Vorwort

Dieses Buch ist aus einer Vorlesung hervorgegangen, die der ältere der beiden Autoren im Wintersemester 1987/88 an der Universität Göttingen gehalten hat, die vom jüngeren Autor ausgearbeitet und zunächst vom Mathematischen Institut der Universität Göttingen herausgegeben worden ist. Für die Veröffentlichung als Lehrbuch ist dieses Manuskript noch einmal grundlegend überarbeitet und an zahlreichen Stellen ergänzt worden. Wir hoffen, daß die Darstellung so ausführlich, interessant und klar geraten ist, daß sich das Buch nicht nur als Begleitlektüre zu einer entsprechenden Vorlesung, sondern auch zum Selbststudium, zum Nachlesen oder zur Vorbereitung auf Prüfungen eignet.

Auf die Stoffauswahl gehen wir in der Einleitung ausführlich ein. Hier sei nur bemerkt, daß einerseits der Geometrie und vor allem der Nahtstelle zwischen geometrischer Anschauung und mathematisch-logischer Formulierung ein großes Gewicht eingeräumt wird. Andererseits wird ein großer Teil der grundlegenden Theorie, wie er etwa auch von Studenten der Wirtschaftsmathematik oder Physik benötigt wird, in diesem Buch behandelt, auch wenn es nur den Inhalt des ersten Teils der zweisemestrigen Vorlesung „Analytische Geometrie und lineare Algebra“ wiedergibt. Breiten Raum nehmen die Behandlung linearer Gleichungssysteme (Gaußscher Algorithmus), die Determinantentheorie und die Eigenwerttheorie linearer Abbildungen in „euklidischen“ Vektorräumen (Hauptachsentransformation) ein.

Wir sind bemüht, schnell zu den Kernpunkten der jeweiligen Themenbereiche vorzustoßen und verzichten dabei bewußt auf größtmögliche Allgemeinheit. Zahlreiche Beispiele, die teilweise den Charakter von gelösten Übungsaufgaben haben, sollen das Verständnis neuer Begriffe oder Methoden unmittelbar fördern.

Wir haben dem Verlag und insbesondere den Herren Andreas Türk und Martin Reck zu danken für die jederzeit angenehme Zusammenarbeit. Herrn Peter Lerner danken wir für seine tatkräftige Mitwirkung bei der Erstellung der LaTeX-Dateien und Herrn Christian Meister für seine Hilfe bei der Durchsicht des Manuskripts.

Göttingen und Bayreuth, im März 1999

Hans Grauert
Hans-Christoph Grunau

Zu den Autoren

Hans Grauert wurde am 8. Februar 1930 in Haren (Ems) geboren. Er studierte Mathematik und Physik in Mainz, Münster und an der ETH Zürich. Er wurde 1958 an die Universität Göttingen als Nachfolger von C. L. Siegel berufen. Er verbrachte insgesamt drei Jahre in Amerika (u. a. am Institute for Advanced Study in Princeton), war aber auch in anderen Ländern wie vor allem in Frankreich (Institut des Hautes Études Scientifiques) und u. a. in Japan, China, der UdSSR und Indien. Seine wissenschaftlichen Arbeiten befassen sich mit Fragen der komplexen Analysis.

Hans-Christoph Grunau wurde am 23. März 1961 in Wuppertal geboren. Er studierte Mathematik in Marburg und Göttingen und hat sich 1996 an der Universität Bayreuth habilitiert. Sein wissenschaftliches Interesse gilt der reellen Analysis und insbesondere der Theorie der partiellen Differentialgleichungen. Auf diesem Gebiet arbeitet er mit Wissenschaftlern u. a. aus den Niederlanden, Spanien und Italien zusammen.

Inhaltsverzeichnis

Einleitung

In der Anfängervorlesung „Analytische Geometrie und lineare Algebra“ ist es heute vielfach üblich, einfach mit der abstrakten Theorie der Vektorräume und evtl. sogar der Moduln über Ringen zu beginnen. Man hat das Bestreben, sofort von den mathematischen Grundbegriffen auszugehen und aus ihnen alles zusammenzubauen. Das soll zu einer logischen Übersichtlichkeit führen. Aber der Anfänger wird staunen, was es so alles gibt. Er kann sich bei manchem nur wenig denken.

In der Technik, aber auch in der Physik wird im allgemeinen erst nur der 3-dimensionale anschauliche Raum auftreten. Er hat noch keine Koordinaten. Man hat noch keine Tripel von reellen Zahlen, dafür aber gibt es eine Vorstellung von Geraden, Winkeln, Translationen, Drehungen. Vektoren gibt es. Sie sind Pfeile einer bestimmten Länge. Man kann mit ihnen rechnen. Warum geht man dann nicht von diesem Raum aus? Man kann die Vektorraumeigenschaften ableiten und so den Studenten zu den in der Mathematik natürlich notwendigen Abstraktionen hinführen. Die Abstraktionen werden dann als natürlich empfunden. Man kann dann auch die Existenz von rechtwinkligen Koordinaten zeigen und die Isomorphie mit dem Zahlenraum herstellen. Die Existenz einer solchen Isomorphie bedeutet, daß der anschauliche Raum die gleiche Gestalt wie der Zahlenraum hat.

Wer seinen Baum beschneiden will, wird nicht erst Koordinaten in seinen Garten legen.

Das hier vorgestellte Axiomensystem ist auch ein vollständiges Axiomensystem der euklidischen Geometrie. Anders als bei Hilbert werden jedoch keine Geraden, sondern Pfeilvektoren verwendet. Das ist praktischer. Die Pfeilvektoren finden ja überall Anwendung. Natürlich wird nicht die alte hyperbolische Geometrie mit eingeschlossen. Diese dürfte aber auch heute ziemlich ohne Bedeutung sein. Wenn man sie machen will, ist es ohnehin besser, gleich die Quotienten von Liegruppen zu untersuchen! Die nicht-euklidischen Geometrien werden hier gleich dadurch ausgeschlossen, daß man fordert, daß die Relation „parallel“ eine Äquivalenzrelation ist und daß das Parallelogramm der Kräfte gilt.

Es sei noch erwähnt, daß die Griechen keinen Beweis des Satzes von Pythagoras geliefert haben, der den heutigen Anforderungen entspricht. Sie haben nämlich den Flächeninhalt benutzt, ohne ihn zu definieren. Auch hätten sie zeigen müssen, daß er invariant gegen Drehungen ist. An sich hat der Pythagoräische Satz mit dem Flächeninhalt nichts zu tun. Dieses Buch enthält einen exakten Beweis.

In dem ersten Kapitel dieses Buches wird zwar zunächst der n-dimensionale reelle Zahlenraum $\mathbb{R}^n$ als Menge der n-tupel von reellen Zahlen und der Begriff des abstrakten Vektorraums eingeführt, damit man die späteren Begriffe einordnen kann. Dann geht man jedoch im zweiten Kapitel sofort zur Anschaulichkeit über. Insbesondere wird die anschauliche Ebene betrachtet. Hier wird unser Axiomensystem der anschaulichen Geometrie soweit entwickelt, daß die anschauliche Ebene als zweidimensionaler reeller Vektorraum erkannt werden kann.

Hierauf aufbauend entwickeln wir im 3. Kapitel die Grundlagen der Theorie endlichdimensionaler (zunächst ausschließlich reeller) Vektorräume sowie der linearen Abbildungen zwischen diesen. Wir werden zeigen, daß durch Einführung geeigneter „Isomorphismen “ (Koordinatensysteme) die Betrachtung linearer Abbildungen zwischen endlichdimensionalen Vektorräumen zurückgeführt werden kann auf die Betrachtung linearer Abbildungen $\mathbb{R}^n \rightarrow \mathbb{R}^m$. Diese sind dem ebenso einfachen wie wirkungsvollen Matrizenkalkül (dabei handelt es sich lediglich um die Manipulation von Zahlenschemata) zugänglich, den wir simultan und gleichberechtigt zur Theorie linearer Abbildungen entwickeln und mit dessen Hilfe wir zahlreiche theoretische Resultate herleiten werden. Breiten Raum nimmt die Behandlung linearer Gleichungssysteme bzw., was, wie wir in Kapitel 3.7 darlegen werden, dazu gleichwertig ist, allgemeiner s-dimensionaler Ebenen ein. In diesem Zusammenhang wird der Gaußsche Algorithmus vorgestellt, aus dem wir neben rechnerischem auch stets großen theoretischen Nutzen ziehen werden.

In Kapitel 4 sprechen wir speziellere Fragen der Vektorraumtheorie an. Im Mittelpunkt stehen dort duale Vektorräume und duale Abbildungen sowie in diesem Zusammenhang einige Besonderheiten unendlichdimensionaler Vektorräume. Die Idee, zusammen mit einem Vektorraum V die Gesamtheit aller linearen Abbildungen $V \rightarrow \mathbb{R}$ zu betrachten, ist grundlegend in vielen Bereichen der modernen Mathematik, führt zum einen von einer abstrakten Seite hin auf den Begriff der Orthogonalität und bietet zum anderen in allgemeinen Situationen einen Ersatz für bzw. eine Verallgemeinerung von Orthogonalität. Ausführlichere Erklärungen versuchen wir zu Beginn und im Verlauf von Kapitel 4.

Kapitel 5 ist der klassischen Determinantentheorie und dabei wegen ihrer enormen Wichtigkeit unter anderem der Cramerschen Regel gewidmet. Auch dieses Kapitel führt wegen der geometrischen Bedeutung der Determinante hin zu Kapitel 6, in dem unser in Kapitel 2 begonnenes Axiomensystem der anschaulichen Geometrie um diejenigen Axiome ergänzt wird, die Orthogonalität, Länge, Winkel, Drehung, u. ä betreffen.

Damit erhalten die anschauliche Ebene, $\mathbb{R}^2$ und allgemeiner $\mathbb{R}^n$ die volle Struktur „euklidischer Vektorräume“. Das Vorhandensein eines Skalarprodukts erlaubt es uns, in Kapitel 7 etwa „othogonale“, „selbstadjungierte“ oder „normale“ Abbildungen zu betrachten. Die starken Eigenschaften solcher speziellen Abbildungen erlauben die Herleitung sehr weitreichender Resultate („Hauptachsentransformation“), deren volle

Eleganz und Allgemeinheit sich aber nur in komplexen Vektorräumen erreichen läßt. Deshalb beginnt dieses Kapitel mit einer Einführung in komplexe Zahlen und komplexe Vektorräume; in der Hoffnung, daß vieles schon aus der gleichzeitig erlernten Analysis vertraut ist, fassen wir uns hier recht kurz. Für die angesprochenen Klassen von Matrizen werden wir orthogonale Koordinatensysteme konstruieren, so daß das Rechnen mit linearen Abbildungen in diesen Koordinaten so einfach wird wie das Rechnen mit reellen oder komplexen Zahlen. Als Beispiele werden wir Wurzeln und Exponentialfunktionen geeigneter linearer Abbildungen bestimmen. Ganz wesentliches Standbein dieses Kapitels ist die „Eigenwerttheorie linearer Abbildungen", die mit all ihren Weiterentwicklungen und Verallgemeinerungen grundlegend für viele Bereiche der modernen Mathematik und Physik bis hin zur Quantenmechanik ist.

Mit dieser Stoffauswahl hoffen wir:

- die geometrischen Wurzeln der linearen Algebra mit logischer Strenge freizulegen,
- die wichtigsten Elemente der Theorie reeller und komplexer Vektorräume und linearer Abbildungen zu entwickeln und
- stets den Blick schon ein wenig zu öffnen in Richtung auf Weiterentwicklungen im Bereich der Geometrie (allgemeine Ebenen etwa als einfachste Prototypen allgemeiner Mannigfaltigkeiten), der Analysis (unendlichdimensionale Vektorräume werden gestreift, um Appetit zu machen auf die Theorie linearer Abbildungen zwischen vollständigen normierten Vektorräumen, die „Funktionalanalysis") und der mathematischen Physik (etwa Lorentztransformationen oder Eigenwerttheorie).

Technische Hinweise zum Lesen des Buches erübrigen sich weitestgehend. Das Kästchen $\square$ kennzeichnet das Ende von Beweisen. Die Numerierung sollte selbsterklärend und vollkommen unzweideutig sein: Sätze, Definitionen, Beispiele, usw. werden unterschiedslos dreistellig numeriert, die erste Stelle gehört zum Kapitel und die zweite zum Teilkapitel (Abschnitt). Am Ende des Buches befinden sich ein ausführliches Symbolverzeichnis und ein Sachverzeichnis, so daß (hoffentlich!) jedes interessierende Thema im Text mühelos aufgefunden werden kann.

Um eventuellen Mißverständnissen vorzubeugen, sei schon hier erwähnt, daß wir n-tupel aus dem $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ zwar grundsätzlich als Zeilen $\vec{b} = (b_1, \ldots, b_n)$ schreiben, um die Lesbarkeit vor allem des fortlaufenden Textes zu gewährleisten, daß wir im Rahmen des Matrizenkalküls darunter jedoch (zumeist stillschweigend) die entsprechenden Spaltenvektoren

$$\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

verstehen.

7 Eigenwerte und Hauptachsentransformation

In diesem Kapitel werden wir eingehend die am Ende von Abschnitt 3.6 aufgeworfene Frage studieren: Sei $F : V \to V$ eine lineare Abbildung des n-dimensionalen Vektorraums V in sich, $G : \mathbb{R}^n \to V$ eine Karte, $G^{-1} : V \to \mathbb{R}^n$ das entsprechende Koordinatensystem und A die zur Abbildung in den Koordinaten $G^{-1} \circ F \circ G$ gehörige $n \times n$-Matrix.

Für welche Matrizen A läßt sich eine Koordinatenwechselabbildung $G^{-1} \circ \hat{G} : \mathbb{R}^n \to \mathbb{R}^n$ mit zugehöriger $n \times n$-Matrix B finden, so daß $\hat{A} := B^{-1} \circ A \circ B$ *Diagonalgestalt*

$$\hat{A} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_n \end{pmatrix}$$

hat? Wir wollen uns hier auf den Fall beschränken, daß die Karten G und $\hat{G}$ auseinander durch Drehungen hervorgehen, daß die Transformationsmatrix B also orthogonal ist. In diesem Fall sagt man, daß B die *Hauptachsentransformation* für die Matrix A vermittelt.

Man ist deshalb so an einer Diagonalform interessiert, weil diese einerseits eine besonders naheliegende Interpretation der entsprechenden Abbildungen gestattet und weil so andererseits Matrizen auf einfache Weise auch komplexen Berechnungen wie Potenzieren oder unendlichen Reihenbildungen zugänglich gemacht werden.

Hat $\hat{A}$ Diagonalgestalt, so gilt $\hat{A} \circ \vec{e}_j = \lambda_j \vec{e}_j$ und deshalb für die Basiselemente $\hat{x}_j := \hat{G}(\vec{e}_j)$ von V:

$$F(\hat{x}_j) = \lambda_j \hat{x}_j.$$

Derartige Vektoren $\hat{x}_j$ heißen *Eigenvektoren* zum jeweiligen *Eigenwert* λ_j, diese werden unter der Anwendung von F lediglich um den Faktor λ_j gestreckt. Solche Eigenvektoren werden beim Aufsuchen der Diagonalform $\hat{A}$ von A eine grundlegende Rolle spielen, dazu vgl. man den Abschnitt 7.3.

Es wird sich herausstellen, daß eine *geschlossene* Behandlung der angesprochenen Fragen für reelle Vektorräume im allgemeinen nicht möglich ist. Deshalb werden wir im folgenden Abschnitt zunächst den Körper der komplexen Zahlen und komplexe Vektorräume einführen und in Abschnitt 7.2 die unitären Matrizen als Verallgemeinerung der orthogonalen Matrizen behandeln.

In Abschnitt 7.4 werden wir, zunächst im *Komplexen*, für die „größtmögliche“ Klasse der „normalen“ Matrizen Diagonalisierungsresultate herleiten. Ein entsprechendes *reelles* Resultat können wir daraus im allgemeinen nur für die wesentlich speziellere Klasse der „symmetrischen (selbstadjungierten)“ Matrizen folgern, da nur in diesem Fall stets die Existenz hinreichend vieler *reeller* Eigenwerte gesichert ist.

7.1 Der allgemeine Körperbegriff, komplexe Vektorräume

Allgemeine Körper

7.1.1 Definition

Eine Menge K, die mit einer Addition

$$+ : K \times K \to K$$

und einer Multiplikation

$$\cdot : K \times K \to K$$

versehen ist, heißt **Körper**, *falls gilt:*

(1) $(K, +)$ *ist eine abelsche Gruppe.*

(2) *Es bezeichne* 0 *das Nullelement von* $(K, +)$ *und* $K^* := K \backslash \{0\}$. *Dann ist* $(K^*, \cdot)$ *ebenfalls eine abelsche Gruppe.*

(3) *Die Distributivgesetze gelten, d. h. für alle* $a, b, c, d \in K$ *ist*

$$(a + b) \cdot c = a \cdot c + b \cdot c, \qquad a \cdot (c + d) = a \cdot c + a \cdot d.$$

Das neutrale Element bzgl. der Multiplikation wird mit 1 bezeichnet und die inversen Elemente bzgl. der Addition mit $(-a)$ und bzgl. der Multiplikation mit a^{-1} oder $\frac{1}{a}$. In jedem Körper gelten insbesondere sowohl für die Addition als auch für die Multiplikation die Regeln, wie wir sie in Abschnitt 1.1 für (abelsche) Gruppen zusammengestellt haben. Darüber hinaus haben wir für die Verknüpfung additiver und multiplikativer Eigenschaften folgende Gesetze:

7.1.2 Satz

Sei K ein Körper, dann gilt für alle $a, b \in K$:

(a) $a \cdot 0 = 0 \cdot a = 0,$

(b) *aus* $a \cdot b = 0$ *folgt* $a = 0$ *oder* $b = 0$, (Nullteilerfreiheit)

(c) $(-a) \cdot b = a \cdot (-b) = -(a \cdot b), \qquad (-a) \cdot (-b) = a \cdot b.$

Für den Beweis sowie für weitere Folgerungen aus den Körperaxiomen verweisen wir auf die Lehrbücher der Analysis etwa von O. Forster [4, § 2] und von H. Grauert, I. Lieb [6, § 3].

Der Körper der reellen Zahlen verfügt zudem über einen Abstandsbegriff (Betrag), bzgl. dessen er vollständig ist, und über eine archimedische Anordnung ($\leq, \geq$).

Der Körper der komplexen Zahlen

Der für das Folgende außerordentlich wichtige Körper $\mathbb{C}$ der komplexen Zahlen wird im allgemeinen in der Differentialrechnung definiert. Daher werden hier manche Eigenschaften nur skizziert und einige Beweise übergangen. Zur weiteren Information sei wieder auf O. Forster [4, § 13] verwiesen.

Wir führen $\mathbb{C}$ als den zweidimensionalen reellen Vektorraum

$$\mathbb{C} := \mathbb{R}^2 = \{(x, y) : \, x, y \in \mathbb{R}\}$$

ein. Damit ist $\mathbb{C}$ bereits eine additive abelsche Gruppe. Eine Multiplikation

$$\cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$$

wird durch

$$(x, y) \cdot (u, v) := (xu - yv, \, xv + yu)$$

erklärt. Damit wird $\mathbb{C}$ zum Körper: das neutrale Element ist bzgl. der Addition $0 = (0, 0)$ und bzgl. der Multiplikation $1 = (1, 0)$. Das Negative von (x, y) ist $(-x, -y)$, für $(x, y) \neq 0$ gilt $(x, y)^{-1} = \frac{1}{x^2+y^2}(x, -y)$.

Die reellen Zahlen werden folgendermaßen in den komplexen Zahlkörper eingebettet: Die Abbildung

$$\Phi : \mathbb{R} \to \mathbb{C}, \quad x \mapsto (x, 0)$$

ist ein injektiver Körperhomomorphismus, d. h. stets ist

$$\Phi(x + y) = \Phi(x) + \Phi(y) \quad \text{und} \quad \Phi(x \cdot y) = \Phi(x) \cdot \Phi(y).$$

Infolgedessen ist $\Phi(\mathbb{R})$ ein Teilkörper von $\mathbb{C}$, d. h. eine Teilmenge von $\mathbb{C}$ und, mit denselben Verknüpfungen wie $\mathbb{C}$ versehen, ein Körper. Man identifiziert

$$\mathbb{R} = \Phi(\mathbb{R}) = \{(x, 0) : \, x \in \mathbb{R}\}.$$

Für $a \in \mathbb{R}$, $(x, y) \in \mathbb{C}$ gilt $a \cdot (x, y) = (a, 0) \cdot (x, y) = (a \cdot x, a \cdot y)$, die Skalarmultiplikation auf $\mathbb{R}^2$ ist also ein Spezialfall der Multiplikation auf $\mathbb{C}$.

Durch Einführung der *imaginären Einheit* $i := (0,1)$ mit $i^2 = -1$ gelangt man zur üblichen Darstellung komplexer Zahlen

$$(x,y) = (x,0) + (0,y) = x \cdot 1 + y \cdot i = x + iy.$$

Da jetzt i stets die imaginäre Einheit bezeichnet, steht es als Summationsindex nicht mehr zur Verfügung. Deshalb ersetzen wir fortan die bisherigen Standard-Summationsindizes i, j durch μ, ν.

Für $z = x + iy \in \mathbb{C}$ ist $\mathrm{Re}(z) := x$ der *Realteil* und $\mathrm{Im}(z) := y$ der *Imaginärteil* von z. Zu jeder komplexen Zahl $z = x + iy$ wird die *konjugiert komplexe Zahl* $\bar{z} = x - iy$ erklärt. Wegen $\bar{\bar{z}} = z$ ist die Konjugation eine bijektive Abbildung. Ferner gilt $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$ und $\overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2$, die Konjugation ist also ein Körperisomorphismus.

Da reelle Zahlen durch die Konjugation nicht verändert werden, ist diese insbesondere eine bijektive lineare Abbildung des $\mathbb{R}$-Vektorraums $\mathbb{C}$ in sich, eine *umkehrbar* $\mathbb{R}$-*lineare Abbildung*. Sie ist anschaulich eine Spiegelung an der x-Achse. Die Elemente $iy, y \in \mathbb{R}$, der y-Achse heißen *rein imaginär*; es gilt: $\mathbb{R} = \{(x,0) : \ x \in \mathbb{R}\} = \{z : z = \bar{z}\}$ und $i\mathbb{R} := \{iy : \ y \in \mathbb{R}\} = \{(0,y) : \ y \in \mathbb{R}\} = \{z : \ z = -\bar{z}\}$.

Mittels der Konjugation erhält man geschlossene Formeln für Real- und Imaginärteil:

$$\mathrm{Re}(z) = \frac{1}{2}(z + \bar{z}), \quad \mathrm{Im}(z) = \frac{1}{2i}(z - \bar{z}).$$

Betrag einer komplexen Zahl

Für $z = x + iy \in \mathbb{C}$ wird entsprechend der euklidischen Länge in $\mathbb{R}^2$ durch

$$|z| := \sqrt{x^2 + y^2}$$

der Betrag erklärt, offensichtlich ist $|z|^2 = z \cdot \bar{z}$. Damit erhält man für $z \neq 0$ eine einfache Formel zur Bestimmung der komplexen Kehrwerte:

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2}.$$

Es gelten die typischen Eigenschaften eines Betrages, für alle $z, z_1, z_2 \in \mathbb{C}$ ist:

(1) $|z| = 0 \Leftrightarrow z = 0$,

(2) $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$,

(3) $|z_1 + z_2| \leq |z_1| + |z_2|$. (Dreiecksungleichung)

Die erste und die dritte Eigenschaft folgen unmittelbar daraus, daß der Betrag der komplexen Zahl $z = x + iy$ mit der euklidischen Länge des Vektors $(x,y) \in \mathbb{R}^2$ übereinstimmt. Für die multiplikative Eigenschaft berechnet man

$$|z_1 \cdot z_2| = |(x_1 + iy_1) \cdot (x_2 + iy_2)| = |(x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)|$$

$$= \sqrt{x_1^2x_2^2 + y_1^2y_2^2 + x_1^2y_2^2 + x_2^2y_1^2} = \sqrt{(x_1^2 + y_1^2) \cdot (x_2^2 + y_2^2)} = |z_1| \cdot |z_2|.$$

Da Konvergenz einer komplexen Zahlenfolge $(z_k)_{k\in\mathbb{N}} = (x_k + iy_k)_{k\in\mathbb{N}}$ äquivalent zur Konvergenz der beiden reellen Zahlenfolgen $(x_k)_{k\in\mathbb{N}}$ und $(y_k)_{k\in\mathbb{N}}$ ist, erhält man aus der Vollständigkeit von $\mathbb{R}$:

Der Körper $\mathbb{C}$ der komplexen Zahlen ist vollständig.

Im Gegensatz zu $\mathbb{R}$ ist jedoch $\mathbb{C}$ nicht angeordnet, man kann also nicht von „größeren“ und „kleineren“ komplexen Zahlen sprechen.

Komplexe Vektorräume

7.1.3 Definition

Auf einer nichtleeren Menge V sei eine Addition

$$+ : V \times V \to V$$

und eine Skalarmultiplikation

$$\cdot : \mathbb{C} \times V \to V$$

gegeben. Es sei $(V, +)$ eine abelsche Gruppe, es gelten also die Eigenschaften (1)-(4) aus Definition 1.1.3. Ferner gelte für alle $c, d \in \mathbb{C}$ und $x, y \in V$:

(5) $1 \cdot x = x,$ (Unitäres Gesetz)

(6) $(c \cdot d) \cdot x = c \cdot (d \cdot x),$ (Assoziativgesetz)

(7) $(c + d) \cdot x = c \cdot x + d \cdot x$ *und* $c \cdot (x + y) = c \cdot x + c \cdot y.$ (Distributivgesetze)

Dann heißt V ein **komplexer Vektorraum** *oder* **$\mathbb{C}$-Vektorraum**.

Man beachte die nahezu wörtliche Übereinstimmung mit der Definition 1.2.1, es wird nur der Skalarenkörper $\mathbb{R}$ durch $\mathbb{C}$ ersetzt. Mit Ausnahme der Abschnitte, in denen Orthogonalität eine Rolle spielt (d. h. Kapitel 4.5 und Kapitel 6), kann im gesamten Buch $\mathbb{R}$ ohne weiteres durch $\mathbb{C}$ ersetzt werden. Hinsichtlich der Orthogonalität werden wir unten zeigen, daß sich die Definition des Skalarproduktes in geeigneter Weise modifizieren läßt. Zur Vermeidung von Mißverständnissen bezüglich des zugrundeliegenden Skalarenkörpers (man vgl. Satz 7.1.6) spricht man von linearer (Un-) Abhängigkeit *über* $\mathbb{C}$, $\mathbb{C}$*-Dimension* (man schreibt $\dim_{\mathbb{C}}$), usw.

Der n-dimensionale komplexe Zahlenraum

$$\mathbb{C}^n := \{(z_1, \ldots, z_n) : z_\nu \in \mathbb{C}\}$$

ist mit den Verknüpfungen

$$\vec{z} + \vec{w} := (z_1 + w_1, \ldots, z_n + w_n) \quad \text{und} \quad c \cdot \vec{z} := (c \cdot z_1, \ldots, c \cdot z_n)$$

ein komplexer Vektorraum mit der $\mathbb{C}$-Basis

$$\{\vec{e}_1 = (1, 0, \ldots, 0),\ \ldots, \vec{e}_n = (0, \ldots, 0, 1)\}.$$

In derselben Weise, wie $\mathbb{R}$ als Teilmenge von $\mathbb{C}$ aufgefaßt werden kann, ist auch der n-dimensionale *reelle* Zahlenraum $\mathbb{R}^n$ in $\mathbb{C}^n$ enthalten; für $\vec{z} \in \mathbb{C}^n$ schreiben wir $\vec{z} = \vec{x} + i\vec{y}$ mit $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Es gilt wörtlich wie in Satz 3.3.14:

7.1.4 Satz

$\dim_{\mathbb{C}} \mathbb{C}^n = n$.

Da $\mathbb{R}$ ein Teilkörper von $\mathbb{C}$ ist, wird jeder komplexe Vektorraum V durch Einschränkung der Skalarmultiplikation von $\mathbb{C} \times V$ auf $\mathbb{R} \times V$ zu einem reellen Vektorraum. Im Folgenden stellen wir dar, wie die Eigenschaften von V als $\mathbb{R}$-Vektorraum und als $\mathbb{C}$-Vektorraum zusammenhängen.

7.1.5 Satz

Es sei V ein $\mathbb{C}$-Vektorraum und $\{z_1, \ldots, z_n\}$ eine Basis von V über $\mathbb{C}$. Dann ist $\{z_1, i \cdot z_1, \ldots, z_n, i \cdot z_n\}$ eine $\mathbb{R}$-Basis von V.

Beweis

Wir zeigen zunächst die lineare Unabhängigkeit über $\mathbb{R}$. Seien $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$ mit

$$\sum_{\nu=1}^{n} a_\nu z_\nu + \sum_{\nu=1}^{n} b_\nu (i \cdot z_\nu) = 0.$$

Für $c_\nu := a_\nu + ib_\nu$ folgt $\sum_{\nu=1}^{n} c_\nu z_\nu = 0$. Die lineare Unabhängigkeit über $\mathbb{C}$ liefert $c_1 = \cdots = c_n = 0$, also

$$a_1 = \ldots = a_n = b_1 = \ldots = b_n = 0.$$

Andererseits wird V auch von dem angegebenen System über $\mathbb{R}$ erzeugt. Da $\{z_1, \ldots, z_n\}$ über $\mathbb{C}$ ein Erzeugendensystem von V ist, existieren zu beliebigem $z \in V$ komplexe Zahlen $c_\nu = a_\nu + ib_\nu \in \mathbb{C}$, $a_\nu, b_\nu \in \mathbb{R}$, $\nu = 1, \ldots, n$, mit

$$z = \sum_{\nu=1}^{n} c_\nu z_\nu = \sum_{\nu=1}^{n} (a_\nu + ib_\nu) z_\nu = \sum_{\nu=1}^{n} a_\nu z_\nu + \sum_{\nu=1}^{n} b_\nu (iz_\nu). \qquad \square$$

Eine unmittelbare Folgerung ist

7.1.6 Satz

Sei V ein komplexer Vektorraum mit $\dim_{\mathbb{C}} V = n$. Dann gilt $\dim_{\mathbb{R}} V = 2n$.

Die Isomorphie von $\mathbb{C}^n$ und $\mathbb{R}^{2n}$ über $\mathbb{R}$ sieht man am deutlichsten, indem man im n-tupel $\vec{z} = (z_1, \ldots, z_n)$, $z_\nu = x_\nu + iy_\nu$ jede Komponente als Zahlenpaar (x_ν, y_ν) schreibt: $((x_1, y_1), \ldots, (x_n, y_n))$. Gemäß Satz 7.1.5 erhält man eine $\mathbb{R}$-Basis von $\mathbb{C}^n$ durch

$$\begin{aligned}
&\{\vec{e}_1, i \cdot \vec{e}_1, \ldots, \vec{e}_n, i \cdot \vec{e}_n\} \\
&= \Big\{ \Big((1,0), (0,0), \ldots, (0,0)\Big),\ \Big((0,1), (0,0), \ldots, (0,0)\Big), \ldots \\
&\quad \ldots, \Big((0,0), \ldots, (0,0), (1,0)\Big),\ \Big((0,0), \ldots, (0,0), (0,1)\Big) \Big\} .
\end{aligned}$$

Komplex lineare Abbildungen

7.1.7 Definition

Unter einer **komplex linearen** *oder* **$\mathbb{C}$-linearen Abbildung** $F : V \to W$ *zwischen den komplexen Vektorräumen* V *und* W *versteht man eine Abbildung, die mit den Vektorraumoperationen verträglich ist, so daß für alle* $x, y \in V$, $c \in \mathbb{C}$ *gilt:*

$$F(x + y) = F(x) + F(y), \qquad F(c \cdot x) = c \cdot F(x).$$

Sei nun $F : \mathbb{C}^n \to \mathbb{C}^n$ eine $\mathbb{C}$-lineare Abbildung, die durch die komplexe $n \times n$-Matrix

$$C = \begin{pmatrix} c_{11} & \ldots & c_{1n} \\ \vdots & & \vdots \\ c_{n1} & \ldots & c_{nn} \end{pmatrix}, c_{\mu\nu} = a_{\mu\nu} + ib_{\mu\nu},\ a_{\mu\nu}, b_{\mu\nu} \in \mathbb{R},$$

gegeben wird.

Man kann F insbesondere auch als $\mathbb{R}$-lineare Abbildung $\mathbb{R}^{2n} \to \mathbb{R}^{2n}$ auffassen. Wir wollen untersuchen, wie die zugehörige reelle $2n \times 2n$-Matrix A aus C hervorgeht. Dazu sind die Bilder der Einheitsvektoren $\vec{e}_1, i \cdot \vec{e}_1, \ldots, \vec{e}_n, i \cdot \vec{e}_n$ zu bestimmen. Für $\nu = 1, \ldots, n$ gilt:

$$\begin{aligned}
F(\vec{e}_\nu) &= (c_{1\nu}, \ldots, c_{n\nu}) = \Big((a_{1\nu}, b_{1\nu}), \ldots, (a_{n\nu}, b_{n\nu})\Big), \\
F(i\vec{e}_\nu) &= i \cdot (c_{i\nu}, \ldots, c_{n\nu}) = (i \cdot c_{1\nu}, \ldots, i \cdot c_{n\nu}) \\
&= \Big((-b_{1\nu}, a_{1\nu}), \ldots, (-b_{n\nu}, a_{n\nu})\Big),
\end{aligned}$$

und man erhält:

$$A = \begin{pmatrix} a_{11} & -b_{11} & \ldots & a_{1n} & -b_{1n} \\ b_{11} & a_{11} & \ldots & b_{1n} & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & -b_{n1} & \ldots & a_{nn} & -b_{nn} \\ b_{n1} & a_{n1} & \ldots & b_{nn} & a_{nn} \end{pmatrix}.$$

Also geht die reelle Matrix A aus der komplexen Matrix C hervor, indem man jede komplexe Komponente $c_{\mu\nu} = a_{\mu\nu} + ib_{\mu\nu}$ von C durch das reelle 2×2-Kästchen

$$\begin{matrix} a_{\mu\nu} & -b_{\mu\nu} \\ b_{\mu\nu} & a_{\mu\nu} \end{matrix}$$

ersetzt.

Auch die Determinantentheorie für komplexe Matrizen läßt sich genauso wie im reellen Fall durchführen. Insbesondere gilt:

7.1.8 Satz

Es sei $F : \mathbb{C}^n \to \mathbb{C}^n$ eine $\mathbb{C}$-lineare Abbildung mit zugehöriger Matrix $C = C^{n,n}$. Dann sind äquivalent:

(a) Die Abbildung F ist bijektiv.

(b) Die Matrix C ist nichtsingulär.

(c) $\det(C) \neq 0$.

Nach diesen Vorbereitungen definieren wir genauso wie in 3.8.5:

7.1.9 Definition

Die Menge aller nichtsingulären komplexen $n \times n$-Matrizen

$$GL(n, \mathbb{C}) := \{C : \ C \text{ ist komplexe } n \times n\text{-Matrix mit } \det(C) \neq 0\}$$

heißt **allgemeine lineare Gruppe** *vom Rang n über $\mathbb{C}$.*

7.1.10 Bemerkung

Im Zusammenhang mit der Determinantentheorie $\mathbb{C}$-linearer *Abbildungen* ist Vorsicht geboten, da sich hier der Wert der Determinante ändern kann, wenn man zu der Interpretation derselben Abbildung als $\mathbb{R}$-linearer Abbildung zwischen $\mathbb{R}$-Vektorräumen mit doppelt so großer $\mathbb{R}$-Dimension wechselt. Als Beispiel betrachten wir $F : \mathbb{C} \to \mathbb{C}$, $F(z) = -z$; die entsprechende komplexe 1×1-Matrix lautet $C = (\,-1\,)$, $\det(C) = -1$. Zu $F : \mathbb{R}^2 \to \mathbb{R}^2$, $F(x, y) = -(x, y)$ gehört dagegen die reelle 2×2-Matrix $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ mit $\det(A) = 1$.

Ein komplexes Skalarprodukt

Um auch $\mathbb{C}^n$ mit einem Skalarprodukt zu versehen, geht man von der Idee aus, daß $\mathbb{C}^n$ die euklidische Norm $\|\,\vec{z}\,\| = \sqrt{|z_1|^2 + \ldots + |z_n|^2}$ tragen soll und daß diese von dem gesuchten Skalarprodukt erzeugt werden soll: $(\vec{z}, \vec{z}\,) \overset{!}{=} \|\,\vec{z}\,\|^2$. Indem man noch das bereits erwähnte Gesetz $|z_\nu|^2 = \bar{z}_\nu \cdot z_\nu$ beachtet, gelangt man zu der Setzung

$$(\vec{z}, \vec{w}) := \sum_{\nu=1}^{n} \bar{z}_\nu \cdot w_\nu .$$

Es stellt sich heraus, daß die axiomatische Beschreibung des Skalarproduktes in $\mathbb{C}^n$ oder allgemeiner in komplexen Vektorräumen gegenüber der Charakterisierung in Satz 6.2.1 für den $\mathbb{R}^n$ bzw. für reelle Vektorräume hinsichtlich der Linearitäts- und Symmetrieeigenschaften zu modifizieren ist.

7.1.11 Satz

Für das durch $(\vec{z}, \vec{w}) = \sum_{\nu=1}^{n} \bar{z}_\nu \cdot w_\nu$ *auf* $\mathbb{C}^n$ *gegebene* **kanonische komplexe Skalarprodukt** *gilt:*

(1) Es ist **konjugiert linear** *in der ersten Komponente, d. h. für* $c \in \mathbb{C}$, $\vec{z}, \vec{z}\,', \vec{w} \in \mathbb{C}^n$ *gilt:*

$$(\vec{z} + \vec{z}\,', \vec{w}) = (\vec{z}, \vec{w}) + (\vec{z}\,', \vec{w}), \quad (c \cdot \vec{z}, \vec{w}) = \bar{c} \cdot (\vec{z}, \vec{w}).$$

(2) Es ist linear in der zweiten Komponente, d. h. für $c \in \mathbb{C}$, $\vec{z}, \vec{w}, \vec{w}\,' \in \mathbb{C}^n$ *gilt:*

$$(\vec{z}, \vec{w} + \vec{w}\,') = (\vec{z}, \vec{w}) + (\vec{z}, \vec{w}\,'), \quad (\vec{z}, c \cdot \vec{w}) = c \cdot (\vec{z}, \vec{w}).$$

(3) Es ist **Hermitesch** *(konjugiert symmetrisch), d. h. für* $\vec{z}, \vec{w} \in \mathbb{C}^n$ *gilt:*

$$(\vec{z}, \vec{w}) = \overline{(\vec{w}, \vec{z})}.$$

(4) Es ist **positiv definit**, *d. h. für* $\vec{z} \in \mathbb{C}^n$ *ist* $(\vec{z}, \vec{z})$ *reell, und es gilt:*

$$(\vec{z}, \vec{z}) \geq 0 \quad \textit{und} \quad (\vec{z}, \vec{z}) > 0, \ \textit{falls } \vec{z} \neq 0.$$

Beweis

Der Beweis folgt durch elementares Nachrechnen aus der Definition des Skalarprodukts. □

Abbildungen $\mathbb{C}^n \oplus \mathbb{C}^n \to \mathbb{C}$ mit den Eigenschaften (1) und (2) nennt man *sesquilinear*. Das kanonische komplexe Skalarprodukt ist also eine Hermitesche, positiv definite *Sesquilinearform*.

Es wird in der Literatur nicht einheitlich gehandhabt, ob das Skalarprodukt in der ersten oder in der zweiten Komponente als konjugiert linear definiert wird. Häufig wird z. B. das Skalarprodukt in $\mathbb{C}^n$ auch durch $\sum_{\nu=1}^{n} z_\nu \bar{w}_\nu$ eingeführt.

7.1.12 Satz

Durch die Setzung

$$\| \vec{z} \| := \sqrt{(\vec{z}, \vec{z})} = \sqrt{\sum_{\nu=1}^{n} \bar{z}_\nu \cdot z_\nu}$$

wird $\mathbb{C}^n$ *zu einem* **normierten komplexen Vektorraum**. *Die Abbildung* $\| \, . \, \| : \mathbb{C}^n \to \mathbb{R}$ *hat also die typischen Eigenschaften einer Norm; für alle* $c \in \mathbb{C}$, $\vec{z}, \vec{w} \in \mathbb{C}^n$ *gilt:*

(1) $\| \vec{z} \| = 0 \Leftrightarrow \vec{z} = 0$,

(2) $\| c \cdot \vec{z} \| = |c| \cdot \| \vec{z} \|$,

(3) $\| \vec{z} + \vec{w} \| \leq \| \vec{z} \| + \| \vec{w} \|$.

Beweis

Man könnte zunächst die auch in $\mathbb{C}^n$ gültige Cauchy-Schwarzsche Ungleichung

$$|(\vec{z}, \vec{w})| \leq \| \vec{z} \| \cdot \| \vec{w} \|$$

herleiten und anschließend genauso wie im Beweis von Satz 6.2.4 verfahren.

Hier wollen wir allerdings direkt auf das dort erzielte Resultat zurückgreifen und verwenden, daß $\mathbb{C}^n$ und $\mathbb{R}^{2n}$ als $\mathbb{R}$-Vektorräume kanonisch isomorph sind. Wie bereits vorher bemerkt, stimmen für

$$\vec{z} = (z_1, \dots, z_n) = (x_1 + iy_1, \dots, x_n + iy_n) = ((x_1, y_1), \dots, (x_n, y_n))$$

die euklidische Norm in $\mathbb{R}^{2n}$ und die euklidische Norm in $\mathbb{C}^n$ überein:

$$\| (x_1, y_1, \dots, x_n, y_n) \|^2 = \sum_{\nu=1}^{n} \left(x_\nu^2 + y_\nu^2 \right) = \sum_{\nu=1}^{n} \bar{z}_\nu \cdot z_\nu = \| \vec{z} \|^2.$$

Die Eigenschaften (1) und (3) folgen also direkt aus dem reellen Resultat. Für $c \in \mathbb{C}$, $\vec{z} \in \mathbb{C}^n$ gilt weiter:

$$\| c \cdot \vec{z} \|^2 = (c \cdot \vec{z}, c \cdot \vec{z}) = \bar{c} \cdot (\vec{z}, c \cdot \vec{z}) = \bar{c} \cdot c \cdot (\vec{z}, \vec{z}) = |c|^2 \cdot \| \vec{z} \|^2. \qquad \square$$

7.2 Unitäre und adjungierte Abbildungen

Die unitäre Gruppe

Die unitären Abbildungen sind die komplexen Verallgemeinerungen der orthogonalen Abbildungen. Zwar könnte man die folgende Theorie auch allgemein in endlichdimensionalen Vektorräumen mit Skalarprodukt (diese nennt man *unitär* bzw. im reellen Fall *euklidisch*) durchführen, wir beschränken uns aber der Einfachheit halber auf Abbildungen $\mathbb{C}^n \to \mathbb{C}^n$. Zudem kann man sich, ganz ähnlich wie bei allgemeinen linearen Abbildungen in Kapitel 3, durch die Einführung *unitärer* Karten stets auf diesen Fall zurückziehen.

Zur Definition wählen wir eine der Eigenschaften aus Satz 6.3.2.

7.2.1 Definition

Eine komplexe $n \times n$-Matrix C heißt **unitär**, *falls $\| C \circ \vec{z} \| = \| \vec{z} \|$ für alle $\vec{z} \in \mathbb{C}^n$ gilt. Eine lineare Abbildung $F : \mathbb{C}^n \to \mathbb{C}^n$ heißt* **unitär**, *falls die zugehörige Matrix unitär ist. Wir nennen*

$$U(n) := \{C : C \text{ ist unitäre } n \times n\text{-Matrix}\}$$

die **unitäre Gruppe** *vom Rang n.*

Tatsächlich ist die Bezeichnung „unitäre Gruppe“ gerechtfertigt:

7.2.2 Satz

$U(n)$ ist eine Untergruppe von $GL(n, \mathbb{C})$.

Beweis

Offenbar ist die zu $C \in U(n)$ gehörige lineare Abbildung injektiv. Wie in Kapitel 3 folgt daraus, daß C nichtsingulär ist und $C^{-1} \in GL(n, \mathbb{C})$ existiert. Es folgt $U(n) \subset GL(n, \mathbb{C})$ und auf Grund von $\| \vec{z} \| = \| C \circ C^{-1} \circ \vec{z} \| = \| C^{-1} \circ \vec{z} \|$ auch $C^{-1} \in U(n)$. Wegen $E^{n,n} \in U(n)$ ist $U(n) \neq \emptyset$. Schließlich ist $U(n)$ auch gegenüber der Matrizenmultiplikation abgeschlossen, denn mit $C_1, C_2 \in U(n)$ ist wegen $\| C_1 \circ C_2 \circ \vec{z} \| = \| C_2 \circ \vec{z} \| = \| \vec{z} \|$ auch $C_1 \circ C_2 \in U(n)$. □

Wir studieren zunächst die Verbindungen mit dem reellen Begriff der orthogonalen Matrix. Eine reelle Matrix $C \in \mathbb{R}^{n \times n}$ ist natürlich auch eine komplexe Matrix $C \in \mathbb{C}^{n \times n}$ und vermittelt damit sowohl Abbildungen $F_{\mathbb{R}} : \mathbb{R}^n \to \mathbb{R}^n$ als auch $F_{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n$. Die komplexe Abbildung $F_{\mathbb{C}}$ geht dabei aus $F_{\mathbb{R}}$ hervor, indem man für $\vec{z} = \vec{x} + i\vec{y} \in \mathbb{C}^n$, $\vec{x}, \vec{y} \in \mathbb{R}^n$ beachtet:

$$F_{\mathbb{C}}(\vec{z}\,) = F_{\mathbb{C}}(\vec{x} + i\vec{y}\,) = F_{\mathbb{C}}(\vec{x}\,) + iF_{\mathbb{C}}(\vec{y}\,) = C \circ \vec{x} + i\, C \circ \vec{y} = F_{\mathbb{R}}(\vec{x}\,) + iF_{\mathbb{R}}(\vec{y}\,).$$

Dieses Verfahren, das man auch unabhängig vom Matrizenkalkül durchführen kann, heißt *Komplexifizierung*.

7.2.3 Satz

$O(n) \subset U(n)$.

Beweis

Sei $C \in O(n)$ und $\vec{z} = \vec{x} + i\vec{y}$, $\vec{x}, \vec{y} \in \mathbb{R}^n$. Die Eigenschaften der euklidischen Norm in $\mathbb{C}^n$ ergeben:

$$\begin{aligned} \| \vec{z} \|^2 &= \| \vec{x} \|^2 + \| \vec{y} \|^2 = \| C \circ \vec{x} \|^2 + \| C \circ \vec{y} \|^2 \\ &= \| C \circ \vec{x} + i\, C \circ \vec{y} \|^2 = \| C \circ (\vec{x} + i\vec{y}\,) \|^2 = \| C \circ \vec{z} \|^2. \end{aligned}$$

□

Wie bereits im vorhergehenden Abschnitt erläutert, können wir $\mathbb{C}^n$ und $\mathbb{R}^{2n}$ als $\mathbb{R}$-Vektorräume identifizieren und entsprechend $\mathbb{C}$-lineare Abbildungen $\mathbb{C}^n \to \mathbb{C}^n$ als $\mathbb{R}$-lineare Abbildungen $\mathbb{R}^{2n} \to \mathbb{R}^{2n}$ auffassen. In diesem Sinne ist der folgende Satz zu verstehen.

7.2.4 Satz

$U(n)$ ist eine Untergruppe von $O(2n)$.

Beweis

Es sei $\Phi : U(n) \to \mathbb{R}^{2n\times 2n}$ die in Abschnitt 7.1 beschriebene Abbildung: Für $C = (c_{\mu\nu})_{\mu,\nu=1,\ldots,n}$ erhält man $A = \Phi(C)$, indem man jede komplexe Komponente $c_{\mu\nu} = a_{\mu\nu} + ib_{\mu\nu}$ durch das reelle 2×2-Kästchen $\begin{pmatrix} a_{\mu\nu} & -b_{\mu\nu} \\ b_{\mu\nu} & a_{\mu\nu} \end{pmatrix}$ ersetzt.

Wir haben also zu zeigen, daß $\Phi : U(n) \to O(2n)$ ein injektiver Gruppenhomomorphismus ist.

Da C und $A = \Phi(C)$ dieselbe Abbildung beschreiben und die euklidische Norm in $\mathbb{C}^n$ mit der in $\mathbb{R}^{2n}$ übereinstimmt, folgt $\Phi(U(n)) \subset O(2n)$ aus Satz 6.3.2.

Wir zeigen nun, daß Φ ein Gruppenhomomorphismus ist. Seien $C_1, C_2 \in U(n)$. Die durch $C_1 \circ C_2$ gegebene Abbildung $\mathbb{C}^n \to \mathbb{C}^n$ stimmt mit der durch $\Phi(C_1) \circ \Phi(C_2)$ gegebenen Abbildung $\mathbb{R}^{2n} \to \mathbb{R}^{2n}$ überein, denn die Komposition der Abbildungen ist unabhängig davon, ob man $\mathbb{C}^n$ als komplexen oder reellen Vektorraum interpretiert. Also gilt $\Phi(C_1 \circ C_2) = \Phi(C_1) \circ \Phi(C_2)$.

Die Injektivität von Φ ergibt sich schließlich aus

$$\Phi(C) = E^{2n,2n} \quad \Rightarrow \quad C = E^{n,n}. \qquad \square$$

Im Spezialfall $n = 1$ läßt sich Genaueres aussagen:

7.2.5 Satz

$U(1) \cong SO(2)$.

Beweis

Für die im vorherigen Beweis betrachtete Abbildung $\Phi : U(1) \to O(2)$ bleibt zu zeigen:

$$\Phi(U(1)) \subset SO(2), \quad \Phi : U(1) \to SO(2) \text{ ist surjektiv.}$$

Sei $c \in U(1)$, es ist also insbesondere $|c|^2 = |c \cdot \bar{c}| = |\bar{c}| = |c|$ und damit $|c| = 1$. Mit $c = a + ib$ gilt $a^2 + b^2 = 1$ und $\Phi(c) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$, also $\Phi(c) \in SO(2)$.

Um noch die Surjektivität von Φ zu zeigen, sei $A \in SO(2)$ gegeben. Gemäß Satz 6.1.8 existieren dann $a, b \in \mathbb{R}$ mit $a^2 + b^2 = 1$ und $A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. Für $c := a + ib$ gilt

$|c| = 1$ und damit für alle $d \in \mathbb{C}$: $|c \cdot d| = |c| \cdot |d| = |d|$, d. h. $c \in U(1)$. Ferner ist $\Phi(c) = A$ und folglich $\mathrm{im}(\Phi) = SO(2)$. □

Eine Drehung des $\mathbb{R}^2$ ist also die Multiplikation mit einer komplexen Zahl c, deren Betrag $|c| = 1$ ist.

Das nächste Ziel ist die komplexe Entsprechung des Satzes 6.3.2, es sollen also einige äquivalente und für das Folgende grundlegende Charakterisierungen unitärer Matrizen zusammengestellt werden. Dazu sind vorbereitend einige aus Kapitel 6 bekannte Begriffe auf den $\mathbb{C}^n$ auszudehnen und einige Hilfsresultate zu beweisen.

Komplexe Orthogonalität

7.2.6 Definition

Die Vektoren $\vec{z}, \vec{w} \in \mathbb{C}^n$ heißen **(komplex) orthogonal**, *wenn $(\vec{z}, \vec{w}) = 0$ gilt. Unter dem* **orthogonalen Komplement** *des Untervektorraums $V \subset \mathbb{C}^n$ verstehen wir*

$$V^{\perp} := \{\vec{z} \in \mathbb{C}^n : \ (\vec{z}, \vec{w}) = 0 \text{ für alle } \vec{w} \in V\}\,.$$

7.2.7 Satz

Die Vektoren $\vec{z}, \vec{w} \in \mathbb{C}^n$ seien (komplex) orthogonal, $\vec{z} = \vec{x} + i\vec{y}$, $\vec{w} = \vec{u} + i\vec{v}$ mit $\vec{x}$, $\vec{y}$, $\vec{u}$, $\vec{v} \in \mathbb{R}^n$. Dann sind die entsprechenden Vektoren $(x_1, y_1, \ldots, x_n, y_n)$, $(u_1, v_1, \ldots, u_n, v_n) \in \mathbb{R}^{2n}$ (reell) orthogonal.

Beweis

Aus der komplexen Orthogonalität folgt

$$\begin{aligned} 0 = (\vec{z}, \vec{w}) &= \sum_{\nu=1}^{n} (x_\nu - i\, y_\nu)(u_\nu + i\, v_\nu) \\ &= \sum_{\nu=1}^{n} (x_\nu u_\nu + y_\nu v_\nu) + i \cdot \sum_{\nu=1}^{n} (x_\nu v_\nu - y_\nu u_\nu) \end{aligned}$$

und damit durch Betrachtung des Realteils die Behauptung. □

Der Beweis zeigt, daß komplexe Orthogonalität in $\mathbb{C}^n$ eine wesentlich stärkere Eigenschaft als reelle Orthogonalität in $\mathbb{R}^{2n}$ ist, denn das entsprechende reelle Skalarprodukt ist lediglich der Realteil des komplexen Skalarprodukts. Dazu geben wir folgendes einfache Beispiel:

7.2.8 Beispiel

Die komplexen Zahlen 1 und i sind in $\mathbb{C}$ natürlich nicht orthogonal, die entsprechenden reellen Vektoren $(1, 0)$ und $(0, 1)$ in $\mathbb{R}^2$ hingegen schon.

7.2.9 Satz

Für jeden Untervektorraum $V \subset \mathbb{C}^n$ gilt

$$V^{\perp\perp} := \left(V^{\perp}\right)^{\perp} = V.$$

Beweis

Offensichtlich ist $\left(V^{\perp}\right)^{\perp} = \{\vec{z} \in \mathbb{C}^n : \ (\vec{z}, \vec{w}) = 0 \text{ für alle } \vec{w} \in V^{\perp}\} \supset V$.

Die umgekehrte Inklusion beweist man mit Hilfe eines Dimensionsarguments. Sei $k = \dim_{\mathbb{C}} V$, wie in Satz 6.3.6 zeigt man $\dim_{\mathbb{C}} V^{\perp} = n - k$. Eine Wiederholung dieses Schlusses ergibt $\dim_{\mathbb{C}} \left(V^{\perp}\right)^{\perp} = n - (n-k) = k$ und damit wegen $V \subset V^{\perp\perp}$ die behauptete Gleichheit. □

In unendlichdimensionalen (unitären) Vektorräumen kann das Dimensionsargument nicht verwendet werden. Tatsächlich behält der obige Satz dort nur unter zusätzlichen Voraussetzungen seine Gültigkeit.

7.2.10 Definition

Man nennt ein k-tupel $(\vec{z}_1, \ldots, \vec{z}_k)$ von Vektoren aus $\mathbb{C}^n$ ein **(komplexes) orthogonales k-Bein**, *wenn für $\mu, \nu = 1, \ldots, k$ gilt:*

$$(\vec{z}_\mu, \vec{z}_\nu) = \delta_{\mu\nu}.$$

Ist ein orthogonales k-Bein gleichzeitig auch Erzeugendensystem eines Untervektorraums V von $\mathbb{C}^n$, so spricht man von einer **Orthonormalbasis** *von V.*

Eine Orthonormalbasis von V ist also eine Basis aus paarweise orthogonalen, normierten (d. h. mit Länge 1) Vektoren.

Wir bezeichnen mit $\bar{C}^t$ die zur $m \times n$-Matrix C konjugiert transponierte $n \times m$-Matrix, die also aus $C = (c_{\mu\nu})_{\substack{\mu=1,\ldots,m \\ \nu=1,\ldots,n}}$ durch Transponieren und komponentenweises Konjugieren hervorgeht:

$$\bar{C}^t = \begin{pmatrix} \bar{c}_{11} & \bar{c}_{21} & \ldots & \bar{c}_{m1} \\ \bar{c}_{12} & \bar{c}_{22} & \ldots & \bar{c}_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{c}_{1n} & \bar{c}_{2n} & \ldots & \bar{c}_{mn} \end{pmatrix}.$$

Wir betrachten nun wieder ausschließlich quadratische Matrizen.

7.2.11 Satz

Für jede $n \times n$-Matrix C und alle Vektoren $\vec{z}, \vec{w} \in \mathbb{C}^n$ gilt:

$$(\vec{w}, C \circ \vec{z}) = (\bar{C}^t \circ \vec{w}, \vec{z}).$$

Beweis

Das komplexe Skalarprodukt lautet in Matrizenschreibweise:

$$(\vec{z}, \vec{w}) = \bar{\vec{z}}^{\,t} \circ \vec{w},$$

daraus erhält man:

$$(\bar{C}^t \circ \vec{w}, \vec{z}\,) = \overline{(\bar{C}^t \circ \vec{w})}^{\,t} \circ \vec{z} = (\bar{\bar{C}}^t \circ \bar{\vec{w}})^t \circ \vec{z} = \bar{\vec{w}}^{\,t} \circ (C \circ \vec{z}\,) = (\vec{w}, C \circ \vec{z}\,). \qquad \square$$

7.2.12 Satz

Ist D eine komplexe $n \times n$-Matrix derart, daß $\bar{\vec{z}}^{\,t} \circ D \circ \vec{z} = 0$ für alle $\vec{z} \in \mathbb{C}^n$ gilt, dann folgt: $D = \mathcal{O}^{n,n}$.

Beweis

Wir zeigen zunächst, daß die Sesquilinearform, die je zwei Vektoren $\vec{z}, \vec{w} \in \mathbb{C}^n$ die komplexe Zahl $(\vec{z}, D \circ \vec{w}) = \bar{\vec{z}}^{\,t} \circ D \circ \vec{w}$ zuordnet, bereits durch die Werte der quadratischen Form $\vec{z} \mapsto \bar{\vec{z}}^{\,t} \circ D \circ \vec{z}$ bestimmt ist. Bei dem folgenden Verfahren der *Polarisierung* setzt man geeignete Kombinationen der beliebig vorgegebenen Vektoren $\vec{z}, \vec{w} \in \mathbb{C}^n$ ein:

$$\begin{aligned}
\overline{(\vec{z} + \vec{w})}^{\,t} \circ D \circ (\vec{z} + \vec{w}) &= \bar{\vec{z}}^{\,t} \circ D \circ \vec{z} + \bar{\vec{z}}^{\,t} \circ D \circ \vec{w} + \bar{\vec{w}}^{\,t} \circ D \circ \vec{z} + \bar{\vec{w}}^{\,t} \circ D \circ \vec{w},\\
\overline{(\vec{z} - \vec{w})}^{\,t} \circ D \circ (\vec{z} - \vec{w}) &= \bar{\vec{z}}^{\,t} \circ D \circ \vec{z} - \bar{\vec{z}}^{\,t} \circ D \circ \vec{w} - \bar{\vec{w}}^{\,t} \circ D \circ \vec{z} + \bar{\vec{w}}^{\,t} \circ D \circ \vec{w},\\
\overline{(\vec{z} + i\vec{w})}^{t} \circ D \circ (\vec{z} + i\vec{w}) &= \bar{\vec{z}}^{\,t} \circ D \circ \vec{z} + i\bar{\vec{z}}^{\,t} \circ D \circ \vec{w} - i\bar{\vec{w}}^{\,t} \circ D \circ \vec{z} + \bar{\vec{w}}^{\,t} \circ D \circ \vec{w},\\
\overline{(\vec{z} - i\vec{w})}^{t} \circ D \circ (\vec{z} - i\vec{w}) &= \bar{\vec{z}}^{\,t} \circ D \circ \vec{z} - i\bar{\vec{z}}^{\,t} \circ D \circ \vec{w} + i\bar{\vec{w}}^{\,t} \circ D \circ \vec{z} + \bar{\vec{w}}^{\,t} \circ D \circ \vec{w}.
\end{aligned}$$

Indem man jeweils die ersten und letzten beiden Gleichungen voneinander subtrahiert, erhält man:

$$\begin{aligned}
&\bar{\vec{z}}^{\,t} \circ D \circ \vec{w} + \bar{\vec{w}}^{\,t} \circ D \circ \vec{z}\\
&= \frac{1}{2}\left(\overline{(\vec{z} + \vec{w})}^{\,t} \circ D \circ (\vec{z} + \vec{w}) - \overline{(\vec{z} - \vec{w})}^{\,t} \circ D \circ (\vec{z} - \vec{w})\right);\\
&\bar{\vec{z}}^{\,t} \circ D \circ \vec{w} - \bar{\vec{w}}^{\,t} \circ D \circ \vec{z}\\
&= \frac{1}{2i}\left(\overline{(\vec{z} + i\vec{w})}^{\,t} \circ D \circ (\vec{z} + i\vec{w}) - \overline{(\vec{z} - i\vec{w})}^{\,t} \circ D \circ (\vec{z} - i\vec{w})\right).
\end{aligned}$$

Diese Gleichungen werden zueinander addiert, und es folgt:

$$\begin{aligned}
\bar{\vec{z}}^{\,t} \circ D \circ \vec{w} = {}& \frac{1}{4}\left(\overline{(\vec{z} + \vec{w})}^{\,t} \circ D \circ (\vec{z} + \vec{w}) - \overline{(\vec{z} - \vec{w})}^{\,t} \circ D \circ (\vec{z} - \vec{w})\right)\\
&+ \frac{1}{4i}\left(\overline{(\vec{z} + i\vec{w})}^{\,t} \circ D \circ (\vec{z} + i\vec{w}) - \overline{(\vec{z} - i\vec{w})}^{\,t} \circ D \circ (\vec{z} - i\vec{w})\right).
\end{aligned}$$

Wir setzen nun in dieses auch an sich interessante Zwischenergebnis unsere Voraussetzung ein, daß stets $\bar{\vec{z}}^{\,t} \circ D \circ \vec{z} = 0$ ist. Dadurch erhalten wir für alle $\vec{z}, \vec{w} \in \mathbb{C}^n$:

$$\bar{\vec{z}}^{\,t} \circ D \circ \vec{w} = 0.$$

Indem wir für $\vec{z}$ und $\vec{w}$ die Elemente $\vec{e}_\mu$ und $\vec{e}_\nu$ der kanonischen Basis des $\mathbb{C}^n$ einsetzen, folgt schließlich für $D = (d_{\mu\nu})_{\mu,\nu=1,\ldots,n}$:

$$d_{\mu\nu} = \vec{e}_\mu^{\,t} \circ D \circ \vec{e}_\nu = 0. \qquad \square$$

7.2.13 Bemerkung

Die komplexe $n \times n$-Matrix D und damit die entsprechende Sesquilinearform $S : \mathbb{C}^n \oplus \mathbb{C}^n \to \mathbb{C},\ S(\vec{z}, \vec{w}) = \bar{\vec{z}}^t \circ D \circ \vec{w} = (\vec{z}, D \circ \vec{w})$ sind durch die quadratische Form $\vec{z} \mapsto \bar{\vec{z}}^t \circ D \circ \vec{z}$, d. h. durch die Werte der Sesquilinearform auf der „Diagonalen" $\vec{z} = \vec{w}$ eindeutig bestimmt.

Ein entsprechendes Resultat gilt im Reellen nur unter Zusatzvoraussetzungen, wie etwa der Symmetrie (man vgl. den Beweis von Satz 6.3.2), wie das Beispiel $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ zeigt.

Unitäre Matrizen

7.2.14 Satz

Für alle komplexen $n \times n$-Matrizen $A = (\vec{a}_1, \ldots, \vec{a}_n)$ sind äquivalent:

(a) $\bar{A}^t \circ A = E$;

(b) *für alle* $\vec{z}, \vec{w} \in \mathbb{C}$ *gilt* $(A \circ \vec{z}, A \circ \vec{w}) = (\vec{z}, \vec{w})$;

(c) A *ist unitär;*

(d) $(\vec{a}_1, \ldots, \vec{a}_n)$ *ist ein komplexes orthogonales n-Bein.*

Beweis

„(a) $\Rightarrow$ (c)": Für alle $\vec{z} \in \mathbb{C}^n$ gilt in Matrizenschreibweise:

$$\begin{aligned} \|\,\vec{z}\,\|^2 &= \bar{\vec{z}}^t \circ \vec{z} = \bar{\vec{z}}^t \circ E \circ \vec{z} = (\bar{\vec{z}}^t \circ \bar{A}^t) \circ (A \circ \vec{z}\,) \\ &= \overline{(A \circ \vec{z}\,)}^{\,t} \circ (A \circ \vec{z}\,) = \|\,A \circ \vec{z}\,\|^2. \end{aligned}$$

„(c) $\Rightarrow$ (b)": In diesem Beweisteil können wir von den in Satz 7.2.12 geleisteten Vorarbeiten profitieren. Wir wollen zeigen, daß die Sesquilinearform

$$S : \mathbb{C}^n \oplus \mathbb{C}^n \to \mathbb{C},\ S(\vec{z}, \vec{w}) = (A \circ \vec{z}, A \circ \vec{w}) - (\vec{z}, \vec{w}) \qquad (*)$$

stets den Wert 0 ergibt. Gemäß Voraussetzung (c) ist der Wert dieser Form auf der „Diagonalen" $\vec{z} = \vec{w}$ stets 0. Indem man noch beachtet, daß für alle $\vec{z}, \vec{w} \in \mathbb{C}^n$

$$\begin{aligned} (A \circ \vec{z}, A \circ \vec{w}) - (\vec{z}, \vec{w}) &= \bar{\vec{z}}^t \circ \big(\bar{A}^t \circ A\big) \circ \vec{w} - \bar{\vec{z}}^t \circ \vec{w} \\ &= \bar{\vec{z}}^t \circ \big(\bar{A}^t \circ A - E\big) \circ \vec{w} \end{aligned}$$

gilt, daß also die Sesquilinearform S in $(*)$ durch die Matrix $\bar{A}^t \circ A - E$ gegeben wird, ergibt Satz 7.2.12 tatsächlich, daß es sich bei dieser Matrix um die Nullmatrix handelt. Mithin verschwindet die Sesquilinearform S in $(*)$ identisch, und wir haben gleichzeitig auch „(c) ⇒ (a)“ mitbewiesen.

„(b) ⇒ (d)“: Gemäß Voraussetzung (b) ist die Matrix A „längen-“ und „winkeltreu“, das gilt insbesondere auch für die Elemente der kanonischen Basis des $\mathbb{C}^n$:

$$\delta_{\mu\nu} = (\vec{e}_\mu, \vec{e}_\nu) = (A \circ \vec{e}_\mu, A \circ \vec{e}_\nu) = (\vec{a}_\mu, \vec{a}_\nu).$$

„(d) ⇒ (a)“: Dieser Schluß basiert allein auf der Definition der Matrizenmultiplikation:

$$\begin{aligned}\bar{A}^t \circ A &= \begin{pmatrix} \bar{\vec{a}}_1{}^t \\ \vdots \\ \bar{\vec{a}}_n{}^t \end{pmatrix} \circ (\vec{a}_1, \ldots, \vec{a}_n) = \left(\bar{\vec{a}}_\mu{}^t \circ \vec{a}_\nu\right)_{\mu,\nu=1,\ldots,n} \\ &= \Big((\vec{a}_\mu, \vec{a}_\nu)\Big)_{\mu,\nu=1,\ldots,n} = (\delta_{\mu\nu})_{\mu,\nu=1,\ldots,n} = E.\end{aligned}$$

□

Adjungierte Abbildungen

Wiederholt hat die zur komplexen $n \times n$-Matrix A konjugiert transponierte Matrix $\bar{A}^t$ eine Rolle gespielt, nicht zuletzt als Umkehrmatrix unitärer Matrizen. Deshalb werden wir nun die entsprechenden Abbildungen eingehender studieren.

7.2.15 Definition

Sei A eine komplexe $n \times n$-Matrix und $F : \mathbb{C}^n \to \mathbb{C}^n$ die zugehörige $\mathbb{C}$-lineare Abbildung.

(a) *Die zur konjugiert transponierten Matrix $\bar{A}^t$ gehörige $\mathbb{C}$-lineare Abbildung $\mathbb{C}^n \to \mathbb{C}^n$ heißt die zu F* **adjungierte Abbildung**.

(b) *Gilt $\bar{A}^t = A$, so heißt A* **Hermitesch** *und F* **selbstadjungiert**.

Ein Spezialfall hiervon ist die entsprechende reelle Definition:

7.2.16 Definition

Sei A eine reelle $n \times n$-Matrix und $F : \mathbb{R}^n \to \mathbb{R}^n$ die zugehörige $\mathbb{R}$-lineare Abbildung.

(a) *Die zur transponierten Matrix A^t gehörige $\mathbb{R}$-lineare Abbildung $\mathbb{R}^n \to \mathbb{R}^n$ heißt die zu F* **adjungierte Abbildung**.

(b) *Ist A symmetrisch, d. h. $A = A^t$, so nennen wir F* **selbstadjungiert**.

7.2.17 Bemerkung

Bei reellen Matrizen $A \in \mathbb{R}^{n\times n}$ und zugehörigen linearen Abbildungen $F : \mathbb{R}^n \to \mathbb{R}^n$ fallen die adjungierte und die duale Abbildung zusammen. Im Komplexen dagegen wird die duale Abbildung F^* unverändert durch A^t, die adjungierte Abbildung $F^{\text{adjungiert}}$ dagegen durch $\bar{A}^t$ gegeben. Hier besteht bei den Abbildungen der Zusammenhang: $F^{\text{adjungiert}}(\vec{z}\,) = \overline{F^*\left(\bar{\vec{z}}\right)}$.

Um im Folgenden nicht unnötig komplizierte Bezeichnungen verwenden zu müssen, werden wir mitunter die zur Matrix A gehörige lineare Abbildung ebenfalls mit A bezeichnen und beispielsweise von $\bar{A}^t$ als der zu A adjungierten Abbildung sprechen.

7.2.18 Satz

Eine komplexe $n \times n$-Matrix A ist genau dann Hermitesch, wenn für alle $\vec{z} \in \mathbb{C}^n$ die Zahl $(\vec{z}, A \circ \vec{z}\,)$ reell ist.

Beweis

Grundlegend ist die folgende, auf den Eigenschaften konjugiert transponierter Matrizen beruhende Gleichheit:

$$(\vec{z}, \bar{A}^t \circ \vec{z}\,) = (A \circ \vec{z}, \vec{z}\,) = \overline{(\vec{z}, A \circ \vec{z}\,)}.$$

Ist A Hermitesch, so gilt für alle $\vec{z} \in \mathbb{C}^n$ also $(\vec{z}, A \circ \vec{z}\,) = \overline{(\vec{z}, A \circ \vec{z}\,)}$, demgemäß sind diese Zahlen stets reell.

Ist dagegen $(\vec{z}, A \circ \vec{z}\,)$ stets reell, d. h. $(\vec{z}, A \circ \vec{z}\,) = \overline{(\vec{z}, A \circ \vec{z}\,)}$, so verschwindet die von $\bar{A}^t - A$ gegebene quadratische Form identisch, es ist also stets $(\vec{z}, (\bar{A}^t - A) \circ \vec{z}\,) = 0$. Der schon wiederholt zitierte Satz 7.2.12 zeigt $\bar{A}^t - A = \mathcal{O}$ und damit die Hermitizität von A. □

Auch um die Symmetrie einer *reellen* $n \times n$-Matrix zu prüfen, ist die quadratische Form $(\vec{z}, A \circ \vec{z}\,)$ für alle *komplexen* Vektoren $\vec{z} \in \mathbb{C}^n$ zu betrachten.

7.2.19 Satz

Für die zur komplexen $n \times n$-Matrix A gehörige lineare Abbildung gilt:

(a) $\ker(A) = \operatorname{im}(\bar{A}^t)^\perp$,

(b) $\operatorname{im}(A) = \ker(\bar{A}^t)^\perp$.

Beweis

Sei $\vec{z} \in \ker(A)$, d. h. $A \circ \vec{z} = \vec{0}$. Für alle $\bar{A}^t \circ \vec{w} \in \operatorname{im}(\bar{A}^t)$, $\vec{w} \in \mathbb{C}^n$, folgt mit Satz 7.2.11

$$(\vec{z}, \bar{A}^t \circ \vec{w}) = (A \circ \vec{z}, \vec{w}) = (\vec{0}, \vec{w}) = 0.$$

Somit ist $\ker(A) \subset \operatorname{im}(\bar{A}^t)^\perp$.

Sei nun umgekehrt $\vec{z} \in \mathrm{im}(\bar{A}^t)^\perp$, für alle $\vec{w} \in \mathbb{C}^n$ haben wir also

$$0 = (\vec{z}, \bar{A}^t \circ \vec{w}) = (A \circ \vec{z}, \vec{w}).$$

Indem man speziell $\vec{w} = A \circ \vec{z}$ wählt, ergibt sich $A \circ \vec{z} = \vec{0}$, $\vec{z} \in \ker(A)$, $\mathrm{im}(\bar{A}^t)^\perp \subset \ker(A)$ und damit insgesamt Behauptung (a).

Indem man in (a) zu den orthogonalen Komplementen übergeht, die Matrix A durch $\bar{A}^t$ ersetzt und $\overline{\bar{A}^t}^t = A$ beachtet, erhält man

$$\ker(\bar{A}^t)^\perp = \mathrm{im}(A)^{\perp\perp}.$$

Vorher hatten wir schon für beliebige Untervektorräume $V \subset \mathbb{C}^n$ die Gültigkeit von $V^{\perp\perp} = V$ gezeigt, und folglich ist auch Behauptung (b) bewiesen. □

Bei selbstadjungierten Abbildungen können wir mit Hilfe dieses Satzes durch Abspaltung des Kerns zu Isomorphismen übergehen.

7.2.20 Satz

Die lineare Abbildung $F : \mathbb{C}^n \to \mathbb{C}^n$ sei selbstadjungiert, es bezeichne $V = \ker(F)^\perp$. Dann ist die Einschränkung von F auf V

$$F|V : V \to V$$

ein Isomorphismus.

Beweis

Gemäß dem vorhergehenden Satz ist $\mathrm{im}(F|V) \subset \mathrm{im}(F) = \ker(F)^\perp = V$. Die Abbildung $F|V$ ist offensichtlich injektiv und damit auf Grund der Dimensionsformel auch surjektiv. □

Normale Abbildungen

Unitäre und selbstadjungierte Abbildungen scheinen zunächst wenig miteinander zu tun zu haben: Unitäre Abbildungen sind längentreu, mithin stets Isomorphismen. Selbstadjungierte Abbildungen können durchaus einen nichttrivialen Kern haben und die eingesetzten Vektoren strecken oder stauchen. Dennoch teilen diese Abbildungen eine Eigenschaft, die sich in Abschnitt 7.4 als entscheidend für die Diagonalisierbarkeit erweisen wird.

7.2.21 Definition

Wir nennen eine $n \times n$-Matrix A und die zugehörige lineare Abbildung $F : \mathbb{C}^n \to \mathbb{C}^n$ **normal**, *wenn die Abbildung und ihre adjungierte kommutieren, falls also die Bedingung*

$$\bar{A}^t \circ A = A \circ \bar{A}^t$$

erfüllt ist.

7.2.22 Satz

Eine $n \times n$-Matrix A ist normal genau dann, wenn für alle $\vec{z}, \vec{w} \in \mathbb{C}^n$ gilt:

$$(\bar{A}^t \circ \vec{z}, \bar{A}^t \circ \vec{w}) = (A \circ \vec{z}, A \circ \vec{w}). \qquad (*)$$

Beweis

Ist A normal, so folgt für alle $\vec{z}, \vec{w} \in \mathbb{C}^n$

$$\begin{aligned}(A \circ \vec{z}, A \circ \vec{w}) &= (\vec{z}, \bar{A}^t \circ A \circ \vec{w}) = (\vec{z}, A \circ \bar{A}^t \circ \vec{w}) \\ &= (\bar{A}^t \circ \vec{z}, \bar{A}^t \circ \vec{w}).\end{aligned}$$

Gilt nun umgekehrt $(*)$ für alle $\vec{z}, \vec{w} \in \mathbb{C}^n$, so folgt

$$\begin{aligned}\left(\vec{z}, \left(\bar{A}^t \circ A\right) \circ \vec{w}\right) &= (A \circ \vec{z}, A \circ \vec{w}) = \left(\bar{A}^t \circ \vec{z}, \bar{A}^t \circ \vec{w}\right) \\ &= \left(\vec{z}, \left(A \circ \bar{A}^t\right) \circ \vec{w}\right), \\ 0 &= \left(\vec{z}, \left(\bar{A}^t \circ A - A \circ \bar{A}^t\right) \circ \vec{w}\right).\end{aligned}$$

Indem man bei beliebigem $\vec{w}$ für $\vec{z} = \left(\bar{A}^t \circ A - A \circ \bar{A}^t\right) \circ \vec{w}$ einsetzt, erhält man $\left(\bar{A}^t \circ A - A \circ \bar{A}^t\right) \circ \vec{w} = \vec{0}$ und damit

$$\bar{A}^t \circ A - A \circ \bar{A}^t = \mathcal{O};$$

die Matrix A ist normal. □

7.3 Eigenwerte

Wir haben stets die Theorie linearer Abbildungen und den Matrizenkalkül simultan und gleichberechtigt entwickelt und haben häufig mit Gewinn zwischen diesen beiden Standpunkten gewechselt. In diesem Abschnitt stellen wir den „Abbildungsstandpunkt“ in den Vordergrund.

Es bezeichne V stets einen n-dimensionalen komplexen Vektorraum und $F : V \to V$ eine $\mathbb{C}$-lineare Abbildung.

Eigenräume

7.3.1 Definition

Ein Vektor $z \in V \setminus \{0\}$ heißt **Eigenvektor** *von F zum* **Eigenwert** $\lambda \in \mathbb{C}$, *wenn die Eigenwertgleichung*

$$F(z) = \lambda \cdot z$$

erfüllt ist.

Da gleichzeitig Eigenwert λ und Eigenvektor z gesucht werden und die Abbildung $\mathbb{C} \times V \to V$, $(\lambda, z) \mapsto \lambda \cdot z$ nicht linear ist, handelt es sich bei der Eigenwertbestimmung um ein nichtlineares Problem.

Diese Begriffe stehen ebenso wie die im Folgenden noch zu definierenden auch für beliebige $n \times n$-Matrizen zur Verfügung, indem man die entsprechende $\mathbb{C}$-lineare Abbildung betrachtet. So ist $\vec{z} \in \mathbb{C}^n \setminus \{\vec{0}\}$ Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{C}$, falls $A \circ \vec{z} = \lambda \cdot \vec{z}$ gilt. Das Bestehen dieser Eigenwertgleichung ist äquivalent zu $(A - \lambda E) \circ \vec{z} = \vec{0}$, bzw. auf Abbildungsebene zu $(F - \lambda \cdot id)(z) = 0$. Also gilt

$$\begin{aligned} &\{z \in V : \ z \text{ ist Eigenvektor von } F : V \to V \text{ zu } \ \lambda \in \mathbb{C}\} \cup \{0\} \\ &= \{z : \ (F - \lambda \cdot id)(z) = 0\} = \ker(F - \lambda \cdot id). \end{aligned}$$

7.3.2 Definition

Für $\lambda \in \mathbb{C}$ heißt die Menge aller Eigenvektoren zu λ, ergänzt um den Nullvektor,

$$V_\lambda := V_\lambda(F) := \ker(F - \lambda \cdot id)$$

Eigenraum *von F zum Eigenwert $\lambda \in \mathbb{C}$.*

7.3.3 Bemerkung

Als Kern einer linearen Abbildung ist V_λ ein $\mathbb{C}$-Untervektorraum von V. Genau dann ist λ Eigenwert von F, wenn $\dim V_\lambda \neq 0$ gilt. Ist z Eigenvektor zu λ von F, so auch jedes komplexe Vielfache $c \cdot z$, $c \in \mathbb{C}$.

7.3.4 Satz

Es seien $z_1, \ldots, z_k$ Eigenvektoren von F zu den paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$:

$$F(z_\nu) = \lambda_\nu z_\nu, \quad z_\nu \neq 0, \quad \lambda_\mu \neq \lambda_\nu \text{ für } \mu \neq \nu.$$

Dann sind $z_1, \ldots, z_k$ linear unabhängig.

Insbesondere hat F höchstens n paarweise verschiedene Eigenwerte.

Beweis

Wir führen den Beweis durch Induktion nach der Anzahl k der Eigenvektoren. Der Fall $k = 1$ ist offensichtlich, denn Eigenvektoren sind definitionsgemäß vom Nullvektor verschieden. Sei nun die Behauptung für $k - 1$ gezeigt, und es seien $z_1, \ldots, z_k$ Eigenvektoren zu den paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$. Wir betrachten Linearkombinationen des Nullvektors aus $z_1, \ldots, z_k$:

$$\sum_{\nu=1}^{k} a_\nu z_\nu = 0, \qquad a_\nu \in \mathbb{C}.$$

Man erhält durch Multiplikation mit λ_k:

$$\sum_{\nu=1}^{k} a_\nu \lambda_k z_\nu = 0$$

und durch die Anwendung von F:

$$0 = F(0) = F\left(\sum_{\nu=1}^{k} a_\nu z_\nu\right) = \sum_{\nu=1}^{k} a_\nu F(z_\nu) = \sum_{\nu=1}^{k} a_\nu \lambda_\nu z_\nu .$$

Durch Subtraktion der beiden Gleichungen folgt:

$$\sum_{\nu=1}^{k} a_\nu(\lambda_k - \lambda_\nu) z_\nu = \sum_{\nu=1}^{k-1} a_\nu(\lambda_k - \lambda_\nu) z_\nu = 0.$$

Nach Induktionsvoraussetzung sind $z_1, \ldots, z_{k-1}$ linear unabhängig und somit $a_\nu(\lambda_k - \lambda_\nu) = 0$. Da die Eigenwerte als paarweise verschieden vorausgesetzt sind, erhalten wir $a_1 = \ldots = a_{k-1} = 0$. Aus der ursprünglichen Linearkombination des Nullvektors wird $0 = a_k z_k$, hieraus folgt schließlich auch $a_k = 0$. $\square$

Das charakteristische Polynom

Die Bestimmung von Eigenwerten soll auf die Bestimmung der Nullstellen eines komplexen Polynoms reduziert werden. Wir fassen hier komplexe Polynome im Sinne der analytischen Definition als Funktionen $\mathbb{C} \to \mathbb{C}$ auf.

Sei $G : \mathbb{C}^n \to V$ eine Karte, d. h. ein $\mathbb{C}$-Vektorraumisomorphismus. Die Abbildung F in den Koordinaten, $G^{-1} \circ F \circ G$, wird durch eine $n \times n$-Matrix A gegeben. Wegen der Linearität des Koordinatensystems G^{-1} lautet die Abbildung $F - \lambda \cdot id$ in den Koordinaten: $G^{-1} \circ (F - \lambda \cdot id) \circ G = G^{-1} \circ (F \circ G - \lambda \cdot G) = G^{-1} \circ F \circ G - \lambda \cdot G^{-1} \circ G = G^{-1} \circ F \circ G - \lambda \cdot id_{\mathbb{C}^n}$, sie wird also durch die Matrix $A - \lambda \cdot E$ beschrieben. Da die komplexe Zahl λ Eigenwert von A bzw. F genau dann ist, wenn $A - \lambda E$ singulär (nicht invertierbar) ist, liegt es nahe, die entsprechende Determinante zu betrachten:

$$\begin{aligned}
\det(A - \lambda E) &= \sum_{\nu_1, \ldots, \nu_n = 1}^{n} \delta(\nu_1, \ldots, \nu_n)(a_{\nu_1 1} - \lambda \delta_{\nu_1 1}) \cdot \ldots \cdot (a_{\nu_n n} - \lambda \delta_{\nu_n n}) \\
&= (-1)^n \cdot (\lambda^n + Q(\lambda)),
\end{aligned}$$

dabei ist $Q(\lambda)$ ein Polynom in λ vom Grade $\leq n - 1$. Es handelt sich bei $\det(A - \lambda E)$ also um ein Polynom n-ten Grades in λ mit höchtem Koeffizienten $(-1)^n$. Auf Grund der Überlegungen aus Kapitel 5.6 ist $\det(A - \lambda E) = \det(F - \lambda \cdot id)$ unabhängig von der Wahl der Karte G.

7.3.5 Definition

Wir nennen $P(\lambda) := P_F(\lambda) := (-1)^n \cdot \det(F - \lambda \cdot id)$ *das* **charakteristische Polynom** *der linearen Abbildung* F.

Das charakteristische Polynom ist normiert, besitzt also höchsten Koeffizienten 1.

7.3.6 Satz

Sei $P(\lambda)$ *das charakteristische Polynom einer* $\mathbb{C}$*-linearen Abbildung* $F : V \to V$. *Dann ist* $\lambda \in \mathbb{C}$ *Eigenwert von* F *genau dann, wenn* $P(\lambda) = 0$ *gilt.*

Beweis

Die komplexe Zahl $\lambda \in \mathbb{C}$ ist Eigenwert von F genau dann, wenn die Abbildung $F - \lambda \cdot id$ nicht injektiv und damit nicht invertierbar ist. Letzteres ist jedoch äquivalent zum Verschwinden der Determinante: $P(\lambda) = \det(F - \lambda\, id) = 0$. $\square$

Das charakteristische Polynom ist invariant gegenüber $\mathbb{C}$-Isomorphismen:

7.3.7 Satz

Sei $P(\lambda)$ *das charakteristische Polynom einer* $\mathbb{C}$*-linearen Abbildung* $F : V \to V$. *Sei* W *ein weiterer* n*-dimensionaler* $\mathbb{C}$*-Vektorraum und* $H : W \xrightarrow{\sim} V$ *ein* $\mathbb{C}$*-Vektorraumisomorphismus. Für das charakteristische Polynom* $\hat{P}(\lambda)$ *der Abbildung* $\hat{F} := H^{-1} \circ F \circ H$ *gilt* $\hat{P}(\lambda) = P(\lambda)$.

Beweis

Sei $G : \mathbb{C}^n \to V$ eine Karte. Dann ist $H^{-1} \circ G$ eine Karte von W und $\hat{F}$ lautet in diesen Koordinaten

$$(H^{-1} \circ G)^{-1} \circ \hat{F} \circ (H^{-1} \circ G) = G^{-1} \circ H \circ H^{-1} \circ F \circ H \circ H^{-1} \circ G = G^{-1} \circ F \circ G,$$

das ist auch die Abbildung F in den Koordinaten G. Das charakteristische Polynom ist aber von der Wahl der Karte unabhängig. $\square$

Als Folgerung erhalten wir die Existenz einer weiteren Matrixinvariante:

7.3.8 Definition

Unter der **Spur** *einer quadratischen Matrix* $A = (a_{\mu\nu})_{\mu,\nu=1,\dots,n}$ *verstehen wir die Summe der Diagonalelemente:*

$$\mathrm{Sp}(A) := \sum_{\nu=1}^{n} a_{\nu\nu}.$$

7.3.9 Satz

Sei $A \in \mathbb{C}^{n\times n}$ *beliebig,* $B \in \mathbb{C}^{n\times n}$ *invertierbar. Dann haben* A *und* $B^{-1} \circ A \circ B$ *dieselbe Spur.*

Beweis

Eine genaue Inspektion der Definition des charakteristischen Polynoms ergibt

$$\begin{aligned} P_A(\lambda) &= (-1)^n \det(A - \lambda E) \\ &= (-1)^n \sum_{\nu_1,\ldots,\nu_n=1}^{n} \delta(\nu_1,\ldots,\nu_n)(a_{\nu_1 1} - \lambda\delta_{\nu_1 1}) \cdot \ldots \cdot (a_{\nu_n n} - \lambda\delta_{\nu_n n}) \\ &= \lambda^n - \mathrm{Sp}(A)\lambda^{n-1} + \tilde{Q}(\lambda), \end{aligned}$$

das Polynom $\tilde{Q}(\lambda)$ ist vom Grade $\leq n-2$. Die Koeffizienten dieses Polynoms ändern sich unter dem Koordinatenwechsel B nicht. $\square$

7.3.10 Bemerkung (Existenz reeller Eigenvektoren)

Sei $\lambda \in \mathbb{R}$ ein Eigenwert der reellen $n \times n$-Matrix A. Dann ist $\det(A - \lambda E) = 0$, und es existiert ein $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$ mit $(A - \lambda E) \circ \vec{x} = \vec{0}$. Also besitzt A einen reellen Eigenvektor.

Der folgende, für unsere Zwecke zentrale Satz verliert in den reellen Zahlen seine Gültigkeit und ist der Grund dafür, daß die Eigenwerttheorie in *komplexen* Vektorräumen durchgeführt wird.

7.3.11 Satz (Fundamentalsatz der Algebra)

Jedes komplexe Polynom zerfällt über $\mathbb{C}$ in Linearfaktoren. Genauer gilt: Sei $P(\lambda)$ ein Polynom vom Grade n über $\mathbb{C}$. Dann gibt es Zahlen $s \in \mathbb{N}$, $s \leq n$, $r_1,\ldots,r_s \in \mathbb{N}$ mit $r_1 + \cdots + r_s = n$, $a \in \mathbb{C}\setminus\{0\}$, $c_1,\ldots,c_s \in \mathbb{C}$ mit $c_\mu \neq c_\nu$ für $\mu \neq \nu$, so daß gilt:

$$P(\lambda) = a \cdot (\lambda - c_1)^{r_1} \cdot \ldots \cdot (\lambda - c_s)^{r_s}.$$

Man sagt, daß $P(\lambda)$ in c_ν eine Nullstelle r_ν*-ter Ordnung* besitzt.

Der Beweis kann mit den hier zur Verfügung stehenden Mitteln nicht geführt werden und wird im Rahmen der Funktionentheorie oder der Algebra gegeben.

Als Folgerung erhalten wir insbesondere ein Existenzresultat für *komplexe* Eigenwerte:

7.3.12 Satz

Jede $\mathbb{C}$-lineare Abbildung $F : V \to V$ besitzt mindestens einen Eigenwert.

7.3.13 Definition

Sei $P_F(\lambda) = (\lambda - \lambda_1)^{r_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{r_s}$ mit $\lambda_\nu \neq \lambda_\mu$ für $\mu \neq \nu$, $r_1,\ldots,r_s \in \mathbb{N}$ und $r_1 + \ldots + r_s = n$ das charakteristische Polynom einer $\mathbb{C}$-linearen Abbildung F. Man nennt die Nullstellenordnung r_ν die **algebraische** *und die Dimension des Eigenraumes $\dim V_{\lambda_\nu}$ die* **geometrische Vielfachheit** *des Eigenwertes λ_ν.*

7.3.14 Beispiele

(a) Sei

$$A := \begin{pmatrix} a_1 & \dots & * \\ & \ddots & \vdots \\ \mathcal{O} & & a_n \end{pmatrix}$$

eine Dreiecksmatrix. Gemäß Satz 5.4.2 gilt:

$$P(\lambda) = (-1)^n \cdot \det(A - \lambda E) = (-1)^n \cdot \prod_{\nu=1}^{n} (a_\nu - \lambda) = \prod_{\nu=1}^{n} (\lambda - a_\nu).$$

Die Diagonalelemente $a_1, \dots, a_n$ sind die Eigenwerte von A.

(b) Sei

$$A := \begin{pmatrix} a_1 & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & a_n \end{pmatrix}$$

eine Diagonalmatrix. Es gilt $A \circ \vec{e}_\nu = a_\nu \cdot \vec{e}_\nu$, d. h. $\vec{e}_\nu$ ist Eigenvektor von A zum Eigenwert a_ν. Die kanonische Basis ist eine Orthonormalbasis aus Eigenvektoren von A. Die algebraischen und geometrischen Vielfachheiten der Eigenwerte stimmen überein.

(c) Auch reelle Matrizen haben i. a. nur komplexe Eigenwerte. Dazu berechnen wir die Eigenwerte von $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Das charakteristische Polynom lautet $P(\lambda) = \det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1$, die Eigenwerte sind i und $-i$.

(d) Nicht für jede Matrix existiert eine Basis aus Eigenvektoren. Man betrachte dazu $A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$ mit $a \in \mathbb{C} \setminus \{0\}$. Es ist $P(\lambda) = (-1)^2(1-\lambda)^2 = (1-\lambda)^2$, 1 ist also einziger Eigenwert von A. Zur Bestimmung der Eigenvektoren ist folgendes Gleichungssystem zu lösen:

$$\vec{0} = (A - E) \circ \vec{z} = \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} a \cdot z_2 \\ 0 \end{pmatrix}.$$

Der Eigenraum zum einzigen Eigenwert 1 ist also

$$V_1 = \{(z_1, z_2) \in \mathbb{C}^2 : \ z_2 = 0\} = \{(z_1, 0) : z_1 \in \mathbb{C}\}.$$

Insbesondere existieren nicht zwei linear unabhängige Eigenvektoren von A. Die algebraische Vielfachheit des Eigenwertes 1 ist 2, die geometrische 1.

7.3.15 Folgerung

Die Determinante ist das Produkt der Eigenwerte (entsprechend ihren algebraischen Vielfachheiten). Genauer gilt: Sei $P_F(\lambda) = (\lambda - \lambda_1)^{r_1} \cdot \ldots \cdot (\lambda - \lambda_s)^{r_s}$ mit $\lambda_\nu \neq \lambda_\mu$ für $\mu \neq \nu$, $r_1, \ldots, r_s \in \mathbb{N}$ und $r_1 + \ldots + r_s = n$ das charakteristische Polynom der $\mathbb{C}$-linearen Abbildung F. Dann gilt: $\det(F) = \prod_{\nu=1}^{s} \lambda_\nu^{r_\nu}$.

Beweis

Wir haben

$$P_F(\lambda) = (-1)^n \det(F - \lambda\, id) = \prod_{\nu=1}^{s} (\lambda - \lambda_\nu)^{r_\nu}\,.$$

Durch Einsetzen von $\lambda = 0$ folgt

$$\det(F) = \prod_{\nu=1}^{s} \lambda_\nu^{r_\nu}\,. \qquad \square$$

Eigenvektoren normaler Matrizen

Am Ende von Abschnitt 7.2 haben wir die normalen Matrizen auf zunächst wenig anschauliche Weise als Oberbegriff für die Hermiteschen und unitären Matrizen eingeführt. Die nächsten Sätze verdeutlichen, wie wichtig dieser neue Begriff tatsächlich ist.

7.3.16 Satz

Die $n \times n$-Matrix A sei normal. Dann ist λ Eigenwert von A genau dann, wenn $\bar{\lambda}$ Eigenwert der adjungierten Abbildung $\bar{A}^t$ ist; die Eigenräume stimmen überein:

$$V_\lambda(A) = V_{\bar{\lambda}}(\bar{A}^t).$$

Beweis

Zunächst sei bemerkt, daß mit A auch $A - \lambda\, E$ normal ist.

Sei $\vec{z} \in V_\lambda(A)$, also $A \circ \vec{z} - \lambda\, \vec{z} = \vec{0}$. Mit Hilfe von Satz 7.2.22 folgt

$$\begin{aligned} 0 &= \|\, A \circ \vec{z} - \lambda \vec{z}\, \|^2 = ((A - \lambda E) \circ \vec{z}, (A - \lambda E) \circ \vec{z}\,) \\ &= \left(\overline{(A - \lambda E)}^t \circ \vec{z}, \overline{(A - \lambda E)}^t \circ \vec{z}\right) \\ &= \left((\bar{A}^t - \bar{\lambda} \cdot E) \circ \vec{z}, (\bar{A}^t - \bar{\lambda} \cdot E) \circ \vec{z}\right) \\ &= \|\, \bar{A}^t \circ \vec{z} - \bar{\lambda} \cdot \vec{z}\, \|^2; \\ \bar{A}^t \circ \vec{z} &= \bar{\lambda} \cdot \vec{z}. \end{aligned}$$

Somit ist $V_\lambda(A) \subset V_{\bar{\lambda}}(\bar{A}^t)$. Derselbe Schluß, nun angewandt auf $\bar{\lambda}$ und $\bar{A}^t$, zeigt $V_{\bar{\lambda}}(\bar{A}^t) \subset V_{\bar{\bar{\lambda}}}\left(\overline{\bar{A}^t}^t\right) = V_\lambda(A)$ und damit auch $V_\lambda(A) = V_{\bar{\lambda}}(\bar{A}^t)$. $\square$

7.3.17 Satz

Je zwei Eigenvektoren $\vec{z}_1, \vec{z}_2$ einer normalen Matrix A zu verschiedenen Eigenwerten $\lambda_1 \neq \lambda_2$ sind orthogonal.

Beweis

Gemäß Voraussetzung ist $A \circ \vec{z}_1 = \lambda_1 \vec{z}_1$, $A \circ \vec{z}_2 = \lambda_2 \vec{z}_2$; der vorhergehende Satz impliziert $\bar{A}^t \circ \vec{z}_1 = \bar{\lambda}_1 \vec{z}_1$, $\bar{A}^t \circ \vec{z}_2 = \bar{\lambda}_2 \vec{z}_2$. Damit folgt:

$$\begin{aligned} \lambda_2(\vec{z}_1, \vec{z}_2) &= (\vec{z}_1, \lambda_2 \vec{z}_2) = (\vec{z}_1, A \circ \vec{z}_2) \\ &= (\bar{A}^t \circ \vec{z}_1, \vec{z}_2) = (\bar{\lambda}_1 \circ \vec{z}_1, \vec{z}_2) = \lambda_1(\vec{z}_1, \vec{z}_2), \\ 0 &= (\lambda_1 - \lambda_2)(\vec{z}_1, \vec{z}_2). \end{aligned}$$

Aus der Verschiedenheit der Eigenwerte erhalten wir die Orthogonalität der Eigenvektoren: $(\vec{z}_1, \vec{z}_2) = 0$. □

Eigenwerte Hermitescher Matrizen

7.3.18 Satz

Alle Eigenwerte Hermitescher Matrizen sind reell.

Beweis

Sei $A = A^{n,n}$ Hermitesch mit Eigenvektor $\vec{z} \in \mathbb{C}^n \setminus \{\vec{0}\}$ zum Eigenwert $\lambda \in \mathbb{C}$. Satz 7.2.18 zeigt:

$$\lambda \| \vec{z} \|^2 = \lambda(\vec{z}, \vec{z}) = (\vec{z}, \lambda \cdot \vec{z}) = (\vec{z}, A \circ \vec{z}) \in \mathbb{R}.$$

Wegen $\| \vec{z} \|^2 \in \mathbb{R} \setminus \{0\}$ folgt, daß λ reell ist. □

7.3.19 Beispiele

(a) Für die Hermitesche 2×2-Matrix

$$A = \begin{pmatrix} 1 & -3i \\ 3i & 1 \end{pmatrix}$$

lautet das charakteristische Polynom

$$\begin{aligned} P(\lambda) = \det(A - \lambda\, E) = \det \begin{pmatrix} 1-\lambda & -3i \\ 3i & 1-\lambda \end{pmatrix} = (1-\lambda)^2 + 9i^2 \\ = (\lambda - 1)^2 - 9 = (\lambda - 4)(\lambda + 2); \end{aligned}$$

die Eigenwerte sind $\lambda_1 = 4$ und $\lambda_2 = -2$.

(b) Wir betrachten nun die reelle symmetrische 3×3-Matrix

$$A = \begin{pmatrix} \frac{7}{6} & -\frac{2}{3} & \frac{1}{6} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{1}{6} & -\frac{2}{3} & \frac{7}{6} \end{pmatrix}.$$

Das charakteristische Polynom lautet

$$\begin{aligned} P(\lambda) = (-1)^3 \det(A - \lambda E) &= -\frac{1}{216} \det \begin{pmatrix} (7-6\lambda) & -4 & 1 \\ -4 & (4-6\lambda) & -4 \\ 1 & -4 & (7-6\lambda) \end{pmatrix} \\ &= -\frac{1}{216} \det \begin{pmatrix} (7-6\lambda) & -4 & 1 \\ 24(1-\lambda) & -6(2+\lambda) & 0 \\ 12(-3\lambda^2+7\lambda-4) & 24(1-\lambda) & 0 \end{pmatrix} \\ &= -\frac{72}{216} \left\{ 8(1-\lambda)^2 + (2+\lambda)(-3\lambda^2+7\lambda-4) \right\} \\ &= -\frac{1}{3} \left\{ -3\lambda^3 + 9\lambda^2 - 6\lambda \right\} = \lambda(\lambda-1)(\lambda-2), \end{aligned}$$

wir haben also die drei Eigenwerte $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 2$.

Eigenwerte unitärer Matrizen

7.3.20 Satz

Für jeden Eigenwert λ einer unitären Matrix A gilt $|\lambda| = 1$.

Beweis

Sei λ Eigenwert von $A \in U(n)$ zum Eigenvektor $\vec{z} \in \mathbb{C}^n \setminus \{\vec{0}\}$. Die Längentreue der unitären Matrix impliziert

$$\| \vec{z} \|^2 = \| A \circ \vec{z} \|^2 = (A \circ \vec{z}, A \circ \vec{z}) = (\lambda \cdot \vec{z}, \lambda \cdot \vec{z}) = \lambda \bar{\lambda} \| \vec{z} \|^2.$$

Da $\| \vec{z} \|^2 \neq 0$ ist, folgt $|\lambda| = 1$. □

7.3.21 Satz

Sei $A \in O(n, \mathbb{R})$ eine reelle orthogonale Matrix. Die Raumdimension n sei ungerade. Dann ist $+1$ oder -1 Eigenwert von A. Ist A sogar eine Drehung, d. h. $A \in SO(n, \mathbb{R})$, dann ist $+1$ Eigenwert von A.

Wir werden auf S. 257 mit Hilfe der Diagonalisierungsresultate erläutern, daß in $\mathbb{R}^3$ der zu $+1$ gehörige Eigenvektor als Drehachse interpretiert werden kann.

Beweis

Das charakteristische Polynom $P_A(\lambda)$ von A hat reelle Koeffizienten und wird hier als Funktion $\mathbb{R} \to \mathbb{R}$ aufgefaßt. Der Term höchster Ordnung ist λ^n mit der ungeraden Potenz n. Aus der Analysis weiß man $\lim_{\lambda\to\infty} P_A(\lambda) = +\infty$ und $\lim_{\lambda\to-\infty} P_A(\lambda) = -\infty$. Der Zwischenwertsatz ergibt die Existenz wenigstens einer reellen Nullstelle λ_0 mit $|\lambda_0| = 1$, denn A ist orthogonal und insbesondere unitär.

Die nichtreellen Eigenwerte von A treten immer in zueinander konjugierten Paaren auf: Sei λ_1 Eigenwert von A, d. h. Nullstelle des charakteristischen Polynoms

$$0 = P_A(\lambda_1) = \lambda_1^n + \sum_{\nu=0}^{n-1} a_\nu \lambda_1^\nu, \qquad a_\nu \in \mathbb{R}.$$

Durch Konjugation erhält man

$$0 = \overline{P_A(\lambda_1)} = \overline{\lambda_1^n} + \overline{\sum_{\nu=0}^{n-1} a_\nu \lambda_1^\nu} = \bar\lambda_1^n + \sum_{\nu=0}^{n-1} a_\nu \bar\lambda_1^\nu,$$

$\bar\lambda_1$ ist also ebenfalls Eigenwert. Indem man den quadratischen Faktor $(\lambda - \lambda_1)(\lambda - \bar\lambda_1)$ abspaltet, der nur reelle Koeffizienten besitzt, und das Verfahren iteriert, erhält man: Die Matrix A hat $0 \le k \le (n-1)/2$ (nicht notwendigerweise verschiedene) Paare zueinander konjugiert komplexer Eigenwerte $\lambda_1, \bar\lambda_1, \ldots, \lambda_k, \bar\lambda_k$ und eine ungerade Anzahl reeller Eigenwerte $\lambda_{2k+1}, \ldots, \lambda_n \in \{+1, -1\}$; die Eigenwerte werden jeweils entsprechend ihrer algebraischen Vielfachheit aufgeführt. Die Determinante ist das Produkt der Eigenwerte:

$$\det(A) = |\lambda_1|^2 \cdot \ldots \cdot |\lambda_k|^2 \cdot \lambda_{2k+1} \cdot \ldots \cdot \lambda_n = \lambda_{2k+1} \cdot \ldots \cdot \lambda_n.$$

Ist nun $\det(A) = +1$, so können nicht alle reellen Eigenwerte $= -1$ sein, denn n wurde als ungerade vorausgesetzt. In diesem Fall ist $+1$ (zumindest einfacher) Eigenwert von A. $\square$

7.3.22 Beispiel

Durch direktes Nachrechnen sieht man, daß die reelle 3×3-Matrix

$$A = \begin{pmatrix} \frac12 + \frac{\sqrt2}{4} & -\frac12 + \frac{\sqrt2}{4} & -\frac12 \\ -\frac12 + \frac{\sqrt2}{4} & \frac12 + \frac{\sqrt2}{4} & -\frac12 \\ \frac12 & \frac12 & \frac{\sqrt2}{2} \end{pmatrix}$$

orthogonal ist, daß also $A^t \circ A = E$ gilt. Wir berechnen das charakteristische Polynom:

$$P_A(\lambda) = -\det(A - \lambda E) = -\det \begin{pmatrix} \frac12 + \frac{\sqrt2}{4} - \lambda & -\frac12 + \frac{\sqrt2}{4} & -\frac12 \\ -\frac12 + \frac{\sqrt2}{4} & \frac12 + \frac{\sqrt2}{4} - \lambda & -\frac12 \\ \frac12 & \frac12 & \frac{\sqrt2}{2} - \lambda \end{pmatrix}$$

$$
\begin{aligned}
&= -\det\begin{pmatrix} \frac12+\frac{\sqrt2}{4}-\lambda & -1+\lambda & -\frac12 \\ -\frac12+\frac{\sqrt2}{4} & 1-\lambda & -\frac12 \\ \frac12 & 0 & \frac{\sqrt2}{2}-\lambda \end{pmatrix} \\
&= -\det\begin{pmatrix} \frac12+\frac{\sqrt2}{4}-\lambda & -1+\lambda & -\frac12 \\ \frac{\sqrt2}{2}-\lambda & 0 & -1 \\ \frac12 & 0 & \frac{\sqrt2}{2}-\lambda \end{pmatrix} \\
&= (\lambda-1)\cdot\det\begin{pmatrix} \frac{\sqrt2}{2}-\lambda & -1 \\ \frac12 & \frac{\sqrt2}{2}-\lambda \end{pmatrix} = (\lambda-1)\cdot\left(\left(\frac{\sqrt2}{2}-\lambda\right)^2+\frac12\right) \\
&= (\lambda-1)\cdot\left(\lambda-\frac{\sqrt2}{2}-i\frac{\sqrt2}{2}\right)\cdot\left(\lambda-\frac{\sqrt2}{2}+i\frac{\sqrt2}{2}\right),
\end{aligned}
$$

die Eigenwerte von A lauten also $\lambda_1 = \frac{\sqrt2}{2}(1+i)$, $\lambda_2 = \frac{\sqrt2}{2}(1-i)$ und $\lambda_3 = 1$. Daraus folgt auch $\det(A) = \frac{\sqrt2}{2}(1+i)\cdot\frac{\sqrt2}{2}(1-i)\cdot 1 = 1$, also ist A eine Drehung.

Wir bestimmen die Drehachse (sie sei vorbehaltlich der nachzutragenden Erläuterungen von S. 257 bereits so bezeichnet), d. h. den Eigenvektor zum Eigenwert 1, als nichttriviale Lösung des linearen Gleichungssystems

$$
\vec0 = (A-E)\circ\vec x = \begin{pmatrix} -\frac12+\frac{\sqrt2}{4} & -\frac12+\frac{\sqrt2}{4} & -\frac12 \\ -\frac12+\frac{\sqrt2}{4} & -\frac12+\frac{\sqrt2}{4} & -\frac12 \\ \frac12 & \frac12 & \frac{\sqrt2}{2}-1 \end{pmatrix}\circ\vec x.
$$

Etwa mittels des Gaußschen Algorithmus' findet man, daß dieses Gleichungssystem eine linear unabhängige Lösung hat. Der Lösungsraum (die Drehachse) wird von dem Einheitsvektor $\vec x = \left(\frac{\sqrt2}{2}, -\frac{\sqrt2}{2}, 0\right)^t$ aufgespannt.

7.4 Die Hauptachsentransformation

In diesem Abschnitt betrachten wir nur $n\times n$-Matrizen $A\in\mathbb{C}^{n\times n}$ bzw. $A\in\mathbb{R}^{n\times n}$ und die entsprechenden linearen Abbildungen $F:\mathbb{C}^n\to\mathbb{C}^n$ bzw. $F:\mathbb{R}^n\to\mathbb{R}^n$. Lineare Abbildungen allgemeiner endlichdimensionaler Vektorräume in sich könnten, wie schon erwähnt, durch die Einführung *unitärer* Koordinaten auf diesen Fall zurückgeführt werden, die relevanten Eigenschaften wie „normal“, „selbstadjungiert“ oder „unitär“ blieben dabei erhalten.

Sei also $F:\mathbb{C}^n\to\mathbb{C}^n$ eine lineare Abbildung mit zugehöriger $n\times n$-Matrix A. Wir können uns nun der zu Beginn dieses Kapitels wieder aufgeworfenen Frage zuwenden: Unter welchen Bedingungen an die Matrix A läßt sich ein orthonormiertes Koordinatensystem $(\vec b_1,\ldots,\vec b_n)$ in $\mathbb{C}^n$ so geschickt „einzeichnen“, daß die Abbildung

F in den durch die unitäre Matrix $B = (\vec{b}_1, \ldots, \vec{b}_n)$ bzw. die entsprechende unitäre Karte $G : \mathbb{C}^n \to \mathbb{C}^n$ gegebenen Koordinaten

$$\hat{F} = G^{-1} \circ F \circ G \quad \text{mit Matrix} \quad \hat{A} = \bar{B}^t \circ A \circ B$$

Diagonalgestalt hat?

Wir werden zeigen, daß ein solches Diagonalisierungsresultat genau in der Klasse der normalen Matrizen Gültigkeit besitzt. Insbesondere folgt damit die Diagonalisierbarkeit unitärer und Hermitescher Matrizen. Da die Eigenwerte reeller symmetrischer Matrizen sämtlich reell sind, erhalten wir daraus unmittelbar ein Diagonalisierungsresultat auch im Reellen, während der Übergang vom unitären zum reell orthogonalen Fall größere Modifikationen erfordert.

Transformation auf Dreiecksgestalt, Trigonalisierung

Wir werden zunächst beweisen, daß jede komplexe $n \times n$-Matrix durch eine unitäre Transformationsmatrix auf Dreiecksgestalt gebracht werden kann, und daraus später unsere Diagonalisierungsresultate herleiten.

Zur Vorbereitung dienen die folgenden beiden Hilfssätze:

7.4.1 Satz

Sei $m < n$ und $B'' \in \mathbb{C}^{m \times m}$ eine unitäre $m \times m$-Matrix. Dann ist

$$B := \begin{pmatrix} E^{n-m,n-m} & \mathcal{O}^{n-m,m} \\ \mathcal{O}^{m,n-m} & B'' \end{pmatrix} \in \mathbb{C}^{n \times n}$$

ebenfalls unitär.

Beweis

Sei $\vec{z} \in \mathbb{C}^n$ beliebig. Wir bezeichnen $\vec{z}\,' := (z_1, \ldots, z_{n-m})^t \in \mathbb{C}^{n-m}$ und $\vec{z}\,'' := (z_{n-m+1}, \ldots, z_n)^t \in \mathbb{C}^m$. Auf Grund der Nullblöcke in B ist $B \circ \vec{z} = (\vec{z}\,', B'' \circ \vec{z}\,'')^t \in \mathbb{C}^{n-m} \oplus \mathbb{C}^m$. Wir greifen noch auf die Definition der euklidischen Norm in $\mathbb{C}^n$ zurück und erhalten

$$\| B \circ \vec{z} \|^2 = \| \vec{z}\,' \|^2 + \| B'' \circ \vec{z}\,'' \|^2 = \| \vec{z}\,' \|^2 + \| \vec{z}\,'' \|^2 = \| \vec{z} \|^2. \qquad \square$$

7.4.2 Satz

Zu je zwei Einheitsvektoren $\vec{z}, \vec{w} \in \mathbb{C}^n$ (d. h. $\| \vec{z} \| = \| \vec{w} \| = 1$) existiert eine unitäre $n \times n$-Matrix $B \in U(n)$ mit $\vec{w} = B \circ \vec{z}$.

Beweis

Genauso wie in Satz 6.3.5 beweist man: es gibt Vektoren $\vec{z}_2, \ldots, \vec{z}_n$ derart, daß $(\vec{z}, \vec{z}_2, \ldots, \vec{z}_n)$ ein orthogonales n-Bein ist, d. h. nach Satz 7.2.14 ist $B_1 :=$

$(\vec{z}, \vec{z}_2, \ldots, \vec{z}_n) \in U(n)$. Es gilt

$$B_1 \circ \vec{e}_1 = \vec{z}, \text{ also } \vec{e}_1 = B_1^{-1} \circ \vec{z} = \bar{B}_1^t \circ \vec{z}.$$

Entsprechend konstruiert man ein $B_2 \in U(n)$ mit $B_2 \circ \vec{e}_1 = \vec{w}$. Für $B := B_2 \circ \bar{B}_1^t \in U(n)$ ist dann in der Tat $B \circ \vec{z} = B_2 \circ \vec{e}_1 = \vec{w}$. □

7.4.3 Satz

Zu jeder komplexen $n \times n$-Matrix A existiert eine unitäre $n \times n$-Matrix $B \in U(n)$ derart, daß $\hat{A} := \bar{B}^t \circ A \circ B$ eine (obere) Dreiecksmatrix ist, daß also $\hat{a}_{\mu\nu} = 0$ gilt für $\mu > \nu$.

Beweis

Wir führen den Beweis durch vollständige Induktion nach n. Bei der Induktionsbasis $n = 1$ ist nichts zu zeigen, denn jede 1×1-„Matrix“ ist eine Dreiecks-„Matrix“.

Sei nun $n > 1$, und wir nehmen an, die Behauptung sei für $n - 1$ bewiesen. Das charakteristische Polynom $P(\lambda) = P_A(\lambda)$ der $n \times n$-Matrix A zerfällt über $\mathbb{C}$ in Linearfaktoren, es gibt also einen Eigenwert $\lambda_1 \in \mathbb{C}$ mit zugehörigem Eigenvektor $\vec{z}_1 \neq \vec{0}$. Wir können den Eigenvektor normieren und ohne Einschränkung $\| \vec{z}_1 \| = 1$ annehmen. Gemäß Satz 7.4.2 existiert eine unitäre Abbildung $B_1 \in U(n)$, die den ersten Einheitsvektor auf diesen Eigenvektor abbildet: $B_1 \circ \vec{e}_1 = \vec{z}_1$.

Für

$$A_1 := \bar{B}_1^t \circ A \circ B_1$$

gilt

$$A_1 \circ \vec{e}_1 = \bar{B}_1^t \circ A \circ (B_1 \circ \vec{e}_1) = \bar{B}_1^t \circ (A \circ \vec{z}_1) = \bar{B}_1^t \circ (\lambda_1 \vec{z}_1) = \lambda_1 \vec{e}_1,$$

also

$$A_1 = \left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & \\ \vdots & A' \\ 0 & \end{array} \right)$$

mit einer $(n-1) \times (n-1)$-Matrix A'. Nach Induktionsvoraussetzung existiert ein $B' \in U(n-1)$, so daß in $\bar{B'}^t \circ A' \circ B'$ unterhalb der Diagonalen ausschließlich Nullen stehen. Auf die in Satz 7.4.1 beschriebene Weise erhalten wir aus B' eine n-reihige unitäre Matrix:

$$B_2 := \begin{pmatrix} 1 & \mathcal{O}^{1,n-1} \\ \mathcal{O}^{n-1,1} & B' \end{pmatrix} \in U(n).$$

Wir betrachten die Hintereinanderschaltung

$$B := B_1 \circ B_2 \in U(n)$$

der unitären Transformationen B_1 und B_2. Für

$$\hat{A} := \bar{B}_2^t \circ A_1 \circ B_2 = \bar{B}_2^t \circ \bar{B}_1^t \circ A \circ B_1 \circ B_2 = \overline{(B_1 \circ B_2)}^t \circ A \circ (B_1 \circ B_2) = \bar{B}^t \circ A \circ B$$

erhalten wir nun Dreiecksgestalt:

$$\begin{aligned}\hat{A} &= \begin{pmatrix} 1 & \mathcal{O} \\ \mathcal{O} & \bar{B}'^t \end{pmatrix} \circ \begin{pmatrix} \lambda_1 & * \\ \mathcal{O} & A' \end{pmatrix} \circ \begin{pmatrix} 1 & \mathcal{O} \\ \mathcal{O} & B' \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & * \\ \mathcal{O} & \bar{B}'^t \circ A' \end{pmatrix} \circ \begin{pmatrix} 1 & \mathcal{O} \\ \mathcal{O} & B' \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \\ \mathcal{O} & \bar{B}'^t \circ A' \circ B' \end{pmatrix}. \qquad \square\end{aligned}$$

Im Beweis haben wir wesentlichen Gebrauch vom Fundamentalsatz der Algebra gemacht. Betrachtet man eine reelle Matrix A, so hat deren charakteristisches Polynom $P_A(\lambda)$ zwar nur reelle Koeffizienten, der Fundamentalsatz der Algebra besitzt im Reellen jedoch keine Gültigkeit. Tatsächlich wird der soeben bewiesene Satz in dieser Situation im allgemeinen falsch:

7.4.4 Beispiel

Wir betrachten die reelle normale 2×2-Matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Für beliebiges $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in O(2, \mathbb{R})$ berechnet man

$$B^t \circ A \circ B = \begin{pmatrix} 0 & -ad + bc \\ -bc + ad & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\det(B) \\ \det(B) & 0 \end{pmatrix};$$

Dreiecksgestalt läßt sich nicht erreichen. Mit der komplexen unitären Matrix $C = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$ ist dagegen $\bar{C}^t \circ A \circ C = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}$.

Um das reelle Analogon von Satz 7.4.3 zu beweisen, haben wir also die Existenz von reellen Eigenwerten vorauszusetzen.

7.4.5 Satz

Die reelle $n \times n$-Matrix A habe nur reelle Eigenwerte, d. h. das charakteristische Polynom $P_A(\lambda)$ zerfalle über $\mathbb{R}$ *in Linearfaktoren. Dann existiert eine (reelle) Matrix $B \in SO(n)$ derart, daß $\hat{A} = B^{-1} \circ A \circ B = B^t \circ A \circ B$ eine Dreiecksmatrix ist.*

Beweis

Weil wir hier weitgehend parallel zum Beweis von Satz 7.4.3 vorgehen, wird nur der Induktionsschritt skizziert.

Gemäß Voraussetzung existiert ein Eigenwert $\lambda_1 \in \mathbb{R}$ von A, wegen Bemerkung 7.3.10 gibt es zu λ_1 einen reellen Eigenvektor:

$$A \circ \vec{x}_1 = \lambda_1 \cdot \vec{x}_1 \text{ mit } \vec{x}_1 \in \mathbb{R}^n, \ \| \vec{x}_1 \| = 1.$$

In Satz 6.3.5 haben wir eine Drehung $B_1 \in SO(n)$ konstruiert mit $\vec{x}_1 = B_1 \circ \vec{e}_1$. Die transformierte Matrix $A_1 := B_1^t \circ A \circ B_1$ hat dasselbe Aussehen wie oben und gemäß Satz 7.3.7 dasselbe charakteristische Polynom wie A. Ist $Q(\lambda)$ das charakteristische Polynom von A', so liefert die Entwicklung von $\det(A_1 - \lambda E)$ nach der ersten Spalte

$$P(\lambda) = (\lambda - \lambda_1) \cdot Q(\lambda).$$

Also zerfällt auch $Q(\lambda)$ über $\mathbb{R}$ in Linearfaktoren. Die Induktionsvoraussetzung liefert ein $B' \in SO(n-1)$, so daß $B'^t \circ A' \circ B'$ Dreiecksmatrix ist. Der Rest des Beweises verläuft wie oben, wobei die Konjugationsstriche entfallen und $U(n)$ durch $SO(n)$ zu ersetzen ist. □

7.4.6 Folgerung

Die geometrische Vielfachheit eines Eigenwertes kann seine algebraische Vielfachheit nicht übersteigen.

Beweis

Wir können uns auf den komplexen Fall beschränken: Die geometrische Vielfachheit im Reellen kann die im Komplexen nicht übersteigen, denn *reelle* Vektoren, also solche aus $\mathbb{R}^n$, die über $\mathbb{R}$ linear unabhängig sind, sind auch über $\mathbb{C}$ linear unabhängig. Die algebraischen Vielfachheiten stimmen im Reellen und Komplexen überein.

Sei λ_0 Eigenwert der $n \times n$-Matrix A mit geometrischer Vielfachheit $s = \dim V_{\lambda_0}$ und algebraischer Vielfachheit r. Ähnlich wie im Beweis von Satz 7.4.3 wollen wir den Eigenraum V_{λ_0} von $\mathbb{C}^n$ abspalten. Sei dazu $\vec{z}_1, \ldots, \vec{z}_s$ eine Orthonormalbasis von V_{λ_0}, diese können wir durch Verwendung des Schmidtschen Orthonormalisierungsverfahrens zu einem orthogonalen n-Bein $\vec{z}_1, \ldots, \vec{z}_n$ ergänzen. Die Matrix $B_1 := (\vec{z}_1, \ldots, \vec{z}_n)$ ist also unitär. Mit der Wahl von $A_1 := \bar{B}_1^t \circ A \circ B_1$ folgt für $\nu = 1, \ldots, s$:

$$A_1 \circ \vec{e}_\nu = \bar{B}_1^t \circ A \circ (B_1 \circ \vec{e}_\nu) = \bar{B}_1^t \circ (A \circ \vec{z}_\nu) = \lambda_0 \left(\bar{B}_1^t \circ \vec{z}_\nu\right) = \lambda_0 \vec{e}_\nu,$$

also

$$A_1 = \begin{array}{c|c} \begin{smallmatrix}\text{Zeile/}\\\text{Spalte}\end{smallmatrix} & \begin{matrix} 1 & \cdots & s & s{+}1 & \cdots & n \end{matrix} \\ \begin{matrix} 1 \\ \vdots \\ s \\ s{+}1 \\ \vdots \\ n \end{matrix} & \begin{pmatrix} \lambda_0 & & \mathcal{O} & & & \\ & \ddots & & & * & \\ \mathcal{O} & & \lambda_0 & & & \\ & & & & & \\ & \mathcal{O} & & & A' & \\ & & & & & \end{pmatrix} \end{array}.$$

Gemäß Satz 7.4.3 existiert eine unitäre Matrix $B' \in U(n-s)$, so daß $\bar{B}'^t \circ A' \circ B'$ Dreiecksgestalt hat. Mit

$$B_2 := \begin{pmatrix} E^{s,s} & \mathcal{O} \\ \mathcal{O} & B' \end{pmatrix} \in U(n)$$

erhalten wir wie in Satz 7.4.3

$$\hat{A} := \bar{B}_2^t \circ A_1 \circ B_2 = \begin{array}{cc} \begin{array}{c}\text{Zeile/}\\ \text{Spalte}\end{array} & \begin{array}{cccccc} 1 & \cdots & s & s+1 & \cdots & n \end{array} \\ \begin{array}{c} 1 \\ \vdots \\ s \\ s+1 \\ \vdots \\ n \end{array} & \begin{pmatrix} \lambda_0 & & \mathcal{O} & & & \\ & \ddots & & & * & \\ \mathcal{O} & & \lambda_0 & & & \\ & & & & & \\ & \mathcal{O} & & & \bar{B'}^t \circ A' \circ B' & \\ & & & & & \end{pmatrix} \end{array}$$

$$= \begin{array}{cc} \begin{array}{c}\text{Zeile/}\\ \text{Spalte}\end{array} & \begin{array}{cccccc} 1 & \cdots & s & s+1 & \cdots & n \end{array} \\ \begin{array}{c} 1 \\ \vdots \\ s \\ s+1 \\ \vdots \\ n \end{array} & \begin{pmatrix} \lambda_0 & & \mathcal{O} & & & \\ & \ddots & & & * & \\ \mathcal{O} & & \lambda_0 & & & \\ & & & \lambda_{s+1} & & * \\ & \mathcal{O} & & & \ddots & \\ & & & \mathcal{O} & & \lambda_n \end{pmatrix} \end{array}.$$

Die Matrizen A und $\hat{A}$ haben dasselbe charakteristische Polynom, es lautet $P(\lambda) = (\lambda - \lambda_0)^s \cdot \prod_{\nu=s+1}^{n} (\lambda - \lambda_\nu)$. Die Nullstellen λ_ν sind i. a. *nicht* paarweise und auch nicht von λ_0 verschieden, es folgt $s \leq r$. □

Diagonalisierbarkeit normaler Matrizen

Wir wollen zeigen, daß die Dreiecksmatrix aus Satz 7.4.3 sogar eine Diagonalmatrix ist, sofern eine normale Matrix vorgegeben wurde. Hier liegt die Beobachtung zugrunde, daß Normalität unter unitären Koordinatentransformationen erhalten bleibt:

7.4.7 Satz

Die Matrix $A \in \mathbb{C}^{n \times n}$ sei normal. Dann ist $\bar{B}^t \circ A \circ B$ für jede unitäre Matrix $B \in U(n)$ ebenfalls normal.

Beweis

$$\begin{aligned} &\overline{(\bar{B}^t \circ A \circ B)}^t \circ \bar{B}^t \circ A \circ B - \bar{B}^t \circ A \circ B \circ \overline{(\bar{B}^t \circ A \circ B)}^t \\ &= \bar{B}^t \circ \bar{A}^t \circ B \circ \bar{B}^t \circ A \circ B - \bar{B}^t \circ A \circ B \circ \bar{B}^t \circ \bar{A}^t \circ B \\ &= \bar{B}^t \circ \bar{A}^t \circ A \circ B - \bar{B}^t \circ A \circ \bar{A}^t \circ B \\ &= \bar{B}^t \circ \left(\bar{A}^t \circ A - A \circ \bar{A}^t\right) \circ B = \mathcal{O}. \end{aligned}$$

□

7.4.8 Satz

Zu jeder normalen Matrix $A \in \mathbb{C}^{n\times n}$ existiert eine unitäre Matrix $B \in U(n)$ derart, daß $\hat{A} := \bar{B}^t \circ A \circ B$ Diagonalgestalt hat. Die Spalten von $B = \left(\vec{b}_1, \dots, \vec{b}_n\right)$ bilden eine Orthonormalbasis von $\mathbb{C}^n$ aus Eigenvektoren von A. Die Diagonalelemente von $\hat{A}$ sind genau die Eigenwerte von A entsprechend ihren geometrischen Vielfachheiten.

Beweis

Gemäß Satz 7.4.3 finden wir eine unitäre Matrix $B \in U(n)$, so daß $\hat{A} := \bar{B}^t \circ A \circ B$ Dreiecksgestalt hat; $\hat{A}$ ist ebenfalls normal. Es bleibt zu zeigen, daß normale Dreiecksmatrizen schon Diagonalmatrizen sind.

Wir haben $\hat{a}_{\mu\nu} = 0$ für $\mu > \nu$. Die Normalität ergibt für $\mu = 1, \dots, n$:

$$\begin{aligned} 0 &= \sum_{\sigma=1}^{n} \left(\bar{\hat{a}}_{\sigma\mu}\hat{a}_{\sigma\mu} - \hat{a}_{\mu\sigma}\bar{\hat{a}}_{\mu\sigma}\right) = \sum_{\sigma=1}^{\mu} |\hat{a}_{\sigma\mu}|^2 - \sum_{\sigma=\mu}^{n} |\hat{a}_{\mu\sigma}|^2 \\ &= \sum_{\sigma=1}^{\mu-1} |\hat{a}_{\sigma\mu}|^2 - \sum_{\sigma=\mu+1}^{n} |\hat{a}_{\mu\sigma}|^2. \end{aligned}$$

Die Betrachtung von $\mu = 1$ ergibt

$$0 = \sum_{\sigma=2}^{n} |\hat{a}_{1\sigma}|^2, \qquad \hat{a}_{1\sigma} = 0 \text{ für } \sigma = 2, \dots, n.$$

Wir fahren induktiv fort und nehmen an, wir hätten im k-ten Schritt $\hat{a}_{\mu\sigma} = 0$ für $\sigma = \mu + 1, \dots, n$, $\mu = 1, \dots, k$ erreicht. Durch Betrachtung von $\mu = k + 1$ folgt:

$$0 = \sum_{\sigma=1}^{k} |\hat{a}_{\sigma,k+1}|^2 - \sum_{\sigma=k+2}^{n} |\hat{a}_{k+1,\sigma}|^2 = - \sum_{\sigma=k+2}^{n} |\hat{a}_{k+1,\sigma}|^2,$$

also ist auch $\hat{a}_{k+1,\sigma} = 0$ für $\sigma = k + 2, \dots, n$.

Nach $(n-1)$ Schritten erhalten wir, daß $\hat{a}_{\mu\nu} = 0$ auch für alle μ, ν mit $\mu < \nu$ ist. Auch oberhalb der Diagonalen enthält $\hat{A}$ nur Nullen und ist folglich eine Diagonalmatrix.

Da die Matrix B unitär ist, bilden deren Spalten ein orthogonales n-Bein, also eine Orthonormalbasis von $\mathbb{C}^n$. Wir haben zu zeigen, daß es sich dabei um Eigenvektoren von A handelt. Aus $\hat{A} = \bar{B}^t \circ A \circ B$ folgt $B \circ \hat{A} = A \circ B$. Wir betrachten jeweils die ν-te Spalte:

$$A \circ \vec{b}_\nu = B \circ \vec{\hat{a}}_\nu = B \circ (\hat{a}_{\nu\nu}\vec{e}_\nu) = \hat{a}_{\nu\nu}\,(B \circ \vec{e}_\nu) = \hat{a}_{\nu\nu} \circ \vec{b}_\nu,$$

in der Tat ist $\vec{b}_\nu$ Eigenvektor von A zum Eigenwert $\lambda_\nu := \hat{a}_{\nu\nu}$. Jeder Eigenwert tritt gemäß seiner geometrischen Vielfachheit auf, denn die Anzahl der entsprechenden Eigenvektoren in B ist gerade die Dimension des Eigenraums. Somit sind $\hat{a}_{11}, \dots, \hat{a}_{nn}$

genau die Eigenwerte von A, denn A und $\hat{A}$ haben dieselben Eigenwerte mit denselben geometrischen Vielfachheiten. □

Mit den normalen Matrizen haben wir die größtmögliche Klasse derjenigen Matrizen bestimmt, die mit Hilfe unitärer Koordinatenwechselabbildungen diagonalisiert werden können:

7.4.9 Satz

Für die $n \times n$-Matrix A existiere eine unitäre $n \times n$-Matrix $B \in U(n)$ derart, daß $\bar{B}^t \circ A \circ B$ Diagonalgestalt hat. Dann ist A normal.

Beweis

Es bezeichne

$$D = \begin{pmatrix} d_1 & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & d_n \end{pmatrix} = \bar{B}^t \circ A \circ B;$$

diese Diagonalmatrix ist normal:

$$\begin{aligned} \bar{D}^t \circ D &= \begin{pmatrix} \bar{d}_1 & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & \bar{d}_n \end{pmatrix} \circ \begin{pmatrix} d_1 & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & d_n \end{pmatrix} \\ &= \begin{pmatrix} \bar{d}_1 \cdot d_1 & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & \bar{d}_n \cdot d_n \end{pmatrix} = D \circ \bar{D}^t. \end{aligned}$$

Nun ist $A = B \circ D \circ \bar{B}^t = \overline{\bar{B}^t}^t \circ D \circ \bar{B}^t$ mit $\bar{B}^t \in U(n)$, und A ist gemäß Satz 7.4.7 normal. □

Das reelle Analogon zu Satz 7.4.8 leitet man mit wörtlich demselben Beweis aus Satz 7.4.5 her.

7.4.10 Satz

Die reelle $n \times n$-Matrix A sei normal und habe nur reelle Eigenwerte. Dann existiert eine reelle Drehung $B \in SO(n)$ derart, daß $\hat{A} := B^t \circ A \circ B$ eine reelle Diagonalmatrix ist.

Bezüglich jeder normalen Matrix können wir mit Hilfe des Diagonalisierungssatzes den $\mathbb{C}^n$ in eine orthogonale Summe über deren Eigenräume zerlegen:

7.4.11 Satz

Es seien $\lambda_1, \ldots, \lambda_s$ sämtliche paarweise verschiedenen Eigenwerte einer normalen $n \times n$-Matrix A. Dann besitzt $\mathbb{C}^n$ eine orthogonale Zerlegung in die Eigenräume von A:

$$\mathbb{C}^n = \oplus_{\nu=1}^s V_{\lambda_\nu}.$$

Ist $P(\lambda) = \prod_{\nu=1}^s (\lambda - \lambda_\nu)^{r_\nu}$ das charakteristische Polynom von A, so ist $r_\nu = \dim V_{\lambda_\nu}$, die geometrischen und algebraischen Vielfachheiten der Eigenwerte stimmen also überein.

Beweis

Die Orthogonalität der Eigenräume wurde bereits in Satz 7.3.17 nachgewiesen.

Sei $B \in U(n)$ derart, daß $\hat{A} := \bar{B}^t \circ A \circ B$ Diagonalgestalt hat, die Spalten $B = \left(\vec{b}_1, \ldots, \vec{b}_n\right)$ bilden eine Orthonormalbasis von Eigenvektoren von A, in der Diagonalen von $\hat{A}$ treten die Eigenwerte von $\hat{A}$ und damit die von A entsprechend ihrer geometrischen Vielfachheit $\dim V_{\lambda_\nu}$ auf; man beachte, daß sich geometrische Vielfachheiten bei Koordinatenwechseln nicht ändern. In $\hat{A}$ und folglich auch in A stimmen somit offensichtlich geometrische und algebraische Vielfachheiten r_ν überein: $r_\nu = \dim V_{\lambda_\nu}$. Indem man ggfs. die Spalten von B vertauscht, kann man erreichen:

$$\begin{array}{ll} \vec{b}_1, \ldots, \vec{b}_{r_1} & \text{sind Eigenvektoren zu } \lambda_1, \\ \vec{b}_{r_1+1}, \ldots, \vec{b}_{r_1+r_2} & \text{sind Eigenvektoren zu } \lambda_2, \\ & \vdots \\ \vec{b}_{r_1+\ldots+r_{s-1}+1}, \ldots, \vec{b}_n & \text{sind Eigenvektoren zu } \lambda_s. \end{array}$$

Da jedes Element $\vec{z} \in \mathbb{C}^n$ eine Darstellung in der Basis $\{\vec{b}_1, \ldots, \vec{b}_n\}$ besitzt, folgt durch entsprechende Klammerung die Zerlegung von $\vec{z}$ in Komponenten in den jeweiligen Eigenräumen:

$$\begin{aligned} \vec{z} &= \sum_{\nu=1}^n c_\nu \vec{b}_\nu = \left(\sum_{\nu=1}^{r_1} c_\nu \vec{b}_\nu\right) + \left(\sum_{\nu=r_1+1}^{r_1+r_2} c_\nu \vec{b}_\nu\right) + \ldots + \left(\sum_{\nu=r_1+\ldots+r_{s-1}+1}^{n} c_\nu \vec{b}_\nu\right) \\ &=: \vec{z}_1 + \ldots + \vec{z}_s, \qquad \text{mit } \vec{z}_\nu \in V_{\lambda_\nu}. \end{aligned}$$

Die Eindeutigkeit dieser Darstellung folgt aus der Orthogonalität der Eigenräume. □

Die Hauptachsentransformation für Hermitesche und reell symmetrische Matrizen

Wir erhalten den folgenden Satz direkt aus dem Diagonalisierungsresultat für normale Matrizen Satz 7.4.8 zusammen mit der Beobachtung aus Satz 7.3.18, daß Eigenwerte Hermitescher Matrizen stets reell sind:

7.4.12 Satz (Hauptachsentransformation)

Zu jeder Hermiteschen $n \times n$-Matrix A existiert eine unitäre Matrix $B \in U(n)$ derart, daß $\hat{A} := \bar{B}^t \circ A \circ B$ eine reelle Diagonalmatrix ist.

Man erhält das reelle Analogon, indem man auf Satz 7.4.10 anstatt auf Satz 7.4.8 verweist.

7.4.13 Satz (Hauptachsentransformation)

Sei A eine symmetrische reelle $n \times n$-Matrix. Dann existiert eine Drehung $B \in SO(n, \mathbb{R})$ derart, daß $\hat{A} := B^t \circ A \circ B$ eine (reelle) Diagonalmatrix ist.

7.4.14 Bemerkung

Um den Begriff „Hauptachsentransformation" zu erläutern, betrachten wir zunächst eine reelle $n \times n$-Diagonalmatrix $\hat{A}$ mit den (nicht notwendigerweise verschiedenen) Diagonalelementen $\lambda_1, \ldots, \lambda_n$. Unter dieser Abbildung wird die Einheitskugel

$$K = \left\{ \vec{x} \in \mathbb{R}^n : \; \| \vec{x} \|^2 \leq 1 \right\}$$

auf ein Ellipsoid

$$E = \left\{ \vec{x} \in \mathbb{R}^n : \; \sum_{\nu=1}^{n} \frac{x_\nu^2}{\lambda_\nu^2} \leq 1 \right\}$$

abgebildet. Die Basisvektoren $\lambda_\nu \vec{e}_\nu$ sind die Hauptachsen dieses Ellipsoids.

Sei nun $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix und $B \in SO(n, \mathbb{R})$ eine Transformationsmatrix derart, daß die Abbildung in diesen Koordinaten Diagonalgestalt $\hat{A} := B^t \circ A \circ B$ wie oben hat. Man kann also in $\mathbb{R}^n$ eine Orthonormalbasis $B = \left(\vec{b}_1, \ldots, \vec{b}_n \right)$ „einzeichnen", so daß unter der Abbildung A die Einheitskugel K in ein Ellipsoid mit den Basisvektoren $\lambda_\nu \vec{b}_\nu$ als Hauptachsen übergeht.

7.4.15 Beispiel

Eine wichtige Anwendung des Diagonalisierungssatzes für reell symmetrische Matrizen findet man in der Theorie der Drehbewegung starrer Körper. Für eine ausführliche Darstellung verweisen wir auf die Lehrbücher der Mechanik wie etwa das von Á. Budó [1, §§ 51 ff].

Wir stellen uns einen solchen Körper vor und legen den Ursprung und ein orthogonales Koordinatensystem in dessen Schwerpunkt. Wir betrachten Drehbewegungen nur um Achsen, die durch diesen Ursprung gehen, um insgesamt resultierende *Kräfte* auf diese Achsen zu vermeiden. Der Zusammenhang zwischen dem Drehimpulsvektor $\vec{L}$ und der Winkelgeschwindigkeit $\vec{\omega}$ wird durch den „Trägheitstensor"

$$J = \begin{pmatrix} j_{11} & -j_{12} & -j_{13} \\ -j_{21} & j_{22} & -j_{23} \\ -j_{31} & -j_{32} & j_{33} \end{pmatrix} \in \mathbb{R}^{3 \times 3}$$

vermittelt:

$$\vec{L} = J \circ \vec{\omega}.$$

Dabei sind die Diagonalelemente $j_{\nu\nu}$ die „Trägheitsmomente“ bzgl. der Rotation um die x_ν-Achse und $j_{\mu\nu}$ für $\mu \neq \nu$ „Deviationsmomente“. Die Deviationsmomente sind ein Maß dafür, wie sehr bei der Rotation um die Koordinatenachsen *Drehmomente* auf diese Achsen auftreten, die deren Lage zu ändern versuchen. Bei realen Bewegungen von Maschinen oder Fahrzeugen müssen diese Momente von den Lagern der Achse kompensiert werden und führen zu einem „unrunden“ Lauf und erhöhtem Verschleiß. Deshalb ist man bemüht, die Drehachse und damit $\vec{\omega}$ so einzurichten, daß Drehimpuls und Winkelgeschwindigkeit parallel sind. Es sind also die Eigenvektoren des Trägheitstensors J zu bestimmen.

Den definierenden Gleichungen für dessen Komponenten entnimmt man (vgl. [1, § 51]), daß der Trägheitstensor J symmetrisch ist. Er besitzt somit eine Orthonormalbasis $\left\{\vec{b}_1, \vec{b}_2, \vec{b}_3\right\}$ aus Eigenvektoren, den „Hauptträgheitsachsen“ des Körpers im Schwerpunkt. Führt man die durch $B = \left(\vec{b}_1, \vec{b}_2, \vec{b}_3\right)$ gegebenen Koordinaten ein, so lautet der Trägheitstensor in diesen Koordinaten:

$$\hat{J} := B^t \circ J \circ B = \begin{pmatrix} \hat{j}_{11} & 0 & 0 \\ 0 & \hat{j}_{22} & 0 \\ 0 & 0 & \hat{j}_{33} \end{pmatrix}.$$

Bezüglich der Hauptträgheitsachsen $\vec{b}_1, \vec{b}_2, \vec{b}_3$ treten keine Deviationsmomente auf.

> Unter „Auswuchten“ eines Körpers versteht man dessen Veränderung derart, daß die (fest vorgegebene) Drehachse zu einer Hauptträgheitsachse im Schwerpunkt wird.

7.4.16 Beispiele

Wir setzen hier die Behandlung der Beispiele 7.3.19 fort.

(a) Für $A = \begin{pmatrix} 1 & -3i \\ 3i & 1 \end{pmatrix}$ hatten wir dort $\lambda_1 = 4$ und $\lambda_2 = -2$ als Eigenwerte bestimmt. Zugehörige Eigenvektoren findet man als nichttriviale Lösungen der Gleichungssysteme

$$\vec{0} = (A - \lambda_1 E) \circ \vec{x}_1 = \begin{pmatrix} -3 & -3i \\ 3i & -3 \end{pmatrix} \circ \vec{x}_1$$

und

$$\vec{0} = (A - \lambda_2 E) \circ \vec{x}_2 = \begin{pmatrix} 3 & -3i \\ 3i & 3 \end{pmatrix} \circ \vec{x}_2.$$

Man berechnet sofort $\vec{x}_1 = (-i, 1)^t$ und $\vec{x}_2 = (i, 1)^t$ als Lösungsvektoren. Durch Normieren erhält man die Transformationsmatrix

$$B = \begin{pmatrix} -i\frac{\sqrt{2}}{2} & i\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} = \frac{\sqrt{2}}{2}\begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}.$$

Hiermit gilt

$$\begin{aligned} \hat{A} = \bar{B}^t \circ A \circ B &= \begin{pmatrix} i\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -i\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \circ \begin{pmatrix} 1 & -3i \\ 3i & 1 \end{pmatrix} \circ \begin{pmatrix} -i\frac{\sqrt{2}}{2} & i\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}. \end{aligned}$$

(b) Die Eigenwerte von

$$A = \begin{pmatrix} \frac{7}{6} & -\frac{2}{3} & \frac{1}{6} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{1}{6} & -\frac{2}{3} & \frac{7}{6} \end{pmatrix}$$

lauten $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 2$. Als zugehörige Eigenvektoren bestimmt man mit dem Gaußschen Algorithmus beispielsweise $\vec{x}_1 = (1, 2, 1)^t$, $\vec{x}_2 = (-1, 0, 1)^t$ und $\vec{x}_3 = (1, -1, 1)^t$. Durch Normieren gelangt man zu der Transformationsmatrix

$$B = \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 & -\sqrt{3} & \sqrt{2} \\ 2 & 0 & -\sqrt{2} \\ 1 & \sqrt{3} & \sqrt{2} \end{pmatrix},$$

es ist

$$\det(B) = \frac{1}{6^{3/2}} \det\begin{pmatrix} 1 & -\sqrt{3} & \sqrt{2} \\ 0 & 2\sqrt{3} & -3\sqrt{2} \\ 0 & 2\sqrt{3} & 0 \end{pmatrix} = \frac{1}{6^{3/2}} \cdot 6\sqrt{6} = 1.$$

Ohne daß eine Spalte von B durch ihr Negatives ersetzt werden müßte, haben wir bereits $B \in SO(3)$, so daß also B eine Drehung des $\mathbb{R}^3$ ist. In den durch B gegebenen Koordinaten lautet nun A:

$$\begin{aligned} \hat{A} &= B^t \circ A \circ B \\ &= \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix} \circ \begin{pmatrix} \frac{7}{6} & -\frac{2}{3} & \frac{1}{6} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{1}{6} & -\frac{2}{3} & \frac{7}{6} \end{pmatrix} \circ \begin{pmatrix} \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \end{aligned}$$

7.4.17 Beispiel

In der Analysis benötigt man beispielsweise bei der Behandlung gewöhnlicher Differentialgleichungssysteme die „Exponentialfunktion“ quadratischer n-reihiger Matrizen A:

$$\exp(A) := \sum_{k=0}^{\infty} \frac{1}{k!} A^k. \qquad (*)$$

Die Partialsummen $S_N := \left(s_{\mu\nu}^{(N)}\right)_{\mu,\nu=1,\dots,n} := \sum_{k=0}^{N} \frac{1}{k!} A^k$ sind jeweils $n \times n$-Matrizen, von deren Komponenten $s_{\mu\nu}^{(N)}$ man in der Analysis Konvergenz nachweist. Man findet Zahlen $s_{\mu\nu} := \lim_{N\to\infty} s_{\mu\nu}^{(N)} \in \mathbb{C}$ und setzt $\exp(A) := (s_{\mu\nu})_{\mu,\nu=1,\dots,n}$, die Definition $(*)$ ist also im Sinne komponentenweiser Konvergenz zu verstehen. Wir wollen uns hier aber nicht mit Konvergenz, sondern mit der Frage beschäftigen, ob man die Komponenten von $\exp(A)$ auf einfache Weise aus den Komponenten von A bestimmen kann. Im Falle normaler und damit insbesondere auch Hermitescher Matrizen ist das tatsächlich möglich.

Sei A eine normale $n \times n$-Matrix und $B \in U(n)$ derart, daß wir Diagonalgestalt

$$\hat{A} := \bar{B}^t \circ A \circ B = \begin{pmatrix} \lambda_1 & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & \lambda_n \end{pmatrix}, \qquad \lambda_1, \dots, \lambda_n \in \mathbb{C},$$

erreichen. Es folgt:

$$\begin{aligned}
A &= B \circ \hat{A} \circ \bar{B}^t, \\
A^k &= \left(B \circ \hat{A} \circ \bar{B}^t\right)^k \\
&= B \circ \hat{A} \circ \left(\bar{B}^t \circ B\right) \circ \hat{A} \circ \bar{B}^t \circ \ldots \circ B \circ \hat{A} \circ \bar{B}^t \\
&= B \circ \hat{A}^k \circ \bar{B}^t = B \circ \begin{pmatrix} \lambda_1^k & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & \lambda_n^k \end{pmatrix} \circ \bar{B}^t, \\
\exp(A) &= \sum_{k=0}^{\infty} \frac{1}{k!} B \circ \begin{pmatrix} \lambda_1^k & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & \lambda_n^k \end{pmatrix} \circ \bar{B}^t \\
&= B \circ \left(\sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} \lambda_1^k & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & \lambda_n^k \end{pmatrix} \right) \circ \bar{B}^t \\
&= B \circ \begin{pmatrix} \sum_{k=0}^{\infty} \frac{1}{k!} \lambda_1^k & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & \sum_{k=0}^{\infty} \frac{1}{k!} \lambda_n^k \end{pmatrix} \circ \bar{B}^t
\end{aligned}$$

$$= B \circ \begin{pmatrix} \exp(\lambda_1) & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & \exp(\lambda_n) \end{pmatrix} \circ \bar{B}^t.$$

Man erhält die Exponentialfunktion normaler Matrizen, indem man die gewöhnliche komplexe Exponentialfunktion der Eigenwerte berechnet und diese Diagonalmatrix mittels B und $\bar{B}^t$ „zurück"-transformiert. Die Eigenwerte von $\exp(A)$ sind ebenfalls $\exp(\lambda_1), \ldots, \exp(\lambda_n)$. Die Matrizen A und $\exp(A)$ haben mit den Spalten von B dieselben Eigenvektoren.

Wir berechnen $\exp(A)$ für die Hermitesche 2×2-Matrix $A = \begin{pmatrix} 1 & -3i \\ 3i & 1 \end{pmatrix}$ aus dem vorhergehenden Beispiel. Diese hatten wir mittels $B = \frac{\sqrt{2}}{2} \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \in U(2)$ diagonalisiert: $\hat{A} = \bar{B}^t \circ A \circ B = \begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix}$. Daraus erhalten wir:

$$\begin{aligned} \exp(A) &= B \circ \exp(\hat{A}) \circ \bar{B}^t \\ &= \frac{1}{2} \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \circ \begin{pmatrix} e^4 & 0 \\ 0 & e^{-2} \end{pmatrix} \circ \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} e^4 + e^{-2} & -i\left(e^4 - e^{-2}\right) \\ i\left(e^4 - e^{-2}\right) & e^4 + e^{-2} \end{pmatrix}. \end{aligned}$$

Sesquilinearformen

Die Abbildung, die jedem Paar von Vektoren $\vec{z}, \vec{w} \in \mathbb{C}^n$ die komplexe Zahl $(\vec{z}, A \circ \vec{w})$ zuordnet, hat für Hermitesche Matrizen A ähnliche Eigenschaften wie das komplexe Skalarprodukt. Sie ist sesquilinear und Hermitesch:

$$(\vec{z}, A \circ \vec{w}) = (\bar{A}^t \circ \vec{z}, \vec{w}) = (A \circ \vec{z}, \vec{w}) = \overline{(\vec{w}, A \circ \vec{z})},$$

und wegen $(\vec{z}, A \circ \vec{z}) = \overline{(\vec{z}, A \circ \vec{z})}$ ist die entsprechende quadratische Form stets reellwertig. Die Positivität dieser Terme kann aber im allgemeinen nicht gezeigt werden, sondern ist nur von einer Unterklassse der Hermiteschen Matrizen erfüllt:

7.4.18 Definition

(a) *Eine Hermitesche Matrix* $A \in \mathbb{C}^{n \times n}$ *heißt* **positiv definit**, *falls* $(\vec{z}, A \circ \vec{z}) > 0$ *für alle* $\vec{z} \in \mathbb{C}^n \setminus \{\vec{0}\}$ *gilt.*

(b) *Eine symmetrische Matrix* $A \in \mathbb{R}^{n \times n}$ *heißt* **positiv definit**, *falls* $\vec{x} \cdot (A \circ \vec{x}) > 0$ *für alle* $\vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$ *gilt.*

Mit Hilfe der Hauptachsentransformation kann man die positive Definitheit einer Matrix leicht erkennen.

7.4.19 Satz

Eine komplexe Hermitesche (bzw. reelle symmetrische) $n \times n$-Matrix A ist genau dann positiv definit, wenn alle Eigenwerte von A positiv sind.

Beweis

Hier werden nur komplexe Hermitesche Matrizen betrachtet; der reelle Fall ist genauso zu behandeln.

Es gibt eine unitäre Matrix $B \in U(n)$ derart, daß $\hat{A} = \bar{B}^t \circ A \circ B$ eine Diagonalmatrix ist:

$$\hat{A} = \begin{pmatrix} \lambda_1 & & \mathcal{O} \\ & \ddots & \\ \mathcal{O} & & \lambda_n \end{pmatrix}.$$

Die (nicht notwendigerweise verschiedenen) reellen Zahlen λ_ν $(\nu = 1, \ldots, n)$ sind die Eigenwerte von $\hat{A}$ und daher auch von A. Für alle $\vec{z} \in \mathbb{C}^n$ gilt mit der Setzung $\vec{w} := \bar{B}^t \circ \vec{z}$:

$$(\vec{z}, A \circ \vec{z}\,) = \bar{\vec{z}}^{\,t} \circ B \circ \hat{A} \circ \bar{B}^t \circ \vec{z} = \bar{\vec{w}}^{\,t} \circ \hat{A} \circ \vec{w} = \sum_{\nu=1}^{n} \lambda_\nu |w_\nu|^2.$$

Da $\vec{z} = \vec{0} \Leftrightarrow \vec{w} = B \circ \vec{z} = \vec{0}$, ist die Positivität der von A erzeugten quadratischen Form

$$(\vec{z}, A \circ \vec{z}\,) > 0 \text{ für alle } \vec{z} \in \mathbb{C}^n \setminus \{\vec{0}\}$$

äquivalent zur Positivität sämtlicher Eigenwerte

$$\lambda_\nu > 0, \qquad \nu = 1, \ldots, n.$$

□

7.4.20 Folgerung

Die reelle symmetrische $n \times n$-Matrix A sei positiv definit. Dann besitzt A genau eine symmetrische positiv definite Quadratwurzel $C \in \mathbb{R}^{n \times n}$, für die also $C^2 = A$ gilt. Man schreibt auch $C =: A^{1/2}$.

Beweis

Wir diagonalisieren A mittels einer Drehung $B \in SO(n, \mathbb{R})$:

$$\hat{A} = B^t \circ A \circ B = \begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & \mathcal{O} & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_s & & \\ & \mathcal{O} & & & & \ddots & \\ & & & & & & \lambda_s \end{pmatrix}.$$

Dabei sind $\lambda_1, \ldots, \lambda_s$ sämtliche paarweise verschiedenen Eigenwerte von A und von $\hat{A}$; jeder Eigenwert λ_ν tritt in $\hat{A}$ entsprechend seiner Vielfachheit r_ν-mal auf; $r_1 + \ldots + r_s = n$. Wie wir soeben gezeigt haben, sind wegen der positiven Definitheit von A sämtliche Eigenwerte positiv: $\lambda_\nu > 0$. Wir setzen

$$\hat{C} := \begin{pmatrix} \sqrt{\lambda_1} & & & & & & \\ & \ddots & & & & \mathcal{O} & \\ & & \sqrt{\lambda_1} & & & & \\ & & & \ddots & & & \\ & & & & \sqrt{\lambda_s} & & \\ & \mathcal{O} & & & & \ddots & \\ & & & & & & \sqrt{\lambda_s} \end{pmatrix} \text{ und } C := B \circ \hat{C} \circ B^t;$$

es ist

$$C^2 = B \circ \hat{C} \circ B^t \circ B \circ \hat{C} \circ B^t = B \circ \hat{C}^2 \circ B^t = B \circ \hat{A} \circ B^t = A.$$

Die einander entsprechenden Eigenräume von C und von A werden jeweils von denselben Spalten von B erzeugt und stimmen deshalb auf Grund unserer Konstruktion überein:

$$V_{\lambda_\nu}(A) = V_{\sqrt{\lambda_\nu}}(C). \tag{$*$}$$

Es bleibt die Eindeutigkeit der Quadratwurzel in der angegebenen Matrizenklasse nachzuweisen. Sei also $\tilde{C}$ irgendeine positiv definite symmetrische $n \times n$-Matrix mit $\tilde{C}^2 = A$. Wir wollen zeigen, daß die Gleichheit $(*)$ auch für die Eigenräume von $\tilde{C}$ erfüllt ist. Sei dazu $\vec{x} \in \mathbb{R}^n$ Eigenvektor von $\tilde{C}$ zum Eigenwert μ: $\tilde{C} \circ \vec{x} = \mu\,\vec{x}$. Da $\tilde{C}$ als positiv definit vorausgesetzt ist, haben wir $\mu > 0$. Aus $A \circ \vec{x} = \tilde{C} \circ \tilde{C} \circ \vec{x} = \tilde{C} \circ (\mu\,\vec{x}) = \mu^2\,\vec{x}$ folgt, daß μ^2 mit einem Eigenwert λ_ν von A übereinstimmen muß, wegen $\mu > 0$ ist $\mu = \sqrt{\lambda_\nu}$. Folglich sind unter $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_s}$ sämtliche paarweise verschiedenen Eigenwerte von $\tilde{C}$, und das gerade benutzte Argument zeigt auch $V_{\sqrt{\lambda_\nu}}(\tilde{C}) \subset V_{\lambda_\nu}(A)$. Da wegen Satz 7.4.11 bzw. dessen reeller Entsprechung aber die orthogonale Zerlegung

$$\mathbb{R}^n = \oplus_{\nu=1}^s V_{\sqrt{\lambda_\nu}}(\tilde{C}) = \oplus_{\nu=1}^s V_{\lambda_\nu}(A)$$

gilt, erhalten wir die Gleichheit der Eigenräume $V_{\sqrt{\lambda_\nu}}(\tilde{C}) = V_{\lambda_\nu}(A)$. Insbesondere sind genau $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_s}$ die Eigenwerte von $\tilde{C}$, und in Verbindung mit $(*)$ folgt:

$$V_{\sqrt{\lambda_\nu}}(\tilde{C}) = V_{\sqrt{\lambda_\nu}}(C).$$

Für alle

$$\vec{x} = \sum_{\nu=1}^{s} \vec{x}_\nu \in \mathbb{R}^n = \oplus_{\nu=1}^s V_{\sqrt{\lambda_\nu}}(\tilde{C}) = \oplus_{\nu=1}^s V_{\sqrt{\lambda_\nu}}(C)$$

folgt dann

$$\tilde{C} \circ \vec{x} = \sum_{\nu=1}^{s} \tilde{C} \circ \vec{x}_\nu = \sum_{\nu=1}^{s} \sqrt{\lambda_\nu}\, \vec{x}_\nu = \sum_{\nu=1}^{s} C \circ \vec{x}_\nu = C \circ \vec{x},$$

die positiv definiten Quadratwurzeln C und $\tilde{C}$ von A stimmen überein. □

Diagonalisierbarkeit unitärer Matrizen, Normalform für reell orthogonale Matrizen

Hier erhalten wir das Diagonalisierungsresultat ebenfalls durch direkten Verweis auf den entsprechenden Satz 7.4.8 für normale Matrizen und auf Satz 7.3.20.

7.4.21 Satz

Zu jeder unitären Matrix $A \in U(n)$ existiert eine unitäre Transformationsmatrix $B \in U(n)$ derart, daß $\hat{A} := \bar{B}^t \circ A \circ B$ eine Diagonalmatrix ist. Die Diagonalelemente von $\hat{A}$ haben sämtlich den Betrag 1.

Wir wollen im Folgenden noch reelle orthogonale Matrizen untersuchen. Der Fall zweier Raumdimensionen läßt sich elementar diskutieren und wird die Richtung für unser angestrebtes allgemeines reelles Resultat anzeigen.

7.4.22 Beispiele

(a) Seien $\sin_a, \cos_a$ die aus der Analysis bekannten im Bogenmaß erklärten trigonometrischen Funktionen. Für eine Drehung des $\mathbb{R}^2$

$$A = \begin{pmatrix} \cos_a(\varphi) & -\sin_a(\varphi) \\ \sin_a(\varphi) & \cos_a(\varphi) \end{pmatrix}$$

gilt

$$P(\lambda) = (\cos_a(\varphi) - \lambda)^2 + \sin_a^2(\varphi) = 1 - 2\lambda \cos_a(\varphi) + \lambda^2.$$

Die Eigenwerte von A sind $\lambda_1 = e^{i\varphi}$, $\lambda_2 = e^{-i\varphi}$; diese sind für $\varphi \in]0, 2\pi[\setminus \{\pi\}$ nicht reell und verschieden. Im Reellen können wir A also nicht diagonalisieren. Es existiert jedoch eine komplexe unitäre 2×2-Matrix B derart, daß $\hat{A} = \bar{B}^t \circ A \circ B = \begin{pmatrix} e^{i\varphi} & 0 \\ 0 & e^{-i\varphi} \end{pmatrix}$ ist.

(b) Eine Drehspiegelung des $\mathbb{R}^2$ wird durch eine Matrix

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix}, \quad a^2 + b^2 = 1,$$

gegeben. Es ist

$$P(\lambda) = (a - \lambda)(-a - \lambda) - b^2 = \lambda^2 - (a^2 + b^2) = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1).$$

Weil die Eigenwerte von A mit $\lambda_1 = 1$ und $\lambda_2 = -1$ sämtlich reell sind, kann man aus Satz 7.4.10 die Existenz einer orthogonalen Matrix $B \in SO(2)$ folgern, so daß $\hat{A} := B^t \circ A \circ B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ gilt. Bezüglich eines geeigneten Koordinatensystems ist die Drehspiegelung A nur eine Spiegelung an einer Koordinatenachse.

Um Satz 7.4.21 auch für den Fall reeller orthogonaler Matrizen $A \in O(n)$ nutzen zu können, gehen wir von der im Beweis von Satz 7.3.21 gemachten Beobachtung aus, daß die nichtreellen Eigenwerte in Paaren zueinander konjugierter Eigenwerte derselben (geometrischen und algebraischen) Vielfachheit auftreten: $\lambda_1, \bar{\lambda}_1, \ldots, \lambda_k, \bar{\lambda}_k \in \mathbb{C} \setminus \mathbb{R}$ mit einer geeigneten Zahl $k \leq n/2$; die Eigenwerte werden entsprechend ihrer Vielfachheit aufgeführt. Dabei können wir erreichen, daß unter $\lambda_1, \lambda_2, \ldots, \lambda_k$ nicht zwei Eigenwerte zueinander konjugiert sind: $\lambda_\nu \neq \bar{\lambda}_\mu$, $\mu, \nu = 1, \ldots, k$. An das Ende dieser Eigenwertliste schreiben wir die reellen Eigenwerte $\lambda_{2k+1}, \ldots, \lambda_n \in \{\pm 1\} \subset \mathbb{R}$. Bei den Eigenvektoren können wir eine ganz ähnliche Beobachtung machen: Sind $\vec{b}_\nu \in \mathbb{C}^n$ $(1 \leq \nu \leq k)$ paarweise orthogonale normierte Eigenvektoren zu den Eigenwerten λ_ν $(1 \leq \nu \leq k)$, so folgt aus der Reellwertigkeit der Komponenten von A:

$$(A - \lambda_\nu E) \circ \vec{b}_\nu = \vec{0} \;\Rightarrow\; \overline{(A - \lambda_\nu E)} \circ \bar{\vec{b}}_\nu = \vec{0} \;\Rightarrow\; (A - \bar{\lambda}_\nu E) \circ \bar{\vec{b}}_\nu = \vec{0};$$

die konjugierten Vektoren $\bar{\vec{b}}_\nu$ $(1 \leq \nu \leq k)$ sind also normierte Eigenvektoren zu den Eigenwerten $\bar{\lambda}_\nu$. Wir wählen nun die Transformationsmatrix zur Diagonalisierung von A gemäß

$$B = \left(\vec{b}_1, \bar{\vec{b}}_1, \ldots, \vec{b}_k, \bar{\vec{b}}_k, \vec{b}_{2k+1}, \ldots, \vec{b}_n\right) \in U(n),$$

dabei sind $\vec{b}_{2k+1}, \ldots, \vec{b}_n$ paarweise orthogonale normierte *reelle* Eigenvektoren zu den reellen Eigenwerten $\lambda_{2k+1}, \ldots, \lambda_n$. Wir haben noch zu beachten, daß aus der Orthogonalität von $\vec{b}_\nu, \vec{b}_\mu$, $1 \leq \nu \neq \mu \leq k$ auch die Orthogonalität der konjugierten Vektoren folgt: $\left(\bar{\vec{b}}_\nu, \bar{\vec{b}}_\mu\right) = \overline{\left(\vec{b}_\nu, \vec{b}_\mu\right)} = 0$. Wegen der Verschiedenheit $\lambda_\nu \neq \bar{\lambda}_\mu$ gilt für $\mu, \nu = 1, \ldots, k$ ohnehin: $\left(\vec{b}_\nu, \bar{\vec{b}}_\mu\right) = 0$. Bezüglich dieser Transformation B lautet die Diagonalmatrix

$$\hat{A} = \bar{B}^t \circ A \circ B = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & & & \cdots & 0 \\ 0 & \bar{\lambda}_1 & 0 & \cdots & & & \cdots & 0 \\ \vdots & & \ddots & & & & & \vdots \\ & \cdots & 0 & \lambda_k & 0 & 0 & \cdots & \\ & \cdots & 0 & 0 & \bar{\lambda}_k & 0 & \cdots & \\ & \cdots & 0 & 0 & 0 & \lambda_{2k+1} & \cdots & \\ \vdots & & & & & & \ddots & \vdots \\ 0 & \cdots & & & & & \cdots & \lambda_n \end{pmatrix}.$$

Wir gelangen nun zu einer reellen orthogonalen Transformationsmatrix C, indem wir jeweils bei Paaren zueinander konjugierter Eigenvektoren $\vec{b}_\nu, \bar{\vec{b}}_\nu$, $\nu = 1, \ldots, k$ zu deren normierten Real- und Imaginärteilen übergehen:

$$\vec{c}_{2\nu-1} := \frac{\sqrt{2}}{2}\left(\vec{b}_\nu + \bar{\vec{b}}_\nu\right), \qquad \vec{c}_{2\nu} := \frac{\sqrt{2}}{2i}\left(\vec{b}_\nu - \bar{\vec{b}}_\nu\right) \in \mathbb{R}^n,$$

wegen der Orthogonalität von $\vec{b}_\nu$ und $\bar{\vec{b}}_\nu$ handelt es sich dabei nicht um die Nullvektoren. Ferner ist in $\mathbb{C}^n$ auf Grund dieser Orthogonalität:

$$\begin{aligned}
\left(\vec{b}_\nu + \bar{\vec{b}}_\nu, \vec{b}_\nu - \bar{\vec{b}}_\nu\right) &= \left(\vec{b}_\nu, \vec{b}_\nu\right) - \left(\vec{b}_\nu, \bar{\vec{b}}_\nu\right) + \left(\bar{\vec{b}}_\nu, \vec{b}_\nu\right) - \left(\bar{\vec{b}}_\nu, \bar{\vec{b}}_\nu\right) \\
&= \| \vec{b}_\nu \|^2 - \overline{\left(\vec{b}_\nu, \vec{b}_\nu\right)} = 0,
\end{aligned}$$

also sind $\vec{c}_{2\nu-1}$ und $\vec{c}_{2\nu}$ in $\mathbb{R}^n$ zueinander orthogonal. Indem man die reellen Eigenvektoren unverändert läßt,

$$\vec{c}_{2k+1} := \vec{b}_{2k+1}, \ldots, \vec{c}_n := \vec{b}_n,$$

und nur ggfs. einen Eigenvektor durch sein Negatives ersetzt, erhält man mit

$$C := (\vec{c}_1, \ldots, \vec{c}_{2k}, \vec{c}_{2k+1}, \ldots, \vec{c}_n) \in SO(n, \mathbb{R})$$

eine reelle Transformationsmatrix. Um die Matrix der Abbildung A in den durch C gegebenen Koordinaten zu ermitteln, berechnen wir für $\nu = 1, \ldots, k$:

$$\begin{aligned}
A \circ \vec{c}_{2\nu-1} &= A \circ \frac{\sqrt{2}}{2}\left(\vec{b}_\nu + \bar{\vec{b}}_\nu\right) = \frac{\sqrt{2}}{2}\left(\lambda_\nu \vec{b}_\nu + \bar{\lambda}_\nu \bar{\vec{b}}_\nu\right) \\
&= \frac{\sqrt{2}}{4}\left(\left(\lambda_\nu + \bar{\lambda}_\nu\right)\left(\vec{b}_\nu + \bar{\vec{b}}_\nu\right) + \left(\lambda_\nu - \bar{\lambda}_\nu\right)\left(\vec{b}_\nu - \bar{\vec{b}}_\nu\right)\right) \\
&= \mathrm{Re}(\lambda_\nu) \cdot \vec{c}_{2\nu-1} - \mathrm{Im}(\lambda_\nu) \cdot \vec{c}_{2\nu}, \\
A \circ \vec{c}_{2\nu} &= A \circ \frac{\sqrt{2}}{2i}\left(\vec{b}_\nu - \bar{\vec{b}}_\nu\right) = \frac{\sqrt{2}}{2i}\left(\lambda_\nu \vec{b}_\nu - \bar{\lambda}_\nu \bar{\vec{b}}_\nu\right) \\
&= \frac{\sqrt{2}}{4i}\left(\left(\lambda_\nu - \bar{\lambda}_\nu\right)\left(\vec{b}_\nu + \bar{\vec{b}}_\nu\right) + \left(\lambda_\nu + \bar{\lambda}_\nu\right)\left(\vec{b}_\nu - \bar{\vec{b}}_\nu\right)\right) \\
&= \mathrm{Im}(\lambda_\nu) \cdot \vec{c}_{2\nu-1} + \mathrm{Re}(\lambda_\nu) \cdot \vec{c}_{2\nu}.
\end{aligned}$$

Für die Matrix

$$\tilde{A} := C^t \circ A \circ C,$$

die die Abbildung $A : \mathbb{R}^n \to \mathbb{R}^n$ in den durch die orthogonale Matrix C gegebenen

Koordinaten beschreibt, haben wir also gefunden:

$$\tilde{A} = \begin{pmatrix} \mathrm{Re}(\lambda_1) & \mathrm{Im}(\lambda_1) & & & & & & \\ -\mathrm{Im}(\lambda_1) & \mathrm{Re}(\lambda_1) & & & & \mathcal{O} & & \\ & & \ddots & & & & & \\ & & & \mathrm{Re}(\lambda_k) & \mathrm{Im}(\lambda_k) & & & \\ & & & -\mathrm{Im}(\lambda_k) & \mathrm{Re}(\lambda_k) & & & \\ & & & & & \lambda_{2k+1} & & \\ & & \mathcal{O} & & & & \ddots & \\ & & & & & & & \lambda_n \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

Diese Normalform zeigt, daß für Matrizen $A \in SO(3,\mathbb{R})$ der Eigenvektor zum Eigenwert $+1$ als Drehachse interpretiert werden kann: Indem man ein geeignetes Koordinatensystem einführt, erhält die Abbildung A in diesen Koordinaten die folgende Darstellung mit geeigneten reellen Zahlen a, b, $a^2 + b^2 = 1$:

$$\tilde{A} = \begin{pmatrix} a & -b & 0 \\ b & a & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\Theta & -\sin\Theta & 0 \\ \sin\Theta & \cos\Theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

die Zahl $\Theta \in [0, 360[$ ist geeignet zu wählen. Die Abbildung $\tilde{A}$ läßt die dritte Komponente fest, die ersten beiden Komponenten werden durch die 2×2-Untermatrix $\begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix}$ abgebildet. Nach unseren Überlegungen am Ende von Kapitel 6.1 (insbesondere Satz 6.1.12 und Definition 6.1.15) bedeutet das eine Drehung der x_1-x_2-Ebene um den Winkel Θ, die x_3-Achse (in den ursprünglichen Koordinaten dem Eigenvektor zum Eigenwert 1 entsprechend) ist die Drehachse.

7.4.23 Beispiel

Wir bestimmen sowohl die komplexe als auch die reelle Normalform für die schon in Beispiel 7.3.22 betrachtete Matrix

$$A = \begin{pmatrix} \frac{1}{2} + \frac{\sqrt{2}}{4} & -\frac{1}{2} + \frac{\sqrt{2}}{4} & -\frac{1}{2} \\ -\frac{1}{2} + \frac{\sqrt{2}}{4} & \frac{1}{2} + \frac{\sqrt{2}}{4} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \in SO(3,\mathbb{R}).$$

Die Eigenwerte hatten wir als $\lambda_1 = \frac{\sqrt{2}}{2}(1+i)$, $\lambda_2 = \frac{\sqrt{2}}{2}(1-i)$ und $\lambda_3 = 1$ bestimmt, wir hatten auch schon die Drehachse als normierten Eigenvektor $\vec{b}_3 = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}, 0\right)^t$ zum Eigenwert 1 berechnet. Mittels des Gaußschen Algorithmus' findet man weiter, daß $\vec{b}_1 = \left(\frac{i}{2}, \frac{i}{2}, \frac{\sqrt{2}}{2}\right)^t$ und $\vec{b}_2 = \left(-\frac{i}{2}, -\frac{i}{2}, \frac{\sqrt{2}}{2}\right)^t$ normierte Eigenvektoren zu λ_1 und

λ_2 sind. Mit

$$B := \begin{pmatrix} \frac{i}{2} & -\frac{i}{2} & \frac{\sqrt{2}}{2} \\ \frac{i}{2} & -\frac{i}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \end{pmatrix}$$

erhalten wir die komplexe Diagonalform

$$\hat{A} = \bar{B}^t \circ A \circ B = \begin{pmatrix} \frac{\sqrt{2}}{2}(1+i) & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2}(1-i) & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Um die reelle Normalform zu erhalten, ersetzen wir, wie oben erläutert, die ersten beiden Spalten von B durch deren normierte Real- und Imaginärteile:

$$C := \begin{pmatrix} 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 1 & 0 & 0 \end{pmatrix} \in SO(3).$$

Um mit C eine Drehung, d. h. eine orthogonale Abbildung mit Determinante 1, zu erhalten, haben wir gleichzeitig die letzte Spalte durch ihr Negatives ersetzt. Damit berechnen wir

$$\tilde{A} := C^t \circ A \circ C = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Bei der Abbildung A handelt es sich also um eine Drehung um 315 Grad um die Drehachse $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right)^t$.

Literaturverzeichnis

[1] Budó, Ágoston, *Theoretische Mechanik*, 10. Auflage, Deutscher Verlag der Wissenschaften: Berlin 1980.

[2] Ebbinghaus, Heinz-Dieter, H. Hermes, F. Hirzebruch, M. Koecher, K. Mainzer, A. Prestel, R. Remmert, *Zahlen*, Springer-Verlag: Berlin etc. 1983.

[3] Fischer, Gerd, *Lineare Algebra*, 6. Auflage, Vieweg: Braunschweig, Wiesbaden 1980.

[4] Forster, Otto, *Analysis 1*, 3. Auflage, Vieweg: Braunschweig, Wiesbaden 1980.

[5] Grauert, Hans, Die Axiome der Elementargeometrie und die griechischen Theoreme, *Math. Semesterber.* **44** (1997), 19-36.

[6] Grauert, Hans, I. Lieb, *Differential- und Integralrechnung I*, 4. Auflage, Springer-Verlag: Berlin etc. 1976.

[7] Klingenberg, Wilhelm, *Lineare Algebra und Geometrie*, Springer-Verlag: Berlin etc. 1984.

[8] Koecher, Max, *Lineare Algebra und analytische Geometrie*, Springer-Verlag: Berlin etc. 1983.

[9] Kowalsky, Hans-Joachim, *Lineare Algebra*, 9. Auflage, de Gruyter: Berlin, New York 1979.

[10] Pickert, Günter, *Analytische Geometrie*, 4. Auflage, Akademische Verlagsgesellschaft Geest & Portig: Leipzig 1961.

[11] Rudin, Walter, *Analysis*, Oldenbourg Verlag: München, Wien 1998.

[12] Sperner, Emanuel, *Einführung in die Analytische Geometrie und Algebra*, 7. Auflage, Vandenhoeck & Ruprecht: Göttingen 1969.

[13] Tietz, Horst, *Lineare Geometrie*, Verlag Aschendorff: Münster 1967.

[14] van der Waerden, B. L., *Erwachende Wissenschaft*, 2. Auflage, Birkhäuser Verlag: Basel, Stuttgart 1966.

Symbolverzeichnis

$:=$	Die linke Seite der Gleichung wird durch die rechte Seite definiert.
$\stackrel{\text{def}}{=}$	Gleichheit auf Grund einer Definition
$\equiv$	identisch gleich; $f(x) \equiv 0$ bedeutet, daß $f(x) = 0$ für *alle* x gilt
$\Rightarrow$	Folgerung: Wenn die linke Aussage gilt, dann auch die rechte.
$\Leftrightarrow$	Äquivalenz zweier Aussagen
$\emptyset$	leere Menge
$x \in M$	x ist Element der Menge M
$x \notin M$	x ist nicht Element der Menge M
$M \subset N$	Teilmenge: $x \in M \Rightarrow x \in N$.
$M \supset N$	Obermenge: $N \subset M$
$M = N$	Gleichheit von Mengen: $M \subset N$ und $N \subset M$
$M \cap N$	$= \{x :\ x \in M$ und $x \in N\}$, Durchschnitt
$M \cup N$	$= \{x :\ x \in M$ oder $x \in N\}$, Vereinigung
$M \setminus N$	$= \{x \in M :\ x \notin N\}$, Differenzmenge
$M \times N$	$= \{(x, y) :\ x \in M, y \in N\}$, kartesisches Produkt
$\mathbb{N}$	Menge der natürlichen Zahlen $\{1, 2, 3, \ldots\}$
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \ldots\}$
$\mathbb{Z}$	Menge der ganzen Zahlen
$\mathbb{R}$	Menge der reellen Zahlen
$\mathbb{R}^+$	Menge der positiven reellen Zahlen $\{a \in \mathbb{R} :\ a > 0\}$
$\mathbb{R}_0^+$	Menge der nichtnegativen reellen Zahlen $\{a \in \mathbb{R} :\ a \geq 0\}$
$\mathbb{R}^*$	$= \mathbb{R} \setminus \{0\}$
$\max(a, b)$	die größere der reellen Zahlen a, b
$\min(a, b)$	die kleinere der reellen Zahlen a, b

$\sup$	Supremum
$\inf$	Infimum
$\sum$	Summenzeichen
$\prod$	Produktzeichen
$\mathbb{R}^n$	n-dimensionaler reeller Zahlenraum
$\vec{a}$	$= (a_1, \ldots, a_n) \in \mathbb{R}^n$, n-tupel reeller Zahlen
$\vec{0}$	$= (0, \ldots, 0) \in \mathbb{R}^n$
$\vec{e}_j$	Einheitsvektoren in $\mathbb{R}^n$
$V, V_1, V_2, \ldots$	allgemeine reelle Vektorräume
x, y	Elemente von V
$\left.\begin{array}{l} \text{Span}(x_1, \ldots, x_n) \\ \mathbb{R} \cdot (x_1, \ldots, x_n) \\ <x_1, \ldots, x_n> \\ [x_1, \ldots, x_n] \end{array}\right\}$	von $x_1, \ldots, x_n$ aufgespannter Untervektorraum
$\dim$	Dimension eines Vektorraums
E	anschauliche Ebene
R	anschaulicher Raum
O	Basispunkt, Aufpunkt in E bzw. R
$x = (P, Q)$	Pfeile, Pfeilvektoren in E bzw. R
(P, P)	Nullvektor in P
$(P, Q) \parallel (P', Q')$	Die Pfeile (P, Q) und (P', Q') sind voll parallel.
X	Menge aller Pfeile in E bzw. R
V	$= \{x = (O, P) : P \in E\}$, Vektorraum der Pfeile in O
V^3	$= \{x = (O, P) : P \in R\}$, Vektorraum der Pfeile in O
$\mathcal{O} = (O, O)$	Nullvektor in V bzw. V^3
$F : M_1 \to M_2$	Abbildung
$id_M, id : M \to M$	identische Abbildung
$F : M_1 \twoheadrightarrow M_2$	surjektive Abbildung
$F : M_1 \hookrightarrow M_2$	injektive Abbildung
$F : M_1 \leftrightarrow M_2$	bijektive Abbildung
F^{-1}	Umkehrabbildung einer bijektiven Abbildung
$F \circ G$	Hintereinanderschaltung, Komposition, Zusammensetzung von Abbildungen

$F^{-1}(N)$	$= \{x \in M : F(x) \in N\}$, Urbildmenge
$\mathrm{im}(F) = F(M)$	Bildmenge
$\mathrm{rg}(F)$	$= \dim \mathrm{im}(F)$, Rang einer linearen Abbildung
$\ker(F)$	Kern einer linearen Abbildung
$\mathrm{crg}(F)$	$= \dim \ker(F)$, Corang (Defekt) einer linearen Abbildung
$V_1 \approx V_2$	V_1 und V_2 sind isomorph.
$V_1 \cong V_2, V_1 = V_2$	V_1 und V_2 sind kanonisch isomorph.
$F : V_1 \xrightarrow{\sim} V_2$	F ist ein Isomorphismus.
A, B, C	Matrizen
$A^{m,n} = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$	$m \times n$-Matrix
$\mathbb{R}^{m\times n}$	Menge aller $m \times n$-Matrizen
$E = E^{n,n}; I$	Einheitsmatrix
δ_{ij}	Kroneckersymbol, Komponenten der Einheitsmatrix
$\mathcal{O} = \mathcal{O}^{m,n}$	Nullmatrix
A^t	zu A transponierte Matrix
$\mathrm{Rang}\,(A)$	Rang einer Matrix
$A \circ \vec{c}$	Produkt einer $m \times n$-Matrix und einem Spaltenvektor aus $\mathbb{R}^n$
$A + B$	Matrixaddition
$c \cdot A$	Skalarmultiplikation für Matrizen
$A \circ B$	Matrixprodukt
$GL(n, \mathbb{R})$	allgemeine lineare Gruppe vom Rang n über $\mathbb{R}$, Gruppe der invertierbaren $n \times n$-Matrizen
$SL(n, \mathbb{R})$	spezielle lineare Gruppe
$O(n, \mathbb{R})$	orthogonale Gruppe
$SO(n, \mathbb{R})$	spezielle orthogonale Gruppe
$M,\ N$	in Abschnitt 3.7: allgemeine Ebenen
codim	Codimension von allgemeinen Ebenen
$V_1 \oplus V_2$	direkte Summe der Vektorräume V_1 und V_2
$V_1 + V_2$	von V_1, V_2 aufgespannter Untervektorraum
$F_1 \oplus F_2$	direkte Summe der linearen Abbildungen F_1 und F_2
V^*	Dualraum von V

V^{**}	Bidualraum
$F^* : V_2^* \to V_1^*$	duale Abbildung zu $F : V_1 \to V_2$,
$\delta(i_1, \dots, i_n)$	allgemeines Kroneckersymbol
$\det(A), \det(F)$	Determinante
$A_{k\ell}$	Streichungsmatrix
$d_{k\ell}$	(k, ℓ)-ter Cofaktor
ad (A)	Adjunkte zu A
D_Θ, $\hat{D}_\Theta$	Drehungen der anschaulichen Ebene bzw. der Pfeile
$A_\alpha(\Theta)$	Drehungen in den Koordinaten bzgl. der Karte α
$\vec{x} \cdot \vec{y}$	kanonisches Skalarprodukt in $\mathbb{R}^n$
(x, y)	Skalarprodukt in der anschaulichen Ebene
$\Vert\ .\ \Vert$	Länge bzw. Norm eines Vektors
$\angle(y, x)$	Winkel zwischen $y, x \in V$
$x \perp y$	x, y sind orthogonal
$V^\perp$	orthogonales Komplement des Untervektorraums V
$\mathbb{C}$	Menge der komplexen Zahlen
$\mathbb{C}^n$	n-dimensionaler komplexer Zahlenraum
$\dim_{\mathbb{C}} V$	Dimension eines Vektorraums über $\mathbb{C}$
$\dim_{\mathbb{R}} V$	Dimension eines Vektorraums über $\mathbb{R}$
Re(z)	Realteil der komplexen Zahl z
Im(z)	Imaginärteil der komplexen Zahl z
$\bar{z}$	konjugiert komplexe Zahl
$(\vec{z}, \vec{w})$	kanonisches Skalarprodukt in $\mathbb{C}^n$
$\bar{A}^t$	konjugiert transponierte Matrix
$GL(n, \mathbb{C})$	allgemeine lineare Gruppe vom Rang n über $\mathbb{C}$
$U(n)$	unitäre Gruppe
Sp(A)	Spur der Matrix A
$P_A(\lambda), P_F(\lambda)$	charakteristisches Polynom
λ_ν	Eigenwerte
V_λ	Eigenraum
$\exp(A)$	Exponentialfunktion für Matrizen
$A^{1/2}$	Quadratwurzel positiv definiter Matrizen