λ_k reell, so folgt sofort aus (173), zusammen mit der Hermitezität von P_k, dass $f^\dagger = f$ ist. Sind alle λ_k imaginär, so folgt analog, dass $f^\dagger = -f$ gilt. Unter Ausnutzung von (174) und (175) kann man leicht die beiden verbleibenden Fälle zeigen (alle λ_k haben Betrag $1 \Rightarrow f^\dagger f = \mathrm{Id}_V$, alle $\lambda_k \in \{0, 1\} \Rightarrow f^2 = f = f^\dagger$).

[18] Der Bra-Ket-Formalismus, der weiter unten noch kurz beschrieben wird, stellt eine treffendere Schreibweise dafür zur Verfügung – siehe (184)!

Bra-Ket-Formalismus

In der Quantenmechanik der „Systeme mit endlich vielen Freiheitsgraden“ (z.B. Spin und Polarisation) werden endlichdimensionale unitäre Vektorräume benötigt. Deren Elemente sind die „Zustandsvektoren“, während physikalische Observable durch lineare Abbildungen dargestellt werden, und alles, was sich physikalisch voraussagen lässt, nämlich die möglichen Messwerte und die Wahrscheinlichkeiten, mit denen sie auftreten, werden letztlich durch Eigenwerte und innere Produkte ausgedrückt. Aus welchen Objekten diese Räume im Einzelnen bestehen, spielt dabei keine Rolle, sofern sich mit ihrer Hilfe alle benötigten physikalischen Aussagen gewinnen lassen. Wir haben es also in gewisser Weise mit „abstrakten“ unitären Vektorräumen zu tun.

Soll ein solcher Raum V von Zustandsvektoren die Dimension n haben, so könnte man für ihn einfach den $\mathbb{C}^n$ nehmen, was auch oft getan wird. Eine andere Möglichkeit besteht darin, ihn so „abstrakt“ zu lassen wie möglich, um sich nicht mit individuellen Eigenschaften seiner Elemente, die nichts zur Sache tun, zu belasten. Man beginnt meist mit einer Orthonormalbasis, die wir als $(|1\rangle, \dots, |n\rangle)$ bezeichnen[19], schreibt ein allgemeines Element in der Form

$$|\psi\rangle = b_1|1\rangle + \cdots + b_n|n\rangle \equiv \sum_{j=1}^{n} b_j|j\rangle \tag{180}$$

an und nennt diese Objekte „Kets“. Jedes Element des Dualraums V^* entspricht in natürlicher Weise einem $|\phi\rangle \in V$ – siehe (152) – und wird in der Form $\langle\phi|$ angeschrieben, genannt „Bra“. Bra und Ket gehören zusammen, um in der Form $\langle\phi|\psi\rangle$ („bracket“) ein inneres Produkt zu bilden. Das Schöne an diesem *Bra-Ket-Formalismus* ist, dass ein lineares Funktional wie $\langle\phi|$ gemeinsam mit einem Vektor $|\psi\rangle$ zu einer linearen Abbildung der Form

$$T = |\psi\rangle\langle\phi| \quad \text{wirkt wie} \quad T : |\chi\rangle \mapsto T|\chi\rangle = |\psi\rangle\langle\phi|\chi\rangle \tag{181}$$

kombiniert werden kann[20]. Mit einem Wort: Wir haben hier ein Baukastensystem, das alles Unwesentliche ausklammert und eine wichtige Sache – die kanonische Identifizierung von V^* mit V, die der Konstruktion linearer Operatoren dient – besonders deutlich und rechenfreundlich zur Geltung bringt. Das einzige, worauf man aufpassen muss, ist, dass diese Identifizierung nicht linear, sondern konjugiert linear ist. Ist etwa $|\phi\rangle = \sum_{k=1}^{n} a_k|k\rangle$, so ist $\langle\phi| = \sum_{k=1}^{n} \overline{a_k}\,\langle k|$. Mit $|\psi\rangle$ von (180) berechnen wir

$$\langle\phi|\psi\rangle = \sum_{j,k=1}^{n} \overline{a_k}\, b_j \underbrace{\langle k|j\rangle}_{\delta_{jk}} = \sum_{j=1}^{n} \overline{a_j}\, b_j \tag{182}$$

und können damit innere Produkte durch die Koeffizienten der beteiligten Vektoren ausdrücken. Von besonderer Wichtigkeit für die Quantenphysik sind Orthogonalprojektionen. Für

[19] Oft lässt man die Basis mit $|0\rangle$ beginnen und hört mit $|n-1\rangle$ auf, und manchmal verwendet man andere Bezeichnungen, wie beispielsweise $|e_j\rangle$. Im Fall $n = 2$ sind auch Bezeichnungen wie $|\uparrow\rangle$ und $|\downarrow\rangle$ üblich.

[20] Da die Vektoren bei dieser Darstellung von T nicht in der Reihenfolge Bra-Ket angeschrieben werden, sondern in umgekehrter Reihenfolge, also $|\psi\rangle\langle\phi|$, können wir von der „Ket-Bra-Darstellung“ einer linearen Abbildung sprechen.

jedes $|\psi\rangle \in V$ mit $\langle\psi|\psi\rangle = 1$ ist

$$P = |\psi\rangle\langle\psi| \tag{183}$$

die Orthogonalprojektion auf $|\psi\rangle$ (genauer: auf den eindimensionalen Teilraum $L(|\psi\rangle)$).

> Bemerkung: Solchen eindimensionalen Orthogonalprojektionen sind wir bereits begegnet, und zwar in (176). Im Bra-Ket-Formalismus würde man (176) mit
>
> $$P_k = \sum_{j=1}^{n_k} |w_j^{(k)}\rangle\langle w_j^{(k)}| \tag{184}$$
>
> in der Form
>
> $$P_k|x\rangle = \sum_{j=1}^{n_k} |w_j^{(k)}\rangle\langle w_j^{(k)}|x\rangle \tag{185}$$
>
> anschreiben.

Ist $f : V \to V$ ein linearer Operator, so wird statt $f(|\psi\rangle)$ einfach $f\,|\psi\rangle$ geschrieben, und bilden wir das innere Produkt von $|\phi\rangle$ mit $f\,|\psi\rangle$, so schreiben wir das in der Form $\langle\phi|f|\psi\rangle$. Mit $|\phi\rangle = f^\dagger|\xi\rangle$ kann $\langle\phi|$ interessanterweise in der Form $\langle\phi| = \langle\xi|f$ geschrieben werden[21], also ohne „Dolch"! Der Ausdruck $\langle\xi|f|\psi\rangle$ steht daher gleichzeitig für $\langle f^\dagger(\xi), \psi\rangle$ und $\langle\xi, f(\psi)\rangle$, was aber gemäß (153) ohnehin keinen Unterschied macht! Das ist eben der Vorteil eines klugen Baukastensystems: Um manche Dinge muss man sich nicht eigens kümmern – die erledigt der Formalismus ganz automatisch für uns.

Manche Berechnungen lassen sich aber doch leichter in Matrizenform durchführen. Zu diesem Zweck kann man $|j\rangle$ mit e_j, dem j-ten Vektor der Standardbasis von $\mathbb{C}^n$, identifizieren und $\langle j|$ mit dem entsprechenden linearen Funktional. Verbreitet ist die Konvention, hier für $j = 1$ illustriert:

$$|1\rangle \longleftrightarrow \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{und} \qquad \langle 1| \longleftrightarrow (1 \;\; 0 \;\; \ldots \;\; 0)\,. \tag{186}$$

Wird also $|\psi\rangle \in V$ mit einem als Spaltenvektor angeschriebenen n-Tupel $\psi \in \mathbb{C}^n$ identifiziert, so entspricht $\langle\psi|$ dem Zeilenvektor $\psi^\dagger$. In diesem Sinn ist $(|\psi\rangle)^\dagger = \langle\psi|$. Die Projektion (183) kann dann als Matrix $P = \psi\,\psi^\dagger$ aufgefasst werden (wobei wir – wie in der Literatur vielfach üblich – die Matrix und die lineare Abbildung auf V mit dem gleichen Symbol P bezeichnet haben). Einem Operator der Form $|j\rangle\langle k|$ entspricht eine Matrix, in der in der j-ten Zeile und k-ten Spalte 1 steht und sonst überall 0. (Für $j = k$ ist ein solcher Operator vom Typ (183), also die Orthogonalprojektion auf den von $|j\rangle$ aufgespannten eindimensionalen Teilraum.) Aus Operatoren dieser Form lassen sich dann alle linearen Abbildungen zusammensetzen: Sind a_{jk} die Komponenten der Matrix einer linearen Abbildung f bezüglich der Orthonormalbasis $(|1\rangle, \ldots, |n\rangle)$, so gilt

$$f = \sum_{j,k=1}^{n} a_{jk}|j\rangle\langle k|. \tag{187}$$

[21] Ersetzen wir f durch $f^\dagger$, so ergibt sich: Ist $|\phi\rangle = f|\xi\rangle$, so ist $\langle\phi| = \langle\xi|f^\dagger$. Wir erwähnen noch, dass für beliebige Vektoren $(|\phi\rangle\langle\psi|)^\dagger = |\psi\rangle\langle\phi|$ gilt.

Umgekehrt lässt sich damit aus der Ket-Bra-Darstellung einer linearen Abbildung deren Matrix durch Ablesen der Koeffizienten leicht bestimmen. Für $f = \mathrm{Id}_V$ gilt $a_{jk} = \delta_{jk}$, womit sich aus (187) die Vollständigkeitsrelation für die Basis $(|1\rangle, \ldots, |n\rangle)$, meist angeschrieben in der Form

$$\sum_{j=1}^{n} |j\rangle\langle j| = \mathrm{Id}_V \,, \tag{188}$$

ergibt.

Beispiel: Für $n = 2$ definieren wir die drei *Pauli-Operatoren* und notieren daneben deren Matrizen, die *Pauli-Matrizen*:

$$\sigma_1 = |1\rangle\langle 2| + |2\rangle\langle 1| \qquad \longleftrightarrow \qquad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{189}$$

$$\sigma_2 = -i\,|1\rangle\langle 2| + i\,|2\rangle\langle 1| \qquad \longleftrightarrow \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{190}$$

$$\sigma_3 = |1\rangle\langle 1| - |2\rangle\langle 2| \qquad \longleftrightarrow \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{191}$$

Sie sind hermitesch, ihre Quadrate sind gleich Id_V bzw. E_2, ihre Eigenwerte sind daher[22] gleich -1 und 1. Die Operatoren bzw. Matrizen $S_j = \frac{\hbar}{2}\sigma_j$ stellen den (nichtrelativistischen) quantenmechanischen Spin dar und erfüllen die bedeutsame Beziehung (angeschrieben unter Verwendung der Einsteinschen Summenkonvention)

$$[S_j, S_k] = i\,\hbar\,\varepsilon_{jkl}\,S_l \,, \tag{192}$$

die sogenannte *Drehimpuls-Algebra*. Dabei wurde mit den eckigen Klammern der *Kommutator* zweier Operatoren bzw. Matrizen bezeichnet: $[A, B] = AB - BA$.

Hat man sich einmal an den Bra-Ket-Foralismus gewöhnt, so kann man ihn natürlich auch für Formulierungen und Berechnungen in konkreten unitären Vektorräumen benutzen, nicht nur für den „abstrakten" unitären Vektorraum, an den wir zu Beginn dieser kurzen Beschreibung gedacht haben. Hier kommt einem noch einmal die wichtige Erkenntnis zugute, dass unitäre Vektorräume der gleichen (endlichen) Dimension, durch die „Unitäre-Vektorraum-Brille" betrachtet, ohnehin alle gleich aussehen.

---

[22] Dieses „daher" ist ein schönes Beispiel für eine theoriegeladene Argumentation, die aber das Rechnen abkürzt: Gilt für einen linearen Operator $f^2 = \mathrm{Id}$, so ist $f^2(x) = x$ für alle $x \in V$. Ist x ein Eigenvektor von f zum Eigenwert λ, so folgt $x = f^2(x) = f(f(x)) = f(\lambda x) = \lambda f(x) = \lambda \cdot \lambda\, x = \lambda^2 x$, daher $x = \lambda^2 x$. Da $x \neq 0$ ist, muss $\lambda^2 = 1$ sein, daher $\lambda = -1$ oder $\lambda = 1$. Ist f zudem hermitesch, so ist f diagonalisierbar, und daher gibt es eine diagonale Matrixdarstellung mit lediglich -1 und/oder 1 als Eigenwerten. Folglich gibt es drei Möglichkeiten: $f = \mathrm{Id}$ oder $f = -\mathrm{Id}$ oder sowohl -1 als auch 1 ist Eigenwert von f. Letzteres muss für die Pauli-Operatoren der Fall sein, da sie weder gleich Id noch gleich $-\mathrm{Id}$ sind.

Ergänzung 6: Das Tensorprodukt in a nutshell

Nehmen Sie zwei (endlichdimensionale) Vektorräume V und W über $\mathbb{K}$ und jeweils eine Basis $(v_1, \ldots, v_m)$ und $(w_1, \ldots, w_n)$. Jetzt fassen Sie sich ein Herz und betrachten die „abstrakten Produkte" dieser Basisvektoren:

$$\begin{array}{cccc} v_1 \otimes w_1 & v_1 \otimes w_2 & \ldots & v_1 \otimes w_n \\ v_2 \otimes w_1 & v_2 \otimes w_2 & \ldots & v_2 \otimes w_n \\ \vdots & \vdots & \ddots & \vdots \\ v_m \otimes w_1 & v_m \otimes w_2 & \ldots & v_m \otimes w_n \,. \end{array} \tag{193}$$

Ordnen Sie sie in beliebiger Weise in Form eines mn-Tupels (beispielsweise, indem Sie das obige Schema Zeile für Zeile durchgehen wie einen geschriebenen Text) und *erklären* dieses mn-Tupel zu einer Basis eines neuen („abstrakten") Vektorraums $V \otimes W$. *Voilà*!

Konkreter muss es für viele physikalische Anwendungen nicht sein[23]! $V \otimes W$ heißt das *Tensorprodukt* der Vektorräume V und W. Die Objekte $v_j \otimes w_k$ sind die *Tensorprodukte* der entsprechenden Basisvektoren, Elemente von $V \otimes W$ heißen *Tensoren*. Klarerweise ist $\dim(V \otimes W) = mn$, und Sie können mit diesem neuen Raum nun alles tun, was Sie in jedem Vektorraum tun können. Beispielsweise einen gegebenen Tensor $t \in V \otimes W$ in die soeben konstruierte Basis zu entwickeln:

$$t = \sum_{j=1}^{m} \sum_{k=1}^{n} t_{jk}\, v_j \otimes w_k, \tag{194}$$

wobei wir hier der Konvention gefolgt sind, für die Entwicklungskoeffizienten (Komponenten) t_{jk} das gleiche Symbol wie für den Tensor (also t) mit angehängten Indizes zu verwenden. Die Vektorraum-Operationen Addition und Bilden von skalaren Vielfachen sind dann einfach definiert durch

$$\sum_{j=1}^{m} \sum_{k=1}^{n} t_{jk}\, v_j \otimes w_k + \sum_{j=1}^{m} \sum_{k=1}^{n} s_{jk}\, v_j \otimes w_k := \sum_{j=1}^{m} \sum_{k=1}^{n} (t_{jk} + s_{jk})\, v_j \otimes w_k \tag{195}$$

$$\lambda \sum_{j=1}^{m} \sum_{k=1}^{n} t_{jk}\, v_j \otimes w_k := \sum_{j=1}^{m} \sum_{k=1}^{n} \lambda\, t_{jk}\, v_j \otimes w_k. \tag{196}$$

Mit $x = \sum_{j=1}^{m} x_j v_j \in V$ und $y = \sum_{k=1}^{n} y_k w_k \in W$ können wir nun auch beliebige Tensorprodukte von Vektoren der beiden Ausgangsvektorräume bilden:

$$x \otimes y = \sum_{j=1}^{m} \sum_{k=1}^{n} x_j y_k\, v_j \otimes w_k \,, \tag{197}$$

wobei genau genommen zuvor die Rechenregel vereinbart werden muss, dass das Bilden solcher Tensorprodukte eine bilineare Operation ist, d.h. dass $(x + \lambda u) \otimes y = x \otimes y + \lambda\, u \otimes y$ und $x \otimes (y + \lambda z) = x \otimes y + \lambda\, x \otimes z$ für beliebige $x, u \in V$, $y, z \in W$ und $\lambda \in \mathbb{K}$.

[23] Zur Notation: Unter Verwendung der Bra-Ket-Schreibweise wird statt $|1\rangle \otimes |1\rangle$ und $|1\rangle \otimes |2\rangle$ oft einfach $|11\rangle$ und $|12\rangle$ geschrieben. Manchmal wird das Tensorprodukt $|x\rangle \otimes |y\rangle$ einfach durch Nebeneinanderschreiben der betreffenden Vektoren in der Form $|x\rangle|y\rangle$ ausgedrückt.

Man kann auch Tensorprodukte von linearen Operatoren auf den beiden Ausgangsvektorräumen bilden: Sind $f : V \to V$ und $g : W \to W$ lineare Operatoren, so ist deren Tensoprodukt $f \otimes g : V \otimes W \to V \otimes W$ definiert durch

$$(f \otimes g)(x \otimes y) = f(x) \otimes g(y) \qquad \text{für } x \in V \text{ und } y \in W \tag{198}$$

und lineare Fortsetzung auf *alle* Elemente von $V \otimes W$, was allgemein durch die Formel

$$(f \otimes g)\Big(\sum_{j=1}^{m}\sum_{k=1}^{n} t_{jk}\, v_j \otimes w_k\Big) = \sum_{j=1}^{m}\sum_{k=1}^{n} t_{jk}\, f(v_j) \otimes g(w_k) \tag{199}$$

ausgedrückt werden kann.

Sind V und W jeweils mit einem Skalarprodukt (falls $\mathbb{K} = \mathbb{R}$) oder einem inneren Produkt (falls $\mathbb{K} = \mathbb{C}$) ausgestattet, so definieren wir mit

$$\langle v \otimes w, v' \otimes w' \rangle = \langle v, v' \rangle\, \langle w, w' \rangle \tag{200}$$

(und entsprechender bilinearer bzw. sesquilinearer Fortsetzung) ein Skalarprodukt bzw. ein inneres Produkt auf dem Tensorprodukt $V \otimes W$, das damit zu einem euklidischen bzw. unitären Vektorraum wird.

An dieser Stelle ist ein Check angebracht, dass der so definierte neue Vektorraum (euklidische Vektorraum, unitäre Vektorraum) $V \otimes W$ nicht davon abhängt, welche Basen man zu Beginn in V und W ausgewählt hat. Wären für die Vorlesung mehr Stunden vorgesehen, so könnten wir das ausführlich durchexerzieren und nebenbei eine „schönere“ Definition des Tensorprodukts zweier Vektorräume erhalten, die keinerlei Basiswahl voraussetzt[24]. Wir begnügen uns damit, die Wirkung eines Basiswechsels in V und W zu betrachten: Ist $(v'_1, \ldots, v'_m)$ eine zweite („neue“) Basis von V und $(w'_1, \ldots, w'_n)$ eine zweite („neue“) Basis von W, so gibt es eine invertierbare $m \times m$-Matrix R und eine invertierbare $n \times n$-Matrix S, sodass (vgl. (71))

$$v'_a = \sum_{j=1}^{m} v_j R_{ja} \qquad \text{und} \qquad w'_b = \sum_{k=1}^{n} w_k S_{kb} \tag{201}$$

gilt. Vektorkomponenten transformieren (d.h. ändern sich unter Basiswechsel) dann gemäß dem Muster

$$x = \sum_{j=1}^{m} x_j\, v_j = \sum_{a=1}^{m} x'_a\, v'_a \in V \quad \Rightarrow \quad x'_a = \sum_{j=1}^{m} (R^{-1})_{aj}\, x_j \tag{202}$$

und

$$y = \sum_{k=1}^{n} y_k\, w_k = \sum_{b=1}^{n} y'_b\, w'_b \in W \quad \Rightarrow \quad y'_b = \sum_{k=1}^{n} (S^{-1})_{bk}\, y_k \tag{203}$$

[24] Hier nur so viel: Eine basisfreie Möglichkeit, das Tensorprodukt zweier Vektorräume zu definieren, ist $V \otimes W = \mathrm{Hom}(V^*, W)$. Das Tensorprodukt $v \otimes w$ zweier Vektoren $v \in V$ und $w \in W$ ist dann jene lineare Abbildung $V^* \to W$, die $\varphi \in V^*$ auf $\varphi(v)w \in W$ abbildet. Eine zweite – dazu äquivalente – Möglichkeit besteht darin, das Tensorprodukt $V \otimes W$ als Menge aller bilinearen Abbildungen $V^* \times W^* \to \mathbb{K}$ zu definieren. Das Tensorprodukt $v \otimes w$ zweier Vektoren $v \in V$ und $w \in W$ ist dann die Abbildung $(\varphi, \psi) \in V^* \times W^* \mapsto \varphi(v)\, \psi(w) \in \mathbb{K}$.

(vgl. (72)) und (73)). Dabei haben wir die Symbole für die Indizes so gewählt, dass sie ihre jeweilige Rolle anzeigen: j und a gehören zu V (j zur alten Basis und a zur neuen Basis, beide laufen von 1 bis m), und k und b gehören zu W (k zur alten Basis und b zur neuen Basis, beide laufen von 1 bis n). Wird nun der Tensor (194) bezüglich jener Basis von $V \otimes W$, die aus den Tensorprodukten der neuen Basisvektoren besteht, entwickelt,

$$t = \sum_{j=1}^{m} \sum_{k=1}^{n} t_{jk}\, v_j \otimes w_k = \sum_{a=1}^{m} \sum_{b=1}^{n} t'_{ab}\, v'_a \otimes w'_b\,, \tag{204}$$

so ergibt sich aufgrund der vereinbarten Rechenregeln mit

$$t'_{ab} = \sum_{j=1}^{m} \sum_{k=1}^{n} (R^{-1})_{aj} (S^{-1})_{bk}\, t_{jk} \tag{205}$$

das *Transformationsverhalten von Tensorkomponenten* (versuchen Sie, es nachzurechnen!). Mit (202) und (203) lässt es sich kurz auch so formulieren: Der Übergang $t_{jk} \to t'_{ab}$ verläuft genauso wie der Übergang $x_j\, y_k \to x'_a\, y'_b$ für Vektoren $x \in V$ und $y \in W$. Das Transformationsverhalten von Tensorkomponenten ist wichtig etwa in der Relativitätstheorie (vor allem in der allgemeinen Relativitätstheorie, in der es vor Tensoren nur so wimmelt).

All diese Konzepte lassen sich direkt auf höhere Tensorprodukte $V_1 \otimes V_2 \otimes \cdots \otimes V_r$ verallgemeinern.

Eine letzte Betrachtung wollen wir noch dem Fall $V = W = \mathbb{C}^2$ widmen, der in der Theorie der Quanteninformation von Bedeutung ist. Sind $A, B : \mathbb{C}^2 \to \mathbb{C}^2$ lineare Abbildungen, als komplexe 2×2-Matrizen dargestellt, so kann das durch (198) definierte Tensorprodukt $A \otimes B$, das eine lineare Abbildung $\mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$, also eine lineare Abbildung auf einem vierdimensionalen Vektorraum ist, durch eine 4×4-Matrix dargestellt werden. In Anwendungen ist es oft hilfreich, diese 4×4-Matrix explizit zu kennen. Man kann sie folgendermaßen berechnen: Gehen wir von der Standardbasis in V und W aus und ordnen die Basis von $V \otimes W$ in der Form $(e_1 \otimes e_1,\, e_1 \otimes e_2,\, e_2 \otimes e_1,\, e_2 \otimes e_2)$, so ergibt sich mit

$$\begin{aligned} A \otimes B &= \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix} = \\ &= \begin{pmatrix} a_{11}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} & a_{12}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \\ a_{21}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} & a_{22}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \end{pmatrix} \equiv \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix} \end{aligned} \tag{206}$$

eine schöne Formel, die es erlaubt, die linearen Abbildungen $\mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$, die sich alle als Summen von Abildungen der Form $A \otimes B$ darstellen lassen, ohne großen Aufwand zu studieren.

Schreiben Sie zur Übung $E_2 \otimes E_2$, $E_2 \otimes \sigma_j$, $\sigma_j \otimes E_2$ und $\sigma_j \otimes \sigma_k$ (für $j, k = 1, 2, 3$) als 4×4-Matrizen an! Diese 16 Matrizen bilden eine Basis des Vektorraums aller komplexen 4×4-Matrizen.

Beweisen Sie zur Übung, dass $(A \otimes B)(A' \otimes B') = (AA') \otimes (BB')$, $\mathrm{Tr}(A \otimes B) = \mathrm{Tr}(A)\,\mathrm{Tr}(B)$ und $\det(A \otimes B) = \det(A)^2 \det(B)^2$.

Tensorprodukte werden in der Quantentheorie verwendet, um zusammengesetzte Systeme zu beschreiben. Wird etwa ein Teilchen mit dem Hilbertraum V_1 beschrieben und ein anderes Teilchen mit dem Hilbertraum V_2, so wird das aus beiden Teilchen bestehende System mit dem Hilbertraum $V_1 \otimes V_2$ beschrieben. Befindet sind das erste Teilchen in einem durch $\psi_1 \in V_1$ beschriebenen Zustand, das zweite in einem durch $\psi_2 \in V_2$ beschriebenen Zustand, so befindet sich das Gesamtsystem in dem durch $\psi_1 \otimes \psi_2 \in V_1 \otimes V_2$ beschriebenen Zustand. Alle Zustandvektoren in $V_1 \otimes V_2$, die *nicht* in dieser Form als Tensorprodukt geschrieben werden können, beschreiben sogenannte *verschränkte Zustände*[25]. Wird am ersten Teilchen eine Messung der durch den Operator A dargestellten Observable und am zweiten Teilchen eine Messung der durch den Operator B dargestellten Observable durchgeführt, so wird diese Messung am Gesamtsystem durch den Operator $A \otimes B$ beschrieben.

Haben Sie in diesem Ergänzungsskriptum etwas schmerzlich vermisst? Einen Fehler entdeckt? Dann schreiben Sie mir bitte unter franz.embacher@univie.ac.at.

[25] Ein Beispiel eines verschränkten Zustands wäre $\psi_1 \otimes \psi_2 + \phi_1 \otimes \phi_2$, wobei (ψ_1, ϕ_1) linear unabhängig (in V_1) und (ψ_2, ϕ_2) linear unabhängig (in V_2) ist. Der berühmte EPR-Zustand (nach Einstein, Podolsky und Rosen) ist von dieser Form.