

Skriptum zur Linearen Algebra und Analytischen Geometrie

Universität Kassel, Institut für Mathematik

Friedemann Kemm

Stand vom 25. Juli 2020

Inhaltsverzeichnis

1	Eigenwerte und Dreiecksmatrizen	7
2	Das Innenprodukt kommt ins Spiel	13
2.1	Das Innenprodukt: Wir messen Längen und Winkel	13
2.2	Orthonormalbasen: Endlich schöne Koordinatensysteme	22
2.3	Das orthogonale Komplement	29
2.4	Lineare Funktionale: Was lineare Abbildungen ins Eindimensionale mit Innenprodukten zu tun haben	34
2.5	Wann wir nur reelle Eigenwerte haben	36
2.6	Positive Operatoren und Isometrien und was man lustiges damit anstellen kann	48
2.7	Wie sich das Ganze auf Matrizen überträgt	60
2.7.1	Die Sache mit den orthogonalen und unitären Matrizen und weshalb das Standardskalarprodukt die Hauptrolle spielt	61
2.7.2	Positiv definite Matrizen	64
2.7.3	Polarzerlegung und Singulärwertzerlegung von Matrizen	69
3	Die Jordansche Normalform	73
3.1	Warum wir überhaupt sowas suchen	73
3.2	Vorüberlegungen zu wünschenswerten Matrixstrukturen	75
3.3	Die Sache mit den Haupträumen	79
3.4	. . . und wie wir uns daraus unsere Jordan-Normalform bauen	86
4	Mathematische Grundbegriffe und algebraische Strukturen	91
4.1	Warum wir noch mehr formalisieren wollen	91
4.2	Die mathematischen Grundbegriffe, die wir dafür brauchen	93
4.2.1	Mengen und Relationen	93
4.2.1.1	Allgemeines zu Mengen	93
4.2.1.2	Wenn die Grundmenge ins Spiel kommt	94
4.2.1.3	Kreuzprodukte und Relationen	95
4.2.2	Abbildungen	98
4.3	Was es mit den Verknüpfungen auf sich hat	101
4.3.1	Eine einzelne Verknüpfung	101
4.3.2	Zwei Verknüpfungen in der selben Menge	108
4.3.2.1	Wei bei den Zahlen: Addition und Multiplikation	109

4.3.2.2	Wie bei den Mengen: zwei duale Verknüpfungen	112
4.4	Wenn äußere Verknüpfungen ins Spiel kommen	114
4.5	Allgemeine Innenprodukträume	122
5	Analytische Geometrie: Wir binden die Vektoren an Punkten fest	125
5.1	Affine Räume: Wir bauen uns eine geeignete Mathematische Struktur	125
5.2	Geometrie im \mathbb{R}^n : Koordinatenbehaftete Geometrie	131
5.2.1	Die Ebene: \mathbb{R}^2	132
5.2.1.1	Geraden in der Ebene	132
5.2.1.2	Flächen von Parallelogrammen und Dreiecken	134
5.2.1.3	Wir werden vorsichtig nichtlinear: Quadriken in der Ebene	135
5.2.2	Der Raum: \mathbb{R}^3	139
5.2.2.1	Raum- und Flächeninhalte	139
5.2.2.2	Geraden und Ebenen im Raum	141
5.2.3	Ausblick auf höhere Dimensionen	143
5.3	Und wie das Ganze koordinatenfrei aussieht	143

Dieses Skriptum zur Vorlesung *Lineare Algebra und Analytische Geometrie* wurde zum Sommersemester 2020 an der Universität Kassel erstellt. Es knüpft an das Skriptum zur *Elementaren Linearen Algebra* aus dem Wintersemester 19/20.

Einen Schwerpunkt bilden Normalformen von Matrizen, insbesondere die Schursche Normalform, die Singulärwertzerlegung und die Jordan-Normalform. Da diese im Allgemeinen, d. h. bei großen Matrizen, nicht exakt berechnet werden können, müssen wir deren Approximation für die Vorlesung *Numerik II* (die ersteren beiden) bzw. weiter fortgeschrittene Veranstaltungen zur Numerischen Linearen Algebra (letztere) aufheben.

Bei der Jordan-Normalform gibt es allerdings einige Anwendungen, bei denen die Systemmatrix klein genug oder sonst einfach ist, sodass sie sich berechnen lässt. Hiervon werden Sie in den Vorlesungen zur Analysis gewöhnlicher Differentialgleichungen Gebrauch machen.

Im weiteren Verlauf werden wir dann allgemeine algebraische Strukturen einführen und in diesem Zusammenhang auch allgemeine Vektorräume. Danach werden wir affine Räume betrachten, die es uns endlich erlauben, mit dem bis dahin gelernten echte Geometrie zu betreiben.

Ein paar Lehrbücher zum Thema sind am Ende des Skripts angegeben. Dabei ist das Buch von Axler [Ax15] Hauptzeuge bei den Normalformen und Innenprodukträumen. Das Buch von Scheid und Schwarz [SS16] empfiehlt sich besonders für Lehramtskandidaten, denen die Behandlung von ganzen Zahlen in diesem Skript zu kurz kommt. Schließlich führt das Geometriebuch von Fischer [Fis92] die analytische (und sogar projektive) Geometrie deutlich weiter, als dies im Rahmen der aktuellen Vorlesung möglich wäre.

1 Eigenwerte und Dreiecksmatrizen

In der ELA hatten wir gesehen, dass wir Matrizen mit paarweise verschiedenen (komplexen) Eigenwerten diagonalisieren können. Bei mehrfachen Eigenwerten klappt das nicht immer. Das heißt, wir müssen unsere Ansprüche etwas zurückschrauben.

Zu Beginn werden wir zeigen, dass, zumindest im Komplexen bzw. wenn im reellen alle Eigenwerte selber reell sind, d. h. wenn das charakteristische Polynom in (reelle) Linearfaktoren zerfällt, sich jede quadratische Matrix mindestens *trigonalisieren* lässt. D. h. sie ist ähnlich zu einer oberen Dreiecksmatrix.

Je nachdem, ob wir zusätzliche Ansprüche an die Ähnlichkeitstransformation oder an die resultierende Dreiecksmatrix stellen, werden wir dann die *Schursche* oder die *Jordansche Normalform* erhalten. Das kommt dann aber in den späteren Kapiteln.

Zunächst wollen wir den schon in der ELA kurz benutzten Begriff des *Eigenraums* formal definieren:

Definition 1. Sei $A \in \mathbb{C}^{n \times n}$ und $\lambda \in \mathbb{C}$ ein Eigenwert von A . Dann heißt

$$V_\lambda := \{\mathbf{x} \in \mathbb{C}^n \mid A\mathbf{x} = \lambda\mathbf{x}\} = \ker(A - \lambda E)$$

der *Eigenraum* von A zum Eigenwert λ .

Bemerkung 2.

- Wie man leicht sieht, ist der Eigenraum immer ein Vektorraum.
- Wie wir uns zu Beginn von Kapitel 7 der ELA überlegt hatten, gilt für $\varphi_A : V \rightarrow V$, $\mathbf{x} \mapsto A\mathbf{x}$ und Eigenvektoren $\mathbf{x}_1, \dots, \mathbf{x}_k$ von A

$$\varphi_A(\langle \mathbf{x}_1, \dots, \mathbf{x}_k \rangle) \subseteq \langle \mathbf{x}_1, \dots, \mathbf{x}_k \rangle$$

und damit auch

$$\varphi_A(V_\lambda) \subseteq V_\lambda.$$

Der Eigenraum ist also *invariant* unter der von A induzierten linearen Abbildung.

- Ist $\lambda \neq 0$, dann gilt sogar

$$\varphi_A(V_\lambda) = V_\lambda.$$

- Ist A diagonalisierbar, gilt automatisch

$$\mathbb{C}^n = \bigoplus_{\lambda \text{ Eigenvektor von } A} V_\lambda.$$

Die direkte Summe aller Eigenräume ergibt also den Gesamttraum.

1 Eigenwerte und Dreiecksmatrizen

- Ist A nicht diagonalisierbar, klappt das nicht, sonst könnten wir ja aus den Basen der Eigenräume eine Basis des \mathbb{C}^n zusammenstellen, die aus n linear unabhängigen Eigenvektoren besteht. Damit könnten wir die Matrix dann doch diagonalisieren.
- Jeder unter φ_A invariante eindimensionale Unterraum wird übrigens von einem Eigenvektor aufgespannt. Da sieht man leicht, wenn man sich einen nichtverschwindenden Vektor daraus anschaut. Ihm bleibt nichts anderes übrig, als auf ein skalares Vielfaches seiner selbst abgebildet zu werden.

Beispiel 3. Für die Matrix

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

kennen wir die linear unabhängigen Eigenvektoren $r_1 = (1, 1)^T$ und $r_2 = (1, -1)^T$. Die zugehörigen Eigenwerte sind $\lambda_1 = 0 + 1 = 1$ und $\lambda_2 = 0 - 1 = -1$. Geometrisch werden also alle Vektoren an der Winkelhalbierenden der Koordinatenachsen gespiegelt. Da ein System l. u. Vektoren im \mathbb{C}^2 nicht mehr als zwei Vektoren umfassen kann, haben wir also eine Basis aus Eigenvektoren, und die Matrix ist diagonalisierbar. Es gilt

$$V_1 = \left\langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle, \quad V_{-1} = \left\langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\rangle, \quad \text{und} \quad \mathbb{C}^2 = V_1 + V_{-1}.$$

Beispiel 4. Bei der Matrix

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

klappt das nicht mehr. Das charakteristische Polynom ist $\chi(\lambda) = \lambda^2$, hat also nur die Null als doppelte Nullstelle. Aber für den zugehörigen Eigenraum gilt

$$V_0 = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle,$$

er hat also nur die Dimension Eins.

Interessant ist hier auch die Tatsache, dass nicht nur

$$V_0 = \ker(B - 0E) = \ker B = \langle (1, 0)^T \rangle,$$

sondern auch

$$\text{im } B = \langle (1, 0)^T \rangle$$

weshalb

$$\ker B + \text{im } B = \langle (1, 0)^T \rangle \neq \mathbb{C}^2$$

Würden wir statt B das Quadrat $(B - 0E)^2$ betrachten, so hätten wir wegen $B^2 = 0$

$$\ker(B - 0E)^2 = \mathbb{C}^2, \quad \text{im}(B - 0E)^2 = \{0\}$$

$$\text{und damit} \quad \mathbb{C}^2 = \ker(B - 0E)^2 \oplus \text{im}(B - 0E)^2.$$

Diesem Effekt wollen wir später allgemeiner nachgehen.

Beispiel 5. Bei der Matrix

$$\mathbf{B} = \begin{pmatrix} \alpha & 1 \\ 0 & \alpha \end{pmatrix}$$

mit einer reellen oder komplexen Zahl α sieht die Sache ähnlich aus. Das charakteristische Polynom ist $\chi(\lambda) = (\lambda - \alpha)^2$, hat also nur α als doppelte Nullstelle. Für den zugehörigen Eigenraum gilt wieder

$$V_0 = \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle.$$

Er hat also nur die Dimension Eins.

Auch hier ist $V_0 = \ker(\mathbf{B} - \alpha\mathbf{E}) = \langle (1, 0)^T \rangle$ und $\text{im}(\mathbf{B} - \alpha\mathbf{E}) = \langle (1, 0)^T \rangle$, weshalb $\ker(\mathbf{B} - \alpha\mathbf{E}) + \text{im}(\mathbf{B} - \alpha\mathbf{E}) \neq \mathbb{C}^2$. Würden wir statt $(\mathbf{B} - \alpha\mathbf{E})$ das Quadrat $(\mathbf{B} - \alpha\mathbf{E})^2$ betrachten, so hätten wir wieder

$$\ker(\mathbf{B} - \alpha\mathbf{E})^2 = \mathbb{C}^2, \quad \text{im}(\mathbf{B} - \alpha\mathbf{E})^2 = \{\mathbf{0}\}$$

und damit $\mathbb{C}^2 = \ker(\mathbf{B} - \alpha\mathbf{E})^2 \oplus \text{im}(\mathbf{B} - \alpha\mathbf{E})^2$.

Haben wir einen doppelten Eigenwert, so haben wir, wenn wir die Matrix, etwa \mathbf{A} in ihr charakteristisches Polynom einsetzen, einen Faktor $(\mathbf{A} - \lambda\mathbf{E})^2$, bei dreifachen Nullstellen entsprechend $(\mathbf{A} - \lambda\mathbf{E})^3$ u. s. w. diese wollen wir uns noch etwas genauer ansehen. Doch zuvor definieren wir die *invarianten Unterräume* formal:

Definition 6. Ist $\varphi \in \text{Hom}(\mathbb{C}^n, \mathbb{C}^n)$ (bzw. $\text{Hom}(\mathbb{R}^n, \mathbb{R}^n)$) und V ein Untervektorraum von \mathbb{C}^n (bzw. \mathbb{R}^n) mit

$$\varphi(V) \subseteq V,$$

dann heißt V *invariant* unter dem Endomorphismus φ .

Beispiel 7.

- Wie schon erwähnt, sind die Eigenräume einer quadratischen Matrix \mathbf{A} bzw. des von ihr induzierten Endomorphismus $\varphi_{\mathbf{A}}$ invariant unter diesem.
- Es gibt auch noch einfachere Beispiele, nämlich den Nullraum, den vollen Raum (beides unabhängig von φ), den Kern und das Bild. Bei bijektiven Endomorphismen sind das zwei, ansonsten vier Beispiele.

Ein weiteres nicht ganz so offensichtliches Beispiel ist das Bild von $\mathbf{A} - \lambda\mathbf{E}$ bzw. von $\varphi_{\mathbf{A}} - \lambda \text{id}$:

Lemma 8. Ist φ ein Endomorphismus, so ist für jeden Skalar λ das Bild $\text{im}(\varphi - \lambda \text{id})$ invariant unter φ .

Beweis. Setze $U = \text{im}(\varphi - \lambda \text{id})$ und wähle ein $\mathbf{u} \in U$. Dann ist

$$\varphi(\mathbf{u}) = \varphi(\mathbf{u}) - \lambda\mathbf{u} + \lambda\mathbf{u} = (\varphi - \lambda \text{id})(\mathbf{u}) + \lambda\mathbf{u}.$$

Aufgrund der Definition von U liegt der erste Summand in U , der zweite, weil U ein Vektorraum ist. Damit gilt das aber auch für die Summe, und wir sind fertig. \square

1 Eigenwerte und Dreiecksmatrizen

Zu beachten ist, dass wir λ nicht als Eigenwert vorausgesetzt haben. Ist es kein Eigenwert, dann ist $\varphi - \lambda \text{id}$ bijektiv und das Bild der ganze Raum. Für einen Eigenwert gilt das dann nicht mehr.

An dieser Stelle wollen wir noch einmal zusammenfassen, was wir bisher über die Diagonalisierbarkeit wissen bzw. zeigen können:

Satz 9. Sei $V \subseteq \mathbb{C}^N$ (bzw. \mathbb{R}^N) ein n -dimensionaler Vektorraum und φ ein Endomorphismus auf V mit den paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$. Dann sind die folgenden Bedingungen äquivalent:

(a) φ ist diagonalisierbar, d. h. durch eine Diagonalmatrix darstellbar.

(b) V besitzt eine Basis aus Eigenvektoren von φ .

(c) V besitzt eindimensionale invariante Unterräume U_1, \dots, U_n mit

$$V = \bigoplus_{j=1}^n U_j .$$

(d) V ist die direkte Summe der Eigenräume von φ , also

$$V = \bigoplus_{\ell=1}^k V_{\lambda_\ell} .$$

(e) Es gilt

$$\sum_{\ell=1}^k \dim V_{\lambda_\ell} = n .$$

Beweis. Dass (a) und (b) äquivalent sind, wissen wir schon aus ELA. Aus (b) folgt aber bereits (c) und mit dem letzten Punkt aus Bemerkung 2 auch die Umkehrung. Schließlich wird dann jeder Summand der direkten Summe von einem Eigenvektor aufgespannt, womit wir wieder n linear unabhängige Eigenvektoren, also eine Basis bekommen.

Dass aus (a) bereits (d) folgt hatten wir uns auch schon in Bemerkung 2 überlegt. Daraus folgt aber automatisch (e).

Setze nun (e) voraus und wähle in jedem Eigenraum eine Basis. Damit habe ich n Eigenvektoren. Da Eigenräume zu verschiedenen Eigenwerten immer nur trivialen Schnitt haben können, müssen diese wieder linear unabhängig sein, weshalb (b) gilt. \square

Bevor wir weitermachen, erst noch ein paar Kriterien für die Darstellbarkeit von Endomorphismen durch obere Dreiecksmatrizen. Der Schlüssel dazu ist die Betrachtung invarianter Unterräume:

Lemma 10. Sei V ein Vektorraum mit der Basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ und φ ein Endomorphismus auf V . Dann sind folgende Bedingungen äquivalent:

(a) Die Darstellungsmatrix von φ bezüglich der Basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ hat obere Dreiecksgestalt.

(b) Für jedes j gilt $\varphi(\mathbf{v}_j) \in \langle \mathbf{v}_1, \dots, \mathbf{v}_j \rangle$.

(c) Für jedes j ist $\langle \mathbf{v}_1, \dots, \mathbf{v}_j \rangle$ invariant unter φ .

Beweis. Die Äquivalenz von (a) und (b) kann man sich leicht überlegen (Übung). Dass aus (c) bereits (b) folgt ist noch offensichtlicher. Es bleibt also zu zeigen, dass umgekehrt aus (b) wieder (c) folgt.

Wir nehmen also an, dass (b) gilt und halten ein $j \in \{1, \dots, n\}$ fest. Unter Anwendung von (b) finden wir

$$\begin{aligned}\varphi(\mathbf{v}_1) &\in \langle \mathbf{v}_1 \rangle \subseteq \langle \mathbf{v}_1, \dots, \mathbf{v}_j \rangle, \\ \varphi(\mathbf{v}_2) &\in \langle \mathbf{v}_1, \mathbf{v}_2 \rangle \subseteq \langle \mathbf{v}_1, \dots, \mathbf{v}_j \rangle, \\ &\vdots \\ \varphi(\mathbf{v}_j) &\in \langle \mathbf{v}_1, \dots, \mathbf{v}_j \rangle.\end{aligned}$$

Ist nun $\mathbf{v} \in \langle \mathbf{v}_1, \dots, \mathbf{v}_j \rangle$ Dann ist es eine Linearkombination der $\mathbf{v}_1, \dots, \mathbf{v}_j$ und damit auch $\varphi(\mathbf{v})$. Das wollten wir aber gerade zeigen. \square

Diese Beobachtungen sind nun der Schlüssel zu unserem gewünschten Darstellungssatz:

Satz 11. Sei $V \subseteq \mathbb{C}^N$ ein n -dimensionaler Vektorraum und φ ein Endomorphismus auf V . Dann gibt es eine Basis von V bezüglich der die Darstellungsmatrix von φ obere Dreiecksgestalt hat.

Beweis. Da wir im Komplexen sind, wissen wir, dass das charakteristische Polynom in Linearfaktoren zerfällt, wir also mindestens einen Eigenwert λ haben. Wir machen eine Induktion über die Dimension n von V :

- Für $n = 1$ sind alle Matrizen obere Dreiecksmatrizen. Damit haben wir einen Induktionsanfang.
- Für $n > 1$ wissen wir, dass mit $U = \text{im}(\varphi - \lambda \text{id})$ die Dimension von U kleiner ist als die von V . Das hatten wir uns ja gerade oben überlegt. Außerdem wissen wir aus Lemma 8, dass $\varphi|_U$ ein Endomorphismus auf U ist. Damit gilt für $\varphi|_U$ bereits die Behauptung. Wähle also in U eine Basis $\mathbf{u}_1, \dots, \mathbf{u}_m$ bezüglich der $\varphi|_U$ durch eine obere Dreiecksmatrix dargestellt wird. Damit ist nach Lemma 10 für jedes j

$$\varphi(\mathbf{u}_j) = (\varphi|_U)(\mathbf{u}_j) \in \langle \mathbf{u}_1, \dots, \mathbf{u}_j \rangle. \quad (1.1)$$

Ergänze nun $\mathbf{u}_1, \dots, \mathbf{u}_m$ durch $\mathbf{v}_1, \dots, \mathbf{v}_{n-m}$ zu einer Basis von V . Für jedes passende k wenden wir nun wieder den Trick aus dem Beweis von Lemma 8 an:

$$\varphi(\mathbf{v}_k) = (\varphi - \lambda \text{id})(\mathbf{v}_k) + \lambda \mathbf{v}_k.$$

1 Eigenwerte und Dreiecksmatrizen

Da der erste Summand nach Definition in U liegt, ist also

$$\varphi(\mathbf{v}_k) \in \langle \mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{v}_k \rangle \subseteq \langle \mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{v}_1, \dots, \mathbf{v}_k \rangle. \quad (1.2)$$

Nehmen wir nun (1.1) und (1.2) zusammen und ziehen Lemma 10 hinzu, so folgt, dass φ bezüglich der Basis

$$\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{v}_1, \dots, \mathbf{v}_{n-m}$$

wie gewünscht durch eine obere Dreiecksmatrix dargestellt wird.

□

Bemerkung 12.

- In Anlehnung an die Diagonalisierung einer Matrix spricht man hier auch von einer *Trigonalisierung*.
- Dementsprechend lässt sich nach Satz 11 im Komplexen jede quadratische Matrix *trigonalisieren*.
- Dies funktioniert im Reellen auch, sofern sich das charakteristische Polynom der Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ als Produkt von Linearfaktoren schreiben lässt, also

$$\chi_{\mathbf{A}}(\lambda) = \prod_{k=0}^n (\lambda - \lambda_k),$$

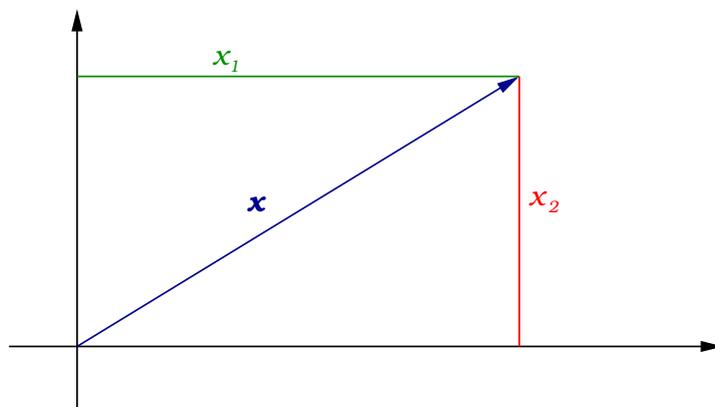
wobei die λ_k nicht notwendig paarweise verschieden sein müssen.

2 Das Innenprodukt kommt ins Spiel

Um unseren Zielen, Geometrie zu treiben und Lineare Abbildungen/Operatoren bezüglich schöner Basen darstellen zu können, näher zu kommen, müssen wir uns überlegen wie wir einerseits die Länge von Vektoren messen und andererseits den Winkel zwischen Vektoren bestimmen können. Damit beides zusammenpasst, wollen wir eine Abbildung von $\mathbb{R}^n \times \mathbb{R}^n$ in die reellen bzw. von $\mathbb{C}^n \times \mathbb{C}^n$ in die komplexen Zahlen suchen, aus der sich dann beides bauen lässt. Wir nennen diese dann *Innenprodukt*.

2.1 Das Innenprodukt: Wir messen Längen und Winkel

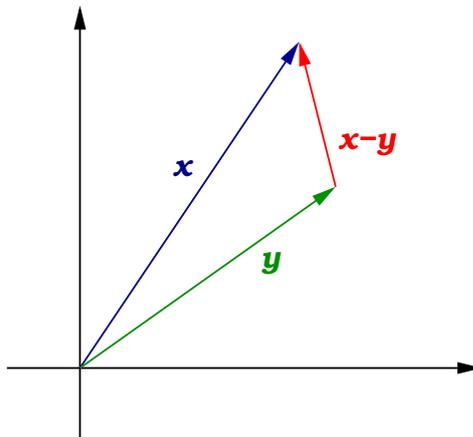
Würden wir den Vektor $\mathbf{x} = (x_1, x_2)$ in einem ebenen kartesischen Koordinatensystem einzeichnen, so wäre nach Pythagoras klar, wie wir seine Länge bestimmen:



$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2} = \sqrt{\mathbf{x}^T \mathbf{x}} .$$

Verwenden wir das und schauen uns zwei Vektoren $\mathbf{x} = (x_1, x_2)$ und $\mathbf{y} = (y_1, y_2)$ an, die wir im selben Punkt ansetzen.

2 Das Innenprodukt kommt ins Spiel



Dann werden die Spitzen der beiden gerade durch den Vektor $\mathbf{x} - \mathbf{y}$ verbunden. Berechnen wir dessen Länge mit obiger Formel, so finden wir für deren Quadrat

$$\begin{aligned} \|\mathbf{x} - \mathbf{y}\|^2 &= (x_1 - y_1)^2 + (x_2 - y_2)^2 \\ &= x_1^2 - 2x_1y_1 + y_1^2 + x_2^2 - 2x_2y_2 + y_2^2 \\ &= (x_1^2 + x_2^2) + (y_1^2 + y_2^2) - 2(x_1y_1 + x_2y_2) \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\mathbf{x}^T \mathbf{y} . \end{aligned}$$

Nach dem Kosinussatz aus der ebenen Geometrie muss aber gelten

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\| \cos(\sphericalangle(\mathbf{x}, \mathbf{y})) .$$

Hieraus schließen wir messerscharf

$$\cos(\sphericalangle(\mathbf{x}, \mathbf{y})) = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} . \quad (2.1)$$

Das wollen wir zum Anlass nehmen, nicht nur die gefundenen Formeln auf beliebige Dimensionen – und ins Komplexe – zu verallgemeinern, sondern auch, statt mit den Winkeln selber mit deren Kosinus zu arbeiten. Das so entstandene *Standardskalarprodukt* $\mathbf{x} \cdot \mathbf{y} := \mathbf{x}^T \mathbf{y}$ wollen wir nun als „role model“ für allgemeine Innenprodukte nehmen.

Das erste Hindernis, das uns dabei begegnet, ist, dass sich im Komplexen und dort schon in \mathbb{C}^1 aus z^2 im allgemeinen keine Wurzel ziehen lässt. So ist für $z = x + iy$ das Quadrat $z^2 = x^2 - y^2 + 2ixy$, also nur in Ausnahmefällen eine nichtnegative reelle Zahl. Was wir bräuchten wären Terme der Gestalt $|z|^2$. Das heißt wir brauchen statt z^2 eher $z\bar{z} = \bar{z}z$. Wir müssen uns also für unser Standardskalarprodukt zwischen $\mathbf{z}^T \bar{\mathbf{w}}$ und $\overline{\mathbf{z}^T \mathbf{w}} = \bar{\mathbf{z}}^T \mathbf{w} =: \mathbf{z}^H \mathbf{w}$ entscheiden. Dabei liest man den oberen Index H als „hermitesch“ (nach dem französischen Mathematiker Charles Hermite (1822–1901)). *In der Literatur sind beide Varianten im Umlauf. Wir richten uns in diesem Kurs nach der Mehrheit und wählen die erstere.*

Damit können wir die wichtigsten Eigenschaften des Standardskalarprodukts zusammentragen und als Definition allgemeiner *Innenprodukte* verwenden:

2.1 Das Innenprodukt: Wir messen Längen und Winkel

Definition 13. Ist V ein reeller oder komplexer Vektorraum, so nennen wir eine Abbildung

$$(\cdot, \cdot): V \times V \rightarrow \mathbb{R} \quad \text{bzw.} \quad \mathbb{C}$$

ein *Innenprodukt* auf V , falls die folgenden Eigenschaften erfüllt sind:

Positivität:

$$(\mathbf{v}, \mathbf{v}) \geq 0 \quad \forall \mathbf{v} \in V.$$

Definitheit:

$$(\mathbf{v}, \mathbf{v}) = 0 \quad \Leftrightarrow \quad \mathbf{v} = \mathbf{0}.$$

Linearität im ersten Argument:

$$(\lambda \mathbf{v} + \mathbf{w}, \mathbf{u}) = \lambda(\mathbf{v}, \mathbf{u}) + (\mathbf{w}, \mathbf{u}) \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V, \quad \lambda \in \mathbb{C} \text{ (bzw. } \mathbb{R}\text{)}.$$

(Konjugierte) Symmetrie:

$$(\mathbf{u}, \mathbf{v}) = \overline{(\mathbf{v}, \mathbf{u})} \quad \forall \mathbf{u}, \mathbf{v} \in V.$$

Einen Vektorraum mit einem Innenprodukt nennen wir dann *Innenproduktraum*.

Wenn nichts anderes angegeben ist, werden wir im Rest dieses Kapitels immer von einem Innenproduktraum ausgehen.

Bemerkung 14.

- Bei der letzten Eigenschaft steht das „Konjugierte“ in Klammern, weil wir es im Reellen nicht brauchen. Da ist dann das Innenprodukt auch automatisch in beiden Argumenten linear, also eine sogenannte *Bilinearform*. (So wie die Determinante eine *Multilinearform* ist.) Allerdings setzt man bei Bilinearformen die ersten beiden Eigenschaften nicht immer voraus. Sind sie erfüllt, so nennt man die Bilinearform *positiv definit*, gilt nur die erste aber nicht die zweite einfach *positiv semidefinit*.
- Oft werden statt der runden Klammern spitze Klammern verwendet, also $\langle \cdot, \cdot \rangle$. Aber die haben wir in unserem Kurs schon für die lineare Hülle verballert.

Beispiel 15. Wie schon angedeutet, ist das Standardskalarprodukt $(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \bar{\mathbf{y}}$ ein Innenprodukt.

Beispiel 16. Sind $c_1, \dots, c_n > 0$, so ist auf \mathbb{R} bzw. \mathbb{C} ein Innenprodukt gegeben durch

$$(\mathbf{x}, \mathbf{y}) := \sum_{j=1}^n c_j x_j \bar{y}_j = \mathbf{x}^T \text{diag}(c_1, \dots, c_n) \bar{\mathbf{y}}.$$

Das heißt, dass wir zwischen die beiden Vektoren eine Diagonalmatrix mit positiven Diagonaleinträgen geschaltet haben. Dieses Beispiel werden wir in Abschnitt 2.5 weiter verallgemeinern.

2 Das Innenprodukt kommt ins Spiel

Weitere Beispiele gibt es in der Übung.

Zunächst tragen wir einige grundlegende Eigenschaften von Innenprodukten zusammen:

Satz 17. In einem Innenproduktraum V gilt:

1. $(\mathbf{0}, \mathbf{u}) = (\mathbf{u}, \mathbf{0}) = 0$ für jedes $\mathbf{u} \in V$.
2. Das Innenprodukt ist auch im zweiten Argument additiv, d. h.

$$(\mathbf{u}, \mathbf{v} + \mathbf{w}) = (\mathbf{u}, \mathbf{v}) + (\mathbf{u}, \mathbf{w}) \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V.$$

3. $(\mathbf{u}, \lambda \mathbf{v}) = \bar{\lambda}(\mathbf{u}, \mathbf{v})$ für jeden Skalar λ und beliebige Vektoren $\mathbf{u}, \mathbf{v} \in V$.

Beweis. Übung. □

Nun könnten wir aus unserem Innenprodukt mit $\|\mathbf{x}\| := \sqrt{(\mathbf{x}, \mathbf{x})}$ zum Messen von Längen eine sogenannte *Norm* bauen und uns deren Eigenschaften anschauen. Wir wollen hier aber lieber umgekehrt erst mal definieren, was wir unter einer Norm verstehen wollen, und dann zeigen, dass mit $\|\mathbf{x}\| := \sqrt{(\mathbf{x}, \mathbf{x})}$ bereits eine solche gegeben ist. Wir schreiben uns also alle wünschenswerten Eigenschaften des Längenmaßes von Vektoren auf:

Definition 18. Sei V ein Vektorraum. Eine Abbildung $\|\cdot\| : V \rightarrow [0, \infty)$ heißt *Norm auf V* , falls sie folgende Eigenschaften hat:

Definitheit: $\|\mathbf{x}\| = 0 \Rightarrow \mathbf{x} = \mathbf{0}$.

Absolute Homogenität: $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$ (absolut, weil wir den Absolutbetrag rausziehen)

Subadditivität: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$. Aus geometrischer Sicht spricht man hier von der *Dreiecksungleichung*. (Siehe Abbildung 2.1.)

Ein Vektorraum mit einer Norm heißt *normierter (Vektor-)Raum*.

Bemerkung 19. Mit $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ lässt sich auf jedem normierten Vektorraum ein Abstands- bzw. Distanzbegriff, also eine sogenannte *Metrik* definieren, was Sie in der Analysis fröhlich benutzen (werden).

Beispiel 20. Unsere oben definierte Länge $\|\mathbf{x}\| := \sqrt{(\mathbf{x}, \mathbf{x})}$ ist eine Norm:

- Sie ist per Definition tatsächlich eine Abbildung von V ins halboffene Intervall $[0, \infty)$.
- $\|\mathbf{x}\| = 0 \Leftrightarrow \|\mathbf{x}\|^2 = (\mathbf{x}, \mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$. D. h. sie erbt die Definitheit direkt vom Innenprodukt.
- $\|\lambda \mathbf{x}\|^2 = (\lambda \mathbf{x}, \lambda \mathbf{x}) = \lambda(\mathbf{x}, \lambda \mathbf{x}) = \lambda \bar{\lambda}(\mathbf{x}, \mathbf{x}) = |\lambda|^2 \|\mathbf{x}\|^2$. Durch Ziehen der Wurzel erhalten wir die absolute Homogenität.

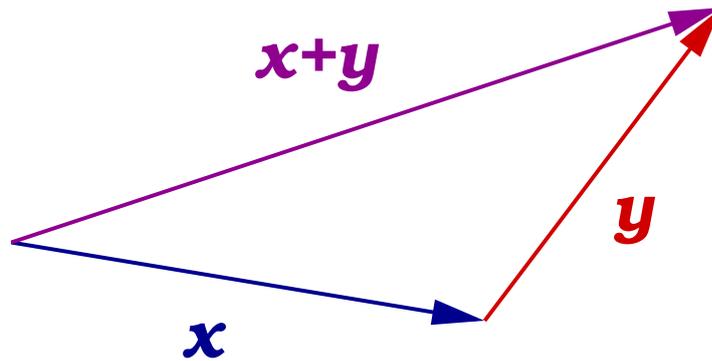


Abbildung 2.1: Veranschaulichung der Dreiecksungleichung

- Wir betrachten wieder das Quadrat, also für die Dreiecksungleichung und verwenden die Tatsache aus Satz 17, dass das Innenprodukt auch im zweiten Argument additiv ist:

$$\begin{aligned}
 \|x + y\|^2 &= (x + y, x + y) \\
 &= (x, x + y) + (y, x + y) \\
 &= (x, x) + (x, y) + (y, x) + (y, y) \\
 &= \|x\|^2 + \|y\|^2 + (x, y) + \overline{(x, y)} \\
 &= \|x\|^2 + \|y\|^2 + 2\Re((x, y)) \\
 &\leq \|x\|^2 + \|y\|^2 + 2|(x, y)|,
 \end{aligned}$$

wobei $\Re(z)$ den *Realteil* einer komplexen Zahl bezeichnet ($\Im(z)$ entsprechend den *Imaginärteil*). Andererseits ist

$$(\|x\| + \|y\|)^2 = \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|.$$

Falls es uns gelingt zu zeigen, dass

$$|(x, y)| \leq \|x\|\|y\|,$$

haben wir es geschafft. Diese Lücke werden wir durch die *Cauchy-Schwarzsche Ungleichung (C. S. U.)* (27) schließen. Aber dafür lassen wir uns noch etwas Zeit, um auf dem Weg dahin noch ein paar „Blumen zu pflücken“.

Beispiel 21. Eine in der Analysis und Numerik häufig verwendete Norm ist die *Betragsnorm*:

$$\|x\|_1 := \sum_{j=1}^n |x_j|. \quad (2.2)$$

Die Positivität und Definitheit sind offensichtlich. Die absolute Homogenität folgt direkt aus den Eigenschaften des Betrags, genauer aus deren Komponentenweiser Anwendung.

2 Das Innenprodukt kommt ins Spiel

Das gilt auch für die Subadditivität:

$$\begin{aligned}\|\mathbf{x} + \mathbf{y}\|_1 &= \sum_{j=1}^n |x_j + y_j| \\ &\leq \sum_{j=1}^n |x_j| + |y_j| \\ &= \|\mathbf{x}\| + \|\mathbf{y}\|_1.\end{aligned}$$

Sie ist also leicht zu verstehen, hat darüber hinaus eine leicht zu berechnende zugehörige *Operatornorm* (zu denen kommen wir später), lässt sich aber nicht durch ein Innenprodukt definieren.

Beispiel 22. Die *Maximumsnorm*

$$\|\mathbf{x}\|_\infty := \max_{j=1,\dots,n} |x_j| \quad (2.3)$$

sieht zwar anders aus als die Betragssummennorm, ist aber ansonsten ziemlich ähnlich gelagert.

Beispiel 23. Mit etwas Analysis kann man ganz allgemein Normen der Form

$$\|\mathbf{x}\|_p = \left(\sum_{j=1}^n |x_j|^p \right)^{\frac{1}{p}} \quad (2.4)$$

mit einem $p \in [1, \infty)$ definieren. Für $p = 1$ ergibt sich gerade wieder die Betragssummennorm, für $p = 2$ die vom Standardskalarprodukt induzierte Norm, die sogenannte *Euklid-Norm*. Der Name kommt daher, dass wir mit ihr gerade wieder den Längenbegriff aus der Euklideschen Geometrie erhalten.

Zuerst verallgemeinern wir den Winkel bzw. dessen Kosinus und die Orthogonalität aus unserer gewohnten Geometrie auf allgemeine Innenprodukträume. In Formel (2.1) sehen wir, dass bei einem rechten Winkel das Standardskalarprodukt verschwindet. Also definieren wir

Definition 24. In einem Innenproduktraum nennen wir zwei Vektoren \mathbf{x} und \mathbf{y} *orthogonal* (bzgl. des Innenprodukts), wenn $(\mathbf{x}, \mathbf{y}) = 0$. Wir verwenden das übliche Symbol hierfür: $\mathbf{x} \perp \mathbf{y}$.

Allgemein definieren wir für den Winkel zwischen zwei Vektoren den *Kosinus* (bzgl. des Innenprodukts) als

$$\cos(\sphericalangle(\mathbf{x}, \mathbf{y})) := \frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\| \|\mathbf{y}\|}. \quad (2.5)$$

Damit wäre der Beweis der ersehnten Cauchy-Schwarzschen Ungleichung auch gleichzeitig der Beweis, dass der Betrag des Kosinus auch in allgemeinen Innenprodukträumen niemals größer als Eins werden kann. Doch zunächst:

2.1 Das Innenprodukt: Wir messen Längen und Winkel

Bemerkung 25. Der Nullvektor steht automatisch auf allen anderen senkrecht. Er ist auch der einzige, der auf sich selbst senkrecht steht. (Dabei haben wir hier *senkrecht* als Synonym für orthogonal verwendet.)

In Innenprodukträumen bekommt der Satz des Pythagoras die folgende Gestalt:

Satz 26 (Satz des Pythagoras). *Ist V ein Innenproduktraum und $\mathbf{u}, \mathbf{v} \in V$ mit $\mathbf{u} \perp \mathbf{v}$. Dann gilt*

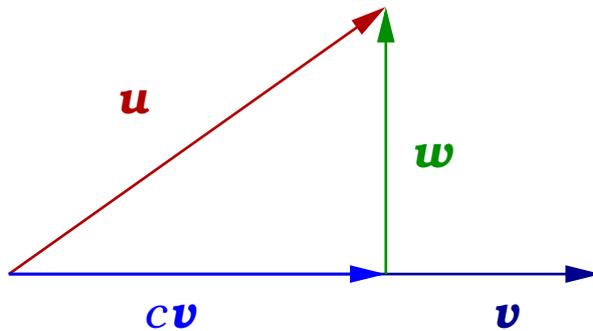
$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 .$$

Beweis. Den Beweis haben wir eigentlich schon fertig: Mit der Rechnung aus Beispiel 20 haben wir

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\Re((\mathbf{u}, \mathbf{v})) .$$

Da aber die Vektoren orthogonal sind, verschwindet der letzte Term. □

Unser nächstes Ziel ist es, von einem Vektor, etwa \mathbf{u} das Lot auf einen anderen Vektor, etwa \mathbf{v} zu fällen. Dazu müssen wir wieder beide gedanklich im selben Punkt ansetzen:



Dementsprechend suchen wir also ein $\mathbf{w} \in V$ und einen Skalar c , sodass

$$\begin{aligned} (\mathbf{w}, \mathbf{v}) &= 0, \\ c\mathbf{v} + \mathbf{w} &= \mathbf{u} . \end{aligned}$$

Lösen wir die zweite Beziehung nach \mathbf{w} auf und setzen sie in die erste ein, so erhalten wir

$$0 = (\mathbf{u} - c\mathbf{v}, \mathbf{v}) = (\mathbf{u}, \mathbf{v}) - c\|\mathbf{v}\|^2 .$$

Das klappt aber gerade für $c = (\mathbf{u}, \mathbf{v})/\|\mathbf{v}\|^2$. Damit ist

$$\mathbf{u} = \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{v}\|^2} \mathbf{v} + \left(\mathbf{u} - \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{v}\|^2} \mathbf{v} \right) \quad (2.6)$$

eine *Orthogonalzerlegung* von \mathbf{u} . Dabei ist der erste Summand die *Orthogonalprojektion* von \mathbf{u} auf die Richtung von \mathbf{v} und der zweite Summand ist das *Lot* von \mathbf{u} auf \mathbf{v} .

Damit können wir aber auch die Cauchy-Schwarzsche Ungleichung zeigen:

2 Das Innenprodukt kommt ins Spiel

Satz 27 (Cauchy-Schwarzsche Ungleichung). Ist V ein Innenproduktraum und $\mathbf{u}, \mathbf{v} \in V$, dann gilt

$$|(\mathbf{u}, \mathbf{v})| \leq \|\mathbf{u}\| \|\mathbf{v}\|. \quad (2.7)$$

Gleichheit gilt genau dann, wenn die beiden Vektoren linear abhängig sind.

Beweis. Ist einer der beiden Vektoren der Nullvektor, so ist die Gleichung trivialerweise erfüllt. Ansonsten betrachten wir entsprechend Gleichung (2.6) die Orthogonalzerlegung

$$\mathbf{u} = \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{v}\|^2} \mathbf{v} + \mathbf{w}$$

und wenden darauf den Satz des Pythagoras 26 an:

$$\begin{aligned} \|\mathbf{u}\|^2 &= \left\| \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{v}\|^2} \mathbf{v} \right\|^2 + \|\mathbf{w}\|^2 \\ &= \frac{|(\mathbf{u}, \mathbf{v})| \|\mathbf{v}\|^2}{\|\mathbf{v}\|^4} + \|\mathbf{w}\|^2 \\ &= \frac{|(\mathbf{u}, \mathbf{v})|}{\|\mathbf{v}\|^2} + \|\mathbf{w}\|^2 \\ &\geq \frac{|(\mathbf{u}, \mathbf{v})|}{\|\mathbf{v}\|^2} \end{aligned}$$

Nun müssen wir nur noch mit $\|\mathbf{v}\|^2$ durchmultiplizieren und erhalten die gewünschte Ungleichung.

Gleichheit gilt offenbar genau dann, wenn $\mathbf{w} = \mathbf{0}$, also genau dann, wenn entsprechend Gleichung (2.6) gilt

$$\mathbf{u} = \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{v}\|^2} \mathbf{v}.$$

Das ist aber gerade unsere Behauptung. □

Beispiel 28. Angewandt auf das Standardskalarprodukt im \mathbb{R}^n liefert die C. S. U.

$$\left| \sum_{j=1}^n x_j y_j \right|^2 \leq \left(\sum_{j=1}^n x_j^2 \right) \cdot \left(\sum_{j=1}^n y_j^2 \right).$$

Nun können wir aber auch unsere Lücke bei der Dreiecksungleichung schließen:

Lemma 29. Die auf einem Innenproduktraum durch $\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$ gegebene Funktion erfüllt die Dreiecksungleichung und ist damit endgültig als Norm nachgewiesen.

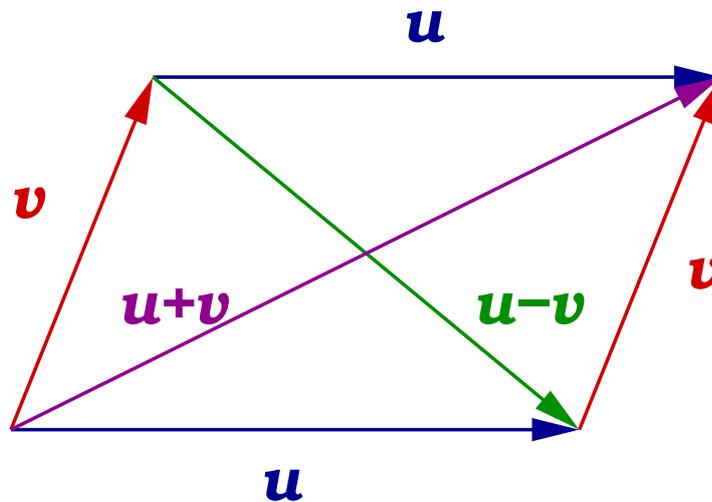
Beweis. Wir mussten nur noch zeigen, dass

$$|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\| \|\mathbf{y}\|.$$

Das ist aber gerade die Cauchy-Schwarzsche Ungleichung, womit wir fertig sind. Den Rest haben wir ja schließlich in Beispiel 20 schon erledigt. □

2.1 Das Innenprodukt: Wir messen Längen und Winkel

Im folgenden Bild sehen wir, wie die beiden Diagonalen eines Parallelogramms dieses jeweils in zwei Dreiecke teilen.



Ganz offensichtlich gilt im allgemeinen für keines der beiden Dreiecke der Satz des Pythagoras, weil beide keinen rechten Winkel haben. Addieren wir aber für beide Dreiecke, so stimmt es wieder:

Satz 30 (Parallelogrammgleichung). Ist V ein Innenproduktraum und $u, v \in V$, dann gilt

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2). \quad (2.8)$$

Beweis. Wir verwenden die Tatsache, dass unsere Norm aus einem Innenprodukt kommt:

$$\begin{aligned} \|u + v\|^2 + \|u - v\|^2 &= (u + v, u + v) + (u - v, u - v) \\ &= \|u\|^2 + \|v\|^2 + (u, v) + (v, u) \\ &\quad + \|u\|^2 + \|v\|^2 - (u, v) - (v, u) \\ &= 2(\|u\|^2 + \|v\|^2). \end{aligned}$$

□

Die *Parallelogrammgleichung* ist ein besonderes Charakteristikum von Innenprodukträumen schließlich gilt

Satz 31. Eine Norm wird genau dann von einem Innenprodukt induziert, wenn sie die *Parallelogrammgleichung* erfüllt.

Den Beweis lassen wir weg. Stattdessen geben wir passende Formeln an, mit denen man aus der Norm das Skalarprodukt zurückgewinnen kann.

Im Reellen: Da reicht

$$(u, v) = \frac{\|u + v\|^2 - \|u - v\|^2}{4}. \quad (2.9)$$

Im Komplexen: Da muss noch was dazu:

$$(\mathbf{u}, \mathbf{v}) = \frac{\|\mathbf{u} + \mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2}{4} + i \frac{\|i\mathbf{u} + \mathbf{v}\|^2 - \|i\mathbf{u} - \mathbf{v}\|^2}{4}. \quad (2.10)$$

Den Beweis, dass diese Formeln es tun, lassen wir als Übungsaufgabe. Man kann die Formeln natürlich noch weiter variieren. Als Inspirationsquelle dafür kann unser Beweis der Parallelogrammgleichung dienen.

2.2 Orthonormalbasen: Endlich schöne Koordinatensysteme

Definition 32. Ein Satz von Vektoren bzw. eine Menge von Vektoren in einem Innenproduktraum heißt *orthonormal*, wenn jeder der Vektoren Norm Eins hat und senkrecht/orthogonal auf allen anderen steht. D. h. die Vektoren $\mathbf{q}_1, \dots, \mathbf{q}_m$ sind orthonormal, falls immer

$$(\mathbf{q}_j, \mathbf{q}_k) = \delta_{jk} := \begin{cases} 1 & \text{falls } j = k, \\ 0 & \text{sonst,} \end{cases} \quad (2.11)$$

wobei δ_{jk} als *Kroneckersymbol* oder *Kroneckerdelta* bezeichnet wird.

Beispiel 33. Beispiele, die sowohl im Reellen als auch im Komplexen bezüglich des Standardskalarprodukts orthonormal sind:

1. Die Standardeinheitsbasis.

2.

$$\begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \quad \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}.$$

3.

$$\begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \quad \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{pmatrix}.$$

Bemerkenswert ist, dass für ein solches orthonormales System mit dem Satz des Pythagoras 26 gilt

$$\begin{aligned}
 \left\| \sum_{j=1}^m \alpha_j \mathbf{q}_j \right\|^2 &= \|\alpha_1 \mathbf{q}_1\|^2 + \left\| \sum_{j=2}^m \alpha_j \mathbf{q}_j \right\|^2 \\
 &= \|\alpha_1 \mathbf{q}_1\|^2 + \|\alpha_2 \mathbf{q}_2\|^2 + \left\| \sum_{j=3}^m \alpha_j \mathbf{q}_j \right\|^2 \\
 &= \dots \\
 &= \sum_{j=1}^m \|\alpha_j \mathbf{q}_j\|^2 \\
 &= \sum_{j=1}^m |\alpha_j|^2 \|\mathbf{q}_j\|^2 \\
 &= \sum_{j=1}^m |\alpha_j|^2,
 \end{aligned} \tag{2.12}$$

woraus wir messerscharf schließen:

Lemma 34. *Jedes orthonormale System ist linear unabhängig.*

Beweis. Ist $\sum_{j=1}^m \alpha_j \mathbf{q}_j = \mathbf{0}$, so muss nach Gleichung (2.12) gelten

$$\sum_{j=1}^m |\alpha_j|^2 = 0.$$

Das geht aber nur, wenn alle α_j verschwinden. □

Für den Fall, dass das System den ganzen Raum aufspannt, machen wir folgende Definition:

Definition 35. Ist V ein Innenproduktraum und $\mathbf{q}_1, \dots, \mathbf{q}_n$ ein Orthonormalsystem und gleichzeitig eine Basis von V , so nennen wir es eine *Orthonormalbasis (ONB)* von V .

Bemerkung 36. Entspricht die Zahl der Vektoren eines Orthonormalsystems gerade der Dimension des Raums, so ist es bereits eine ONB.

Beispiel 37. Die Vektoren

$$\begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}, \quad \begin{pmatrix} 1/2 \\ 1/2 \\ -1/2 \\ -1/2 \end{pmatrix}, \quad \begin{pmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{pmatrix}, \quad \begin{pmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{pmatrix}$$

2 Das Innenprodukt kommt ins Spiel

bilden eine Orthonormalbasis sowohl in \mathbb{R}^4 als auch in \mathbb{C}^4 bezüglich des Standardskalarprodukts. Hierzu müssen wir nur die Norm ausrechnen, die bei allen vieren gleich ist, weil die Komponenten betraglich gleich sind. Diese ist tatsächlich Eins. Dann müssen wir noch paarweise die Innenprodukte bilden und feststellen, dass diese tatsächlich verschwinden. Da es viere sind, müssen sie bereits eine ONB bilden.

Die nächste Frage, die sich stellt, ist die, ob wir es bei einer ONB leichter haben, einen Vektor als Linearkombination über dieser darzustellen:

$$\mathbf{u} = \sum_{j=1}^n \alpha_j \mathbf{q}_j$$

Bilden wir hiervon das Skalarprodukt, etwa mit \mathbf{q}_k aus unserer ONB, so finden wir

$$(\mathbf{u}, \mathbf{q}_k) = \sum_{j=1}^n \alpha_j (\mathbf{q}_j, \mathbf{q}_k) = \alpha_k . \quad (2.13)$$

Damit haben wir also eine Formel für die Koeffizienten. Mit Gleichung (2.12) liefert uns diese aber auch direkt eine Formel für die Norm. Wir fassen dies zusammen als

Lemma 38. *Ist V ein Innenproduktraum mit Orthonormalbasis $\mathbf{q}_1, \dots, \mathbf{q}_n$ und $\mathbf{u} \in V$, so gilt*

$$\mathbf{u} = \sum_{j=1}^n (\mathbf{u}, \mathbf{q}_j) \mathbf{q}_j \quad (2.14)$$

und

$$\|\mathbf{u}\|^2 = \sum_{j=1}^n |(\mathbf{u}, \mathbf{q}_j)|^2 . \quad (2.15)$$

Bemerkung 39. Gleichung (2.14) bzw. (2.13) kann dahingehend interpretiert werden, dass wir von \mathbf{u} das Lot auf jede Koordinatenrichtung fällen und damit entsprechend (2.6) die Orthogonalzerlegung, etwa bzgl. \mathbf{q}_j erhalten als

$$\mathbf{u} = (\mathbf{u}, \mathbf{q}_j) \mathbf{q}_j + (\mathbf{u} - (\mathbf{u}, \mathbf{q}_j) \mathbf{q}_j) .$$

Dabei lassen wir den hinteren Teil, der ja orthogonal auf \mathbf{q}_j steht, weg. Im Vergleich zur allgemeinen Formel (2.6) kann der Nenner wegbleiben, weil er ohnehin Eins ist, schließlich ist \mathbf{q}_j auf Eins normiert.

Nun wissen wir also, wie bequem es ist, mit Orthonormalbasen zu arbeiten. Es bleibt allerdings die Frage, wie wir zu einer solchen kommen. Im \mathbb{R}^n oder \mathbb{C}^n mit dem Standardskalarprodukt könnten wir einfach die Standardeinheitsvektoren nehmen. Aber erstens haben wir es nicht immer mit dem ganzen \mathbb{R}^n oder \mathbb{C}^n zu tun, zweitens nicht immer mit dem Standardskalarprodukt und drittens könnte es sein, dass wir eine ONB suchen, die besser zum zu lösenden Problem passt. Da brauchen wir also eine Strategie, wie wir aus

einem gegebenen (vorzugsweise endlichen) Erzeugendensystem für unseren Innenproduktraum eine ONB basteln können. Hier bietet es sich an, die Strategie von Gram und Schmidt zu verwenden, das sogenannte *Gram-Schmidt-Verfahren* (erfunden von Jørgen Gram, popularisiert durch Erhard Schmidt).

Wir beginnen mit einer Basis $\mathbf{b}_1, \dots, \mathbf{b}_n$. Was passiert, wenn wir ein größeres Erzeugendensystem haben, überlegen wir uns danach. Im ersten Schritt erzeugen wir eine ONB für $\langle \mathbf{b}_1 \rangle$, also die lineare Hülle des ersten Basisvektors. Die einfachste Wahl ist

$$\mathbf{q}_1 = \frac{1}{\|\mathbf{b}_1\|} \mathbf{b}_1.$$

Dieser hat Norm Eins, wie man leicht verifiziert. Im zweiten Schritt erzeugen wir eine ONB für $\langle \mathbf{b}_1, \mathbf{b}_2 \rangle$. Unser bereits gewonnenes \mathbf{q}_1 wollen wir festhalten. Nun zerlegen wir entsprechend (2.6) unseren zweiten Basisvektor in ein Vielfaches von \mathbf{q}_1 und einen Anteil, der senkrecht auf \mathbf{q}_1 steht:

$$\mathbf{b}_2 = (\mathbf{b}_2, \mathbf{q}_1) \mathbf{q}_1 + (\mathbf{b}_2 - (\mathbf{b}_2, \mathbf{q}_1) \mathbf{q}_1).$$

Damit steht also der Vektor

$$\tilde{\mathbf{q}}_2 := \mathbf{b}_2 - (\mathbf{b}_2, \mathbf{q}_1) \mathbf{q}_1$$

senkrecht auf \mathbf{q}_1 . Wir müssen ihn nur noch normieren:

$$\mathbf{q}_2 = \frac{1}{\|\tilde{\mathbf{q}}_2\|} \tilde{\mathbf{q}}_2.$$

Bis dahin war alles einfach. Nun müssen wir (für $n \geq 3$) in $\langle \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3 \rangle$ einen Vektor finden, der sowohl auf \mathbf{q}_1 als auch auf \mathbf{q}_2 senkrecht steht. Damit er auf \mathbf{q}_1 senkrecht steht, würde reichen

$$\mathbf{b}_3 - (\mathbf{b}_3, \mathbf{q}_1) \mathbf{q}_1,$$

um senkrecht auf \mathbf{q}_2 zu stehen

$$\mathbf{b}_3 - (\mathbf{b}_3, \mathbf{q}_2) \mathbf{q}_2.$$

Das legt die Vermutung nahe, dass

$$\tilde{\mathbf{q}}_3 := \mathbf{b}_3 - (\mathbf{b}_3, \mathbf{q}_1) \mathbf{q}_1 - (\mathbf{b}_3, \mathbf{q}_2) \mathbf{q}_2$$

das Lot von \mathbf{b}_3 auf $\langle \mathbf{q}_1, \mathbf{q}_2 \rangle$ bzw. $\langle \mathbf{b}_1, \mathbf{b}_2 \rangle$ ist. Wir probieren das aus:

$$\begin{aligned} (\tilde{\mathbf{q}}_3, \mathbf{q}_1) &= (\mathbf{b}_3 - (\mathbf{b}_3, \mathbf{q}_1) \mathbf{q}_1 - (\mathbf{b}_3, \mathbf{q}_2) \mathbf{q}_2, \mathbf{q}_1) \\ &= 0 - 0 = 0. \end{aligned}$$

Schließlich steht $\mathbf{b}_3 - (\mathbf{b}_3, \mathbf{q}_1) \mathbf{q}_1$ nach Konstruktion senkrecht auf \mathbf{q}_1 und der Rest als Vielfaches von \mathbf{q}_2 auch. Analog sieht es aus, wenn wir das Skalarprodukt mit \mathbf{q}_2 bilden. Unser neuer Vektor steht also senkrecht auf allen Linearkombinationen von \mathbf{q}_1 und \mathbf{q}_2 . Durch Normieren wie oben erhalten wir also unser gewünschtes \mathbf{q}_3 .

2 Das Innenprodukt kommt ins Spiel

Wir hätten auch zuerst das Lot auf \mathbf{q}_1 und von diesem dann das Lot auf \mathbf{q}_2 fällen können. Das Ergebnis wäre, wie man leicht nachrechnet, das gleiche. Bei exakter Rechnung können wir uns also Arbeit sparen und wie oben einfach die *Orthogonalprojektionen* des neu dazugekommenen Vektors auf die bisher gewonnenen Vektoren der ONB von diesem abziehen. Das führt dann auf das *Gram-Schmidt-Verfahren*. Berechnet man sukzessive die Orthogonalzerlegungen, so führt das auf das sogenannte *Modifizierte Gram-Schmidt-Verfahren (MGS)*, das vor allem in der Numerik von Interesse ist, weil sich dadurch die Effekte von Rundungsfehlern reduzieren.

Zusammengefasst haben wir

Algorithmus 40 (Gram-Schmidt).

- Setze $r_{11} = \|\mathbf{b}_1\|$.
- Bilde daraus $\mathbf{q}_1 = \mathbf{b}_1/r_{11}$.
- Für $j = 2, \dots, n$:
 - Bilde für $k = 1, \dots, j-1$ die Skalarprodukte

$$r_{kj} = (\mathbf{b}_j, \mathbf{q}_k).$$

- Berechne daraus

$$\tilde{\mathbf{q}}_j = \mathbf{b}_j - \sum_{k=1}^{j-1} r_{kj} \mathbf{q}_k.$$

- Berechne $r_{jj} = \|\tilde{\mathbf{q}}_j\|$.
- Setze $\mathbf{q}_j = \tilde{\mathbf{q}}_j/r_{jj}$.

Bemerkung 41. Setzt man die r_{kj} in eine Matrix ein, so füllen sie das rechte obere Dreieck. Füllt man den Rest mit Nullen und setzt

$$\mathbf{B} = (\mathbf{b}_1 \mid \cdots \mid \mathbf{b}_n) \quad \text{sowie} \quad \mathbf{Q} = (\mathbf{q}_1 \mid \cdots \mid \mathbf{q}_n),$$

so ergibt sich

$$\mathbf{B} = \mathbf{QR}. \quad (2.16)$$

So etwas verwendet man in der Numerik für reelle Matrizen und mit dem Standardskalarprodukt. Dort trägt diese multiplikative Zerlegung von \mathbf{B} den Namen *QR-Zerlegung* (von \mathbf{B}). Die Matrix \mathbf{Q} hat dann die nette Eigenschaft, dass $\mathbf{Q}^{-1} = \mathbf{Q}^T$, was sich gut zum Lösen linearer Gleichungssysteme eignet. Spannend wird das Ganze, wenn man nicht nur quadratische Matrizen betrachtet. Außerdem gibt es noch andere Algorithmen zum Berechnen einer solchen Zerlegung, die für numerische Zwecke freundlicher sind.

Satz 42. *Startet man mit einer Basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ des Innenproduktraums V , so liefert das Gram-Schmidt-Verfahren eine Orthonormalbasis von V . Insbesondere hat damit jeder endlich-dimensionale Innenproduktraum eine ONB.*

Beweis. Wir müssen nur noch sicher stellen, dass keines der $\tilde{\mathbf{q}}_j$ der Nullvektor ist. Das kann aber nicht passieren, weil sonst bereits \mathbf{b}_j in $\langle \mathbf{q}_1, \dots, \mathbf{q}_{j-1} \rangle = \langle \mathbf{b}_1, \dots, \mathbf{b}_{j-1} \rangle$ liegen müsste, was aber per Voraussetzung ausgeschlossen ist. Wir haben ja schließlich eine Basis, die damit automatisch linear unabhängig ist. \square

Bemerkung 43. Ist das Erzeugendensystem mit dem wir starten nicht l. u., so wird an (mindestens) einer Stelle ein $\tilde{\mathbf{q}}_j = \mathbf{0}$ auftauchen. Dann sortieren wir das zugehörige \mathbf{b}_j einfach aus und machen mit dem Rest weiter.

Beispiel 44. Wir betrachten den \mathbb{R}^3 als Raum der Koeffizienten von quadratischen Polynomen, d. h. der Vektor $\mathbf{v} = (v_0, v_1, v_2)$ (man beachte die „verrutschte“ Nummerierung der Komponenten) steht für das Polynom mit $p_{\mathbf{v}}(x) = v_0 + v_1x + v_2x^2$. Dann können wir dort (siehe Übung) ein Skalarprodukt definieren durch

$$(\mathbf{u}, \mathbf{v}) = \int_{-1}^1 p_{\mathbf{u}}(x)p_{\mathbf{v}}(x) dx .$$

Wir nehmen die Standardeinheitsvektoren als Basis. Diese entspricht den Polynomen $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = x^2$, die man auch gerne als *Monome* bezeichnet. Damit ist

$$r_{11} = \|\mathbf{e}_1\| = \sqrt{\int_{-1}^1 1 dx} = \sqrt{2} .$$

Wir setzen also $\mathbf{q}_1 = \mathbf{e}_1/\sqrt{2}$, was gerade dem konstanten Polynom $\sqrt{2}/2$ entspricht.

Wir berechnen nun

$$r_{12} = (\mathbf{e}_2, \mathbf{q}_1) = \int_{-1}^1 \frac{\sqrt{2}x}{2} dx = 0$$

und daraus

$$\tilde{\mathbf{q}}_2 = \mathbf{e}_2 - 0\mathbf{q}_1 = \mathbf{e}_2 .$$

Das ist gerade wieder der ursprüngliche Basisvektor selber. Für diesen finden wir

$$r_{22}^2 = \|\tilde{\mathbf{q}}_2\|^2 = \|\mathbf{e}_2\|^2 = \int_{-1}^1 x^2 dx = \frac{2}{3} ,$$

weshalb unser zweiter Vektor in der ONB lautet

$$\mathbf{q}_2 = \sqrt{\frac{3}{2}} \mathbf{e}_2 ,$$

was gerade dem Polynom $\sqrt{3/2}x$ entspricht.

Für den dritten Vektor müssen wir zwei Skalarprodukte berechnen:

$$r_{13} = (\mathbf{e}_3, \mathbf{q}_1) = \int_{-1}^1 x^2 \cdot \frac{\sqrt{2}}{2} dx = \frac{\sqrt{2}}{3} ,$$

$$r_{23} = (\mathbf{e}_3, \mathbf{q}_2) = \int_{-1}^1 x^2 \cdot \sqrt{\frac{3}{2}}x dx = 0 .$$

2 Das Innenprodukt kommt ins Spiel

Damit erhalten wir

$$\tilde{\mathbf{q}}_3 = \mathbf{e}_3 - r_{13}\mathbf{q}_1 - r_{23}\mathbf{q}_2 = \mathbf{e}_3 - \frac{1}{3}\mathbf{e}_1.$$

Für die Norm dieses Vektors bzw. das Quadrat davon finden wir damit

$$\begin{aligned} r_{33}^2 = \|\mathbf{e}_3 - \frac{1}{3}\mathbf{e}_1\|^2 &= \int_{-1}^1 \left(x^2 - \frac{1}{3}\right)^2 dx \\ &= \int_{-1}^1 x^4 - \frac{2}{3}x^2 + \frac{1}{9} dx = \frac{2}{5} - \frac{4}{9} + \frac{2}{9} = \frac{8}{45}. \end{aligned}$$

Also hat unser dritter Vektor für die ONB die Gestalt

$$\mathbf{q}_3 = \frac{3\sqrt{5}}{2\sqrt{2}}\tilde{\mathbf{q}}_3 = \frac{3\sqrt{5}}{2\sqrt{2}}\left(\mathbf{e}_3 - \frac{1}{3}\mathbf{e}_1\right) = \frac{3\sqrt{5}}{2\sqrt{2}}\mathbf{e}_3 - \frac{\sqrt{5}}{2\sqrt{2}}\mathbf{e}_1,$$

was gerade dem Polynom $\frac{3\sqrt{5}}{2\sqrt{2}}x^2 - \frac{\sqrt{5}}{2\sqrt{2}}$ entspricht.

Insgesamt erhalten wir bezüglich des gewählten Innenprodukts die Orthonormalbasis

$$\mathbf{q}_1 = \sqrt{2}/2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{q}_2 = \sqrt{3/2} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{q}_3 = \frac{\sqrt{5}}{2\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}.$$

Übersetzt in Polynome hätten wir aus $1, x, x^2$ die „ONB“

$$\sqrt{2}/2, \quad \sqrt{3/2}x, \quad \frac{\sqrt{10}}{4}(3x^2 - 1)$$

gewonnen. Wenn wir später zu allgemeinen Vektorräumen kommen, werden wir sehen, dass es tatsächlich eine ONB im Raum der Polynome vom Grad höchstens zwei ist.

Korollar 45 (zu Satz 42). *In jedem (endlichdimensionalen) Innenproduktraum lässt sich jede orthonormale Liste zu einer Orthonormalbasis ergänzen.*

Beweis. Da die gegebene Liste, etwa $\mathbf{q}_1, \dots, \mathbf{q}_k$ linear unabhängig ist, lässt sie sich nach dem Basisergänzungssatz (siehe ELA Satz 42) zu einer Basis $\mathbf{q}_1, \dots, \mathbf{q}_k, \mathbf{b}_{k+1}, \dots, \mathbf{b}_n$ unseres Innenproduktraums ergänzen. Wenn wir entsprechend Satz 42 mit Gram-Schmidt darauf losgehen, ergibt sich eine ONB, bei der die ersten k Vektoren unverändert bleiben, also unsere gegebenen $\mathbf{q}_1, \dots, \mathbf{q}_k$. \square

Nun kommt der große Moment, wo wir wieder den Bogen zur Trigonalisierung von Matrizen schlagen und die erste der angekündigten Normalformen, genauer deren Existenz, beweisen können. Der Schlüssel ist wie beim Beweis der Trigonalisierbarkeit selber (Satz 11) unser Lemma 10, in dem wir Bedingungen gesammelt haben, die äquivalent dazu sind, dass sich ein Endomorphismus durch eine obere Dreiecksmatrix darstellen lässt. Wir gehen in zwei Stufen vor. Zunächst

Satz 46. Sei V ein Innenproduktraum und φ eine Endomorphismus auf V . Gibt es eine Basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ von V bezüglich der ϕ durch eine obere Dreiecksmatrix dargestellt wird, so gibt es auch eine ONB $\mathbf{q}_1, \dots, \mathbf{q}_n$ bezüglich der das auch der Fall ist.

Beweis. Nach Lemma 10 ist $\langle \mathbf{b}_1, \dots, \mathbf{b}_j \rangle$ für jedes $j = 1, \dots, n$ invariant unter φ . Nun lassen wir Gram-Schmidt auf unsere Basis los und erhalten damit eine ONB $\mathbf{q}_1, \dots, \mathbf{q}_n$. Nun ist aber für jedes j

$$\langle \mathbf{q}_1, \dots, \mathbf{q}_j \rangle = \langle \mathbf{b}_1, \dots, \mathbf{b}_j \rangle$$

invariant unter φ . Nach Lemma 10 wird damit φ auch bezüglich der ONB $\mathbf{q}_1, \dots, \mathbf{q}_n$ durch eine obere Dreiecksmatrix dargestellt. \square

Der eigentliche Satz ergibt sich direkt daraus:

Korollar 47 (von Schur (erste Version)). Sei V ein komplexer Innenproduktraum und φ eine Endomorphismus auf V . Dann gibt es eine Orthonormalbasis $\mathbf{q}_1, \dots, \mathbf{q}_n$ bezüglich der φ durch eine obere Dreiecksmatrix dargestellt wird.

Beweis. Das ist eine direkte Folgerung aus Satz 46 zusammen mit Satz 11, nach dem in V eine Basis existiert bzgl. der sich φ durch eine obere Dreiecksmatrix darstellen lässt. \square

Definition 48. Die so gewonnene Dreiecksmatrix heißt *Schursche Normalform* (nach Issai Schur, der ihre Existenz 1909 bewiesen hat).

Bemerkung 49. Hat eine reelle Matrix bzw. ein Endomorphismus eines reellen Vektorraums nur reelle Eigenwerte, so ist damit klar, dass auch hier die Darstellung als obere Dreiecksmatrix funktioniert. (Siehe Übung)

Bemerkung 50. Leider gibt uns der Satz keine Mittel an die Hand, diese Darstellung konkret auszurechnen. Auf der Diagonalen stehen schließlich die Eigenwerte, von denen wir schon gelernt haben, dass sie die Nullstellen des charakteristischen Polynoms sind. Es gibt aber nur bis zum Polynomgrad vier allgemeine Methoden, die uns die Berechnung der Eigenwerte (garantiert) ermöglichen. Man muss die Schursche Normalform numerisch approximieren. Näheres dazu gibt es im Kurs Numerik II.

2.3 Das orthogonale Komplement

Wenn wir in unserem Innenproduktraum V eine ONB $\mathbf{q}_1, \dots, \mathbf{q}_n$ haben und diese in zwei Teile teilen, etwa $\mathbf{q}_1, \dots, \mathbf{q}_k$ und $\mathbf{q}_{k+1}, \dots, \mathbf{q}_n$, dann gilt für die Unterräume $U := \langle \mathbf{q}_1, \dots, \mathbf{q}_k \rangle$ und $W := \langle \mathbf{q}_{k+1}, \dots, \mathbf{q}_n \rangle$ nicht nur $V = U \oplus W$, sondern auch

$$\mathbf{u} \perp \mathbf{w} \quad \forall \mathbf{u} \in U, \mathbf{w} \in W,$$

die Unterräume stehen also quasi im Gesamten senkrecht aufeinander. Man könnte symbolisch schreiben $U \perp W$. Da W über die direkte Summe U zu ganz V komplettiert und

2 Das Innenprodukt kommt ins Spiel

überdies senkrecht auf U steht, könnte man sagen, W ist das *orthogonale Komplement* von U (und umgekehrt). Diese kurze Betrachtung wollen wir verallgemeinern, indem wir *orthogonale Komplemente* für allgemeine Teilmengen eines Innenproduktraums anschauen.

Zuerst die naheliegende Definition:

Definition 51. Ist U eine Teilmenge eines Innenproduktraums V , so definieren wir ihr *orthogonales Komplement* als

$$U^\perp := \{v \in V : v \perp u \ \forall u \in U\} = \{v \in V : (v, u) = 0 \ \forall u \in U\}.$$

Beispiel 52. Sei im \mathbb{R}^5 , versehen mit dem Standardskalarprodukt die Menge $U = \{u, v, w\}$ gegeben mit

$$\begin{aligned}u &= (1, 2, 3, 4, 5)^T, \\v &= (0, 1, 2, 3, 4)^T, \\w &= (0, 0, 1, 2, 3)^T,\end{aligned}$$

dann bedeutet die Berechnung von U^\perp , die Vektoren x zu finden, für die

$$u^T x = 0, \quad v^T x = 0, \quad w^T x = 0.$$

Um dies simultan zu lösen, schreibt man es am einfachsten als

$$\begin{pmatrix} u^T \\ v^T \\ w^T \end{pmatrix} x = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Offenbar ist also U^\perp die Lösung eines homogenen LGS und damit ein Untervektorraum von V .

Es ist leicht zu sehen, dass in diesem Fall das orthogonale Komplement einer Basis und damit des ganzen Raums gerade der Nullraum ist. Das gilt umgekehrt auch, steht doch der Nullvektor auf allen senkrecht (nicht nur für das Standardskalarprodukt).

Hätten wir einen der drei Vektoren weggelassen, wäre das orthogonale Komplement größer geworden.

Das Beispiel gibt uns den Mut, diese Dinge für den allgemeinen Fall zu zeigen:

Satz 53. In einem Innenproduktraum V gelten

1. Für $U \subseteq V$ ist U^\perp ein UVR von V .
2. $\{0\}^\perp = V$.
3. $V^\perp = \{0\}$.

4. Für $U \subseteq V$ ist $U \cap U^\perp \subseteq \{\mathbf{0}\}$.

5. Ist $U \subseteq W \subseteq V$, so gilt $W^\perp \subseteq U^\perp$.

Beweis.

1. Da der Nullvektor auf allen anderen senkrecht steht, liegt er automatisch in U^\perp .

Seien nun $\mathbf{v}, \mathbf{w} \in U^\perp$, λ ein Skalar und $\mathbf{u} \in U$. Dann gilt wegen der Linearität des Innenprodukts im ersten Argument

$$(\lambda \mathbf{v} + \mathbf{w}, \mathbf{u}) = \lambda(\mathbf{v}, \mathbf{u}) + (\mathbf{w}, \mathbf{u}) = \lambda \cdot 0 + 0 = 0.$$

Weil die Vektoren und der Skalar zwar fest, aber beliebig waren (in den angegebenen Grenzen), ist damit für U^\perp das Untervektorraumkriterium erfüllt.

2. Das folgt, weil der Nullvektor als einziger auf allen senkrecht steht.

3. Das folgt, weil alle Vektoren inklusive ihm selber auf dem Nullvektor senkrecht stehen.

4. Das folgt, weil der Nullvektor als einziger auf sich selber senkrecht steht.

5. In diesem Fall steht ein Vektor, der auf W senkrecht steht, automatisch auch auf jedem Vektor aus U senkrecht. Damit folgt die Behauptung direkt.

□

Da der erste und der vierte Teil natürlich auch gelten, wenn U selber ein UVR von V ist, legt sich die Vermutung nahe, dass sich in diesem Fall der gesamte Raum V als direkte Summe von Unterraum und orthogonalem Komplement schreiben lässt:

Satz 54. Ist V ein Innenproduktraum und U ein UVR von V , so gilt

$$V = U \oplus U^\perp.$$

Beweis. Nach unserer Vorbemerkung müssen wir nur noch zeigen, dass tatsächlich $V = U + U^\perp$.

Wir starten mit einer ONB $\mathbf{q}_1, \dots, \mathbf{q}_k$ unseres Teilraums U und wählen uns einen beliebigen aber festen Vektor $\mathbf{v} \in V$. Ziehen wir von diesem

$$\mathbf{w} := \mathbf{v} - \sum_{j=1}^k (\mathbf{v}, \mathbf{q}_j) \mathbf{q}_j$$

ab, erhalten wir einen Vektor

$$\mathbf{u} := \sum_{j=1}^k (\mathbf{v}, \mathbf{q}_j) \mathbf{q}_j \in U.$$

2 Das Innenprodukt kommt ins Spiel

Jetzt müssen wir nur noch zeigen, dass \mathbf{w} senkrecht auf U steht. Dafür müssen wir es nur für die Vektoren aus der ONB von U testen:

$$\begin{aligned}(\mathbf{w}, \mathbf{q}_\ell) &= \left(\mathbf{v} - \sum_{j=1}^k (\mathbf{v}, \mathbf{q}_j) \mathbf{q}_j, \mathbf{q}_\ell \right) \\ &= (\mathbf{v}, \mathbf{q}_\ell) - \sum_{j=1}^k (\mathbf{v}, \mathbf{q}_j) (\mathbf{q}_j, \mathbf{q}_\ell) \\ &= (\mathbf{v}, \mathbf{q}_\ell) - (\mathbf{v}, \mathbf{q}_\ell) \\ &= 0.\end{aligned}$$

Damit liegt aber unser frei gewähltes $\mathbf{v} \in V$ auch in $U + U^\perp$, und weil es frei gewählt war, folgt die Behauptung. \square

Zusammen mit dem Dimensionssatz aus der ELA folgt hieraus

Korollar 55. *Ist V ein Innenproduktraum und U ein UVR von V , so gilt*

$$\dim U^\perp = \dim V - \dim U.$$

Jetzt fehlt noch die naheliegende Tatsache

Korollar 56. *Ist U ein UVR eines Innenproduktraums V , dann ist*

$$(U^\perp)^\perp = U.$$

Beweis. Die Inklusion

$$U \subseteq (U^\perp)^\perp \tag{2.17}$$

ist offensichtlich. Da aber nach Korollar 55 links und rechts die Dimension gleich sein muss, sind wir schon fertig. \square

Bemerkung 57. Wenn wir den Begriff des Innenproduktraums verallgemeinern und auch unendlichdimensionale Räume bekommen, können wir keine Dimensionsargumente mehr verwenden, und der Beweis wird etwas aufwendiger:

Der Anfang, also die Inklusion (2.17) bleibt gleich. Für die umgekehrte Inklusion wählen wir ein $\mathbf{v} \in (U^\perp)^\perp$. Dies können wir nach Satz 54 in eindeutiger Weise schreiben als $\mathbf{v} = \mathbf{u} + \mathbf{w}$ mit $\mathbf{u} \in U$ und $\mathbf{w} \in U^\perp$. Wegen der Inklusion (2.17) ist damit aber $\mathbf{u} \in (U^\perp)^\perp$. Da das Komplement und somit auch das Komplement des Komplements ein UVR ist, ist damit auch $\mathbf{w} = \mathbf{v} - \mathbf{u} \in (U^\perp)^\perp$. Damit ist $\mathbf{w} \in U^\perp \cap (U^\perp)^\perp = \{\mathbf{0}\}$, also selber der Nullvektor und damit $\mathbf{v} = \mathbf{u} \in U$. Insgesamt also wie behauptet auch $(U^\perp)^\perp \subseteq U$.

Der Beweis von Korollar 56 funktioniert also auch ohne jegliche Betrachtung der Dimension.

Die Aufspaltung eines Vektors in einen Teil der in einem Unterraum und einen, der in dessen orthogonalem Komplement liegt, formulieren wir durch eine Abbildung:

Definition 58. Sei V ein Innenproduktraum und U ein Unterraum von V . Hat ein Vektor $\mathbf{v} \in V$ die Darstellung $\mathbf{v} = \mathbf{u} + \mathbf{w}$ mit $\mathbf{u} \in U$ und $\mathbf{w} \in U^\perp$, so heißt \mathbf{u} die *Orthogonalprojektion* von \mathbf{v} auf U . Die Abbildung, die dieses bewerkstelligt, nennen wir die *Orthogonalprojektion* von V auf U . Wir schreiben dann $P_U(\mathbf{v}) = \mathbf{u}$ (bezogen auf die obige Zerlegung).

Wegen Satz 54 ist die Orthogonalprojektion tatsächlich eine wohldefinierte Abbildung. Sonst wäre die Summe von Unterraum und orthogonalem Komplement nicht direkt.

Beispiel 59. Betrachten wir die Projektion auf einen eindimensionalen Unterraum $U = \langle \mathbf{u} \rangle$, so können wir direkt unsere Orthogonalzerlegung (2.6) verwenden, also

$$\mathbf{v} = \frac{(\mathbf{v}, \mathbf{u})}{\|\mathbf{u}\|^2} \mathbf{u} + \left(\mathbf{v} - \frac{(\mathbf{v}, \mathbf{u})}{\|\mathbf{u}\|^2} \mathbf{u} \right).$$

Die Orthogonalprojektion ist dann gerade der erste Teil, also

$$P_U(\mathbf{v}) = \frac{(\mathbf{v}, \mathbf{u})}{\|\mathbf{u}\|^2} \mathbf{u}.$$

Implizit haben wir das ganze für auch für zwei-, drei-, höherdimensionale Unterräume auch schon bei Gram-Schmidt verwendet.

Die wichtigsten Eigenschaften der Orthogonalprojektion sammeln wir wieder in einem Satz:

Satz 60. Sei V ein Innenproduktraum, U ein Untervektorraum von V und $\mathbf{v} \in V$. Dann gilt

1. P_U ist ein Endomorphismus auf V .
2. Für jedes $\mathbf{u} \in U$ ist $P_U(\mathbf{u}) = \mathbf{u}$.
3. Für jedes $\mathbf{w} \in U^\perp$ ist $P_U(\mathbf{w}) = \mathbf{0}$.
4. $\text{im } P_U = U$,
5. $\ker P_U = U^\perp$.
6. $\mathbf{v} - P_U(\mathbf{v}) \in U^\perp$. (Das Lot von \mathbf{v} auf U)
7. $P_U^2 (= P_U \circ P_U) = P_U$. (Daher der Name Projektion.)
8. Mit der vom Innenprodukt induzierten Norm ist $\|P_U(\mathbf{v})\| \leq \|\mathbf{v}\|$.
9. Ist $\mathbf{q}_1, \dots, \mathbf{q}_k$ eine ONB von U , so ist

$$P_U(\mathbf{v}) = \sum_{j=1}^k (\mathbf{v}, \mathbf{q}_j) \mathbf{q}_j.$$

2 Das Innenprodukt kommt ins Spiel

Den letzten Punkt hatten wir implizit schon bei der Gram-Schmidt-Orthogonalisierung verwendet.

Beweis. Übung. □

Die orthogonale Projektion eines Vektors auf einen Unterraum erfüllt die folgende *Minimalitätseigenschaft*:

Satz 61. Sei U ein Unterraum eines Innenproduktraums V und $\mathbf{u} \in U$ sowie $\mathbf{v} \in V$. Dann ist

$$\|\mathbf{v} - P_U(\mathbf{v})\| \leq \|\mathbf{v} - \mathbf{u}\|. \quad (2.18)$$

Gleichheit gilt genau dann, wenn \mathbf{v} selber in U lag, also genau dann wenn $P_U(\mathbf{v}) = \mathbf{u}$.

Beweis. Da wir in obigem Satz gezeigt hatten, dass $\mathbf{v} - P_U(\mathbf{v})$ orthogonal auf jedem $\mathbf{u} \in U$ steht, können wir durch künstliche Addition und den Satz des Pythagoras folgendes bekommen:

$$\begin{aligned} \|\mathbf{v} - P_U(\mathbf{v})\|^2 &\leq \|\mathbf{v} - P_U(\mathbf{v})\|^2 + \|P_U(\mathbf{v}) - \mathbf{u}\|^2 \\ &= \|(\mathbf{v} - P_U(\mathbf{v})) + (P_U(\mathbf{v}) - \mathbf{u})\|^2 \\ &= \|\mathbf{v} - \mathbf{u}\|^2. \end{aligned}$$

Gleichheit gilt genau dann, wenn $\|P_U(\mathbf{v}) - \mathbf{u}\|$, der künstlich addierte Term, verschwindet. Das ist aber genau für $P_U(\mathbf{v}) = \mathbf{u}$ der Fall. □

Dass das Lot die kürzeste Verbindung eines Punkts zu einer Ebene oder Gerade darstellt, war uns aus der Geometrie bereits klar. Hier haben wir gezeigt, dass es bei jedem Innenprodukt mit der zugehörigen Norm funktioniert.

2.4 Lineare Funktionale: Was lineare Abbildungen ins Eindimensionale mit Innenprodukten zu tun haben

Dieser Abschnitt wird recht kurz. Die Überschrift könnte schon als Einleitung ausreichen. Aber erst mal die Definition:

Definition 62. Ist V ein reeller Vektorraum und $\psi \in \text{hom}(V, \mathbb{R})$ oder V ein komplexer Vektorraum und entsprechend $\psi \in \text{hom}(V, \mathbb{C})$, dann nennen wir ψ ein *lineares Funktional* auf V .

Der Begriff *Funktional* steht auch in anderen Zusammenhängen immer für eine Abbildung von etwas höherdimensionalem ins Eindimensionale.

In diesem Abschnitt wollen wir vor allem schauen, was lineare Funktionale auf einem Innenproduktraum mit dem Innenprodukt zu tun haben.

Beispiel 63. Durch

$$\psi(\mathbf{z}) = 2z_1 - 5z_2 + z_3$$

ist ein lineares Funktional auf \mathbb{C}^3 bzw. \mathbb{R}^3 gegeben. Wir könnten es mittels des Standardskalarprodukts auch schreiben als

$$\psi(\mathbf{z}) = (\mathbf{z}, (2, -5, 1)^T).$$

Wir könnten es aber auch als lineares Funktional auf \mathbb{C}^4 oder \mathbb{R}^4 auffassen. Dann müssten wir bei der Darstellung mittels des Standardskalarprodukts den Vektor im zweiten Argument um eine Null ergänzen zu $(2, -5, 1, 0)^T$ u. s. w.

Wir schauen nun, ob eine solche Darstellung per Innenprodukt in jedem (endlichdimensionalen) Innenproduktraum funktioniert. Dazu wählen wir dort zuerst eine ONB, etwa $\mathbf{q}_1, \dots, \mathbf{q}_n$ und halten ein $\mathbf{v} \in V$ fest. Dann wissen wir

$$\mathbf{v} = \sum_{j=1}^n (\mathbf{v}, \mathbf{q}_j) \mathbf{q}_j$$

und damit

$$\psi(\mathbf{v}) = \sum_{j=1}^n (\mathbf{v}, \mathbf{q}_j) \psi(\mathbf{q}_j).$$

Da $\psi(\mathbf{q}_j)$ können wir es (komplex konjugiert) ins zweite Argument des Innenprodukts reinziehen und danach die Additivität anwenden:

$$\psi(\mathbf{v}) = \sum_{j=1}^n (\mathbf{v}, \overline{\psi(\mathbf{q}_j)} \mathbf{q}_j) = \left(\mathbf{v}, \sum_{j=1}^n \overline{\psi(\mathbf{q}_j)} \mathbf{q}_j \right) =: (\mathbf{v}, \mathbf{u}).$$

Wir haben also das zweite Argument zu einem Vektor $\mathbf{u} \in V$ zusammengefasst. Da \mathbf{v} zwar fest, aber beliebig war, tut es dieses \mathbf{u} für jedes $\mathbf{v} \in V$ und hängt nur vom linearen Funktional ψ ab.

Es bleibt nur noch die Frage, ob es womöglich einen weiteren Vektor mit dieser Eigenschaft gibt, also

$$(\mathbf{v}, \mathbf{w}) = (\mathbf{v}, \mathbf{u}) = \psi(\mathbf{v}).$$

Dann muss aber gelten

$$0 = (\mathbf{v}, \mathbf{w}) - (\mathbf{v}, \mathbf{u}) = (\mathbf{v}, \mathbf{w} - \mathbf{u}) \quad \forall \mathbf{v} \in V,$$

d. h., $(\mathbf{w} - \mathbf{u}) \perp \mathbf{v}$ für jedes $\mathbf{v} \in V$. Der einzige Vektor, der auf allen Vektoren inklusive sich selbst senkrecht steht, ist der Nullvektor, und wir haben gewonnen.

Das Resultat trägt einen eigenen Namen:

Satz 64 (Darstellungssatz von Riesz). *Jedes lineare Funktional auf einem (endlichdimensionalen) Innenproduktraum hat eine Darstellung der Gestalt*

$$\psi(\cdot) = (\cdot, \mathbf{u})$$

mit einem eindeutigen, nur von ψ abhängigen $\mathbf{u} \in V$.

2 Das Innenprodukt kommt ins Spiel

Wichtig ist, zu beachten, dass es im Komplexen wirklich das erste Argument sein muss, das „frei“ bleibt, denn nur dort haben wir Linearität verlangt. Im Reellen ist es egal.

2.5 Wann wir nur reelle Eigenwerte haben

Wir beginnen eiskalt mit einem neuen Begriff:

Definition 65. Seien V und W Innenprodukträume und $\varphi \in \text{hom}(V, W)$. Die Abbildung $\varphi^* \in \text{hom}(W, V)$, die definiert ist durch

$$(\varphi(\mathbf{v}), \mathbf{w}) = (\mathbf{v}, \varphi^*(\mathbf{w})) \quad \forall \mathbf{v} \in V, \mathbf{w} \in W \quad (2.19)$$

heißt die *Adjungierte* oder auch *adjungierte Abbildung* von φ .

(Hat übrigens nichts mit der adjunkten Matrix aus ELA zu tun, auch wenn der Name ähnlich klingt.)

Da wir in der Definition eine Behauptung versteckt haben, nämlich, dass Formel (2.19) tatsächlich eine eindeutige Abbildung definiert, müssen wir hier einen Beweis führen, den der *Wohldefiniertheit*:

Beweis der Wohldefiniertheit.

Halten wir \mathbf{w} fest, so definiert $(\varphi(\mathbf{v}), \mathbf{w})$ als Hintereinanderausführung linearer Abbildungen selber eine lineare Abbildung. Weil sie ins Skalare geht, ein lineares Funktional. Dieses lässt sich nach dem Rieszschen Darstellungssatz 64 darstellen durch einen eindeutigen Vektor $\mathbf{u} \in V$ als $(\varphi(\cdot), \mathbf{w}) = (\cdot, \mathbf{u})$. Unserem $\mathbf{w} \in W$, mit dem wir gestartet sind, ist also ein eindeutiges $\mathbf{u} \in V$ zugeordnet. Es gibt also eine Funktion, die dieses tut, der wir nur noch den Namen φ^* geben müssen. \square

Beispiel 66. Wir betrachten die lineare Abbildung $\varphi_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$; $\mathbf{x} \mapsto \mathbf{Ax}$ mit $\mathbf{A} \in \mathbb{R}^{m \times n}$ und verwenden in beiden Räumen das Standardskalarprodukt. Dann ist

$$(\varphi_A(\mathbf{x}), \mathbf{y}) = (\mathbf{Ax}, \mathbf{y}) = (\mathbf{Ax})^T \mathbf{y} = \mathbf{x}^T \mathbf{A}^T \mathbf{y} = (\mathbf{x}, \mathbf{A}^T \mathbf{y}).$$

Die Adjungierte ist also

$$\varphi_A^* = \varphi_{A^T} : \mathbb{R}^m \rightarrow \mathbb{R}^n; \mathbf{y} \mapsto \mathbf{A}^T \mathbf{y}.$$

Wir müssten also nur die Transponierte Matrix nehmen.

Beispiel 67. Nehmen wir die gleichen Voraussetzungen wie im vorigen Beispiel, außer dass wir überall \mathbb{R} durch \mathbb{C} ersetzen, dann finden wir analog

$$(\varphi_A(\mathbf{x}), \mathbf{y}) = (\mathbf{Ax})^T \bar{\mathbf{y}} = \mathbf{x}^T \mathbf{A}^T \bar{\mathbf{y}} = \mathbf{x}^T \overline{\mathbf{A}^H \mathbf{y}} = (\mathbf{x}, \mathbf{A}^H \mathbf{y}).$$

Wir müssen die Abbildungsmatrix also nicht nur transponieren, sondern zusätzlich komplex konjugieren.

Beiden Beispielen ist gemeinsam, dass die Adjungierte Abbildung selber linear ist. Da die linearen Abbildungen im endlichdimensionalen direkt mit den Matrizen korrespondieren, ist das immer so, wenn wir das Standardskalarprodukt verwenden. Der Verdacht legt sich nahe, dass folgendes gilt:

Satz 68. *Die Adjungierte einer linearen Abbildung zwischen zwei Innenprodukträumen ist selber linear.*

Beweis. Seien wieder V und W unsere Innenprodukträume und w_1, w_2 zwei Vektoren aus W sowie λ ein Skalar. Dann ist für ein beliebiges $v \in V$ mit der Definition der Adjungierten und den Eigenschaften des Innenprodukts

$$\begin{aligned} (v, \varphi^*(\lambda w_1 + w_2)) &= (\varphi(v), \lambda w_1 + w_2) \\ &= \bar{\lambda}(\varphi(v), w_1) + (\varphi(v), w_2) \\ &= \bar{\lambda}(v, \varphi^*(w_1)) + (v, \varphi^*(w_2)) \\ &= (v, \lambda \varphi^*(w_1) + \varphi^*(w_2)). \end{aligned}$$

Da v beliebig war, folgt daraus

$$\varphi^*(\lambda w_1 + w_2) = \lambda \varphi^*(w_1) + \varphi^*(w_2),$$

und weil λ sowie die $w_1, w_2 \in W$ ebenfalls beliebig waren, die Linearität von φ^* . \square

Dadurch ermutigt, versuchen wir die Regeln, die wir vom Transponieren von reellen Matrizen bzw. dem konjugierten Transponieren von komplexen Matrizen kennen, auch für die Adjungierte Abbildung zu nachzuweisen:

Satz 69. *Für die Adjungierte Abbildung gelten folgende Rechenregeln:*

$$(\varphi + \tau)^* = \varphi^* + \tau^*, \tag{A1}$$

$$(\lambda \varphi)^* = \bar{\lambda} \varphi^*, \tag{A2}$$

$$(\varphi^*)^* = \varphi, \tag{A3}$$

$$\text{id}^* = \text{id}, \tag{A4}$$

$$(\varphi \circ \tau)^* = \tau^* \circ \varphi^*, \tag{A5}$$

wobei die φ, τ lineare Abbildungen zwischen passenden Innenprodukträumen sind und λ entsprechend eine reelle oder komplexe Zahl.

Beweis.

zu (A1): Seien etwa $\varphi, \tau \in \text{hom}(V, W)$. Wähle $v \in V$ und $w \in W$ beliebig aber fest. Dann gilt

$$\begin{aligned} (v, (\varphi + \tau)^*(w)) &= ((\varphi + \tau)(v), w) \\ &= (\varphi(v), w) + (\tau(v), w) \\ &= (v, \varphi^*(w)) + (v, \tau^*(w)) \\ &= (v, \varphi^*(w) + \tau^*(w)). \end{aligned}$$

2 Das Innenprodukt kommt ins Spiel

Da \mathbf{v} zwar fest, aber beliebig war, gilt somit

$$(\varphi + \tau)^*(\mathbf{w}) = \varphi^*(\mathbf{w}) + \tau^*(\mathbf{w}),$$

und weil auch \mathbf{w} beliebig war, die Behauptung.

zu (A2)–(A5): Lassen sich analog herleiten (Übung). □

Bislang hatten wir nur die grundlegenden Eigenschaften des Innenprodukts selber verwendet. Betrachten wir Kern und Bild, so kommen plötzlich Orthogonale Komplemente ins Spiel:

Satz 70. Seien V und W Innenprodukträume und $\varphi \in \text{hom}(V, W)$. Dann gilt

1. $\ker \varphi^* = (\text{im } \varphi)^\perp$,
2. $\text{im } \varphi^* = (\ker \varphi)^\perp$,
3. $\ker \varphi = (\text{im } \varphi^*)^\perp$,
4. $\text{im } \varphi = (\ker \varphi^*)^\perp$.

Beweis. 1. Sei $\mathbf{w} \in W$. Dann gilt

$$\begin{aligned} \mathbf{w} \in \ker \varphi^* &\Leftrightarrow \varphi^*(\mathbf{w}) = \mathbf{0} \\ &\Leftrightarrow (\mathbf{v}, \varphi^*(\mathbf{w})) = 0 \quad \forall \mathbf{v} \in V \\ &\Leftrightarrow (\varphi(\mathbf{v}), \mathbf{w}) = 0 \quad \forall \mathbf{v} \in V \\ &\Leftrightarrow \mathbf{w} \in (\text{im } \varphi)^\perp. \end{aligned}$$

2. Folgt aus der vierten Eigenschaft, wenn man φ durch φ^* ersetzt, und verwendet, dass $(\varphi^*)^* = \varphi$.
 3. Ersetzt man in der ersten Eigenschaft φ durch φ^* , so ergibt sich die dritte.
 4. Die letzte ergibt sich, wenn man bei der ersten links und rechts das orthogonale Komplement nimmt.
- Damit kommen wir von der bewiesenen ersten Eigenschaft zu jeder anderen. □

Nun würden wir gerne wieder zu *Darstellungsmatrizen* zurückkehren. Dafür gehen wir wieder von Innenprodukträumen V und W aus und schauen, welchen Gewinn uns dabei Orthonormalbasen bringen. Insbesondere hätten wir gerne, dass sich damit die Matrixdarstellung der adjungierten φ^* direkt aus der von φ selber ergibt, so wie wir es in Beispiel 66 gesehen hatten. Dort war es ja die konjugiert Transponierte (auch Hermitesche genannt) der Matrix. Schön wäre es, wenn dies auch im allgemeinen Fall so wäre.

Dazu gehen wir von Orthonormalbasen B bestehend aus $\mathbf{q}_1, \dots, \mathbf{q}_m$ in V und C bestehend aus $\mathbf{u}_1, \dots, \mathbf{u}_n$ in W aus.

Nun bekommen wir die k -te Spalte von $\mathbf{A} := \mathbf{A}_{\varphi;B,C}$, indem wir $\varphi(\mathbf{q}_k)$ als Linearkombination über C schreiben. Die Entwicklungskoeffizienten sind dann gerade (in der Reihenfolge) die Einträge der Spalte. Da aber C eine ONB ist, kennen wir diese:

$$\varphi(\mathbf{q}_k) = \sum_{j=1}^n (\varphi(\mathbf{q}_k), \mathbf{u}_j) \mathbf{u}_j .$$

Damit ist

$$a_{jk} = (\varphi(\mathbf{q}_k), \mathbf{u}_j) .$$

Treiben wir das selbe Spiel für $\mathbf{A}^* := \mathbf{A}_{\varphi^*;C,B}$, so tauschen sich die Rollen entsprechend durch, und wir finden in der j -ten Spalte

$$a_{kj}^* = (\varphi^*(\mathbf{u}_j), \mathbf{q}_k) = (\mathbf{u}_j, \varphi(\mathbf{q}_k)) = \overline{(\varphi(\mathbf{q}_k), \mathbf{u}_j)} = \overline{a_{jk}} .$$

Das heißt aber gerade, dass es tatsächlich unabhängig vom gewählten Innenprodukt ist: Es ist wie beim Standardskalarprodukt die konjugiert Transponierte. Das halten wir fest:

Satz 71. Seien V und W Innenprodukträume und $\varphi \in \text{hom}(V, W)$. Ist B eine ONB von V und C eine ONB von W , so gilt für die Darstellungsmatrix $\mathbf{A} := \mathbf{A}_{\varphi;B,C}$ von φ bzw. $\mathbf{A}^* := \mathbf{A}_{\varphi^*;C,B}$ von deren Adjungierter φ^* bezüglich dieser Orthonormalbasen

$$\mathbf{A}^* = \overline{\mathbf{A}^T} \quad (= \mathbf{A}^H) . \quad (2.20)$$

Der Wechsel von einem Innenprodukt zum anderen äußert sich also quasi simultan für die Abbildung und ihre Adjungierte. Es ist allerdings zu beachten, dass bei einem anderen Innenprodukt die Adjungierte Abbildung eine andere ist. Insbesondere sind die Orthonormalbasen dann auch andere. Man darf sich also durch Satz 71 nicht zum Leichtsinn verleiten lassen.

Nun betrachten wir einen Sonderfall:

Definition 72. Ein Endomorphismus $\varphi \in \text{hom}(V, V)$ auf einem Innenproduktraum V heißt *selbstadjungiert*, wenn er mit seiner Adjungierten übereinstimmt, also

$$(\varphi(\mathbf{v}), \mathbf{w}) = (\mathbf{v}, \varphi(\mathbf{w})) \quad \forall \mathbf{v}, \mathbf{w} \in V .$$

Damit gilt dann automatisch:

Korollar 73 (zu Satz 71). Für die Darstellungsmatrix eines selbstadjungierten Endomorphismus bezüglich einer Orthonormalbasis gilt

$$\mathbf{A}^* = \mathbf{A} .$$

Bemerkung 74.

2 Das Innenprodukt kommt ins Spiel

Um nicht ständig zwischen reellem und komplexem Fall unterscheiden zu müssen, also zwischen $A^* = A^T$ und $A^* = \overline{A^T} (= A^H)$, verwenden wir im Folgenden vorzugsweise die Schreibweise mit dem Sternchen.

Im Komplexen Fall nennt man die selbstadjungierten Abbildungen wegen $A = A^H$ auch hermitesch und überträgt das dann auch gerne auf die Abbildungsmatrizen.

Eine schöne Eigenschaft von selbstadjungierten Abbildungen ist die folgende:

Satz 75. *Die Eigenwerte einer selbstadjungierten Abbildung sind immer reell.*

Beweis. Das sieht man leicht an folgender Rechnung, wobei λ ein Eigenwert ist und \mathbf{v} ein zugehöriger Eigenvektor:

$$\lambda(\mathbf{v}, \mathbf{v}) = (\lambda\mathbf{v}, \mathbf{v}) = (\varphi(\mathbf{v}), \mathbf{v}) = (\mathbf{v}, \varphi(\mathbf{v})) = (\mathbf{v}, \lambda\mathbf{v}) = \overline{\lambda}(\mathbf{v}, \mathbf{v}).$$

Damit ist aber $\overline{\lambda} = \lambda$ uns somit reell. □

Bemerkung 76. Damit ist eine selbstadjungierte Abbildung aber auch diagonalisierbar:

Sie besitzt dementsprechend eine Schursche Normalform \mathbf{D} mit reellen Diagonaleinträgen. Da diese Darstellung aber bezüglich einer ONB ist, ist $\mathbf{D}^* = \mathbf{D}$, hat also selber obere Dreiecksform. Damit kann es aber außerhalb der Diagonalen nur Nullen geben, \mathbf{D} ist also bereits die Diagonalisierung.

Schön wäre es, wenn wir noch nachweisen können, dass bei selbstadjungierten reellen Matrizen dann auch eine reelle ONB ausreicht.

Dass die Orthogonalität im Komplexen eine deutlich stärkere Eigenschaft ist als im Reellen, sieht man an folgendem

Satz 77. *Sei V ein komplexer Innenproduktraum und φ ein Endomorphismus auf V . Gilt*

$$(\varphi(\mathbf{v}), \mathbf{v}) = 0 \quad \forall \mathbf{v} \in V, \tag{2.21}$$

dann ist φ bereits die Nullabbildung.

Beweis. Wir können im Komplexen den Ausdruck $(\varphi(\mathbf{u}), \mathbf{w})$ umschreiben zu

$$\begin{aligned} (\varphi(\mathbf{u}), \mathbf{w}) = & \frac{(\varphi(\mathbf{u} + \mathbf{w}), \mathbf{u} + \mathbf{w}) - (\varphi(\mathbf{u} - \mathbf{w}), \mathbf{u} - \mathbf{w})}{4} \\ & + i \frac{(\varphi(\mathbf{u} + i\mathbf{w}), \mathbf{u} + i\mathbf{w}) - (\varphi(\mathbf{u} - i\mathbf{w}), \mathbf{u} - i\mathbf{w})}{4}. \end{aligned}$$

Nun haben wir nur noch Skalarprodukte der Form $(\varphi(\mathbf{v}), \mathbf{v})$. Weil diese nach Voraussetzung (2.21) alle verschwinden, müssen auch alle Skalarprodukte der Form $(\varphi(\mathbf{u}), \mathbf{w})$ verschwinden. Das geht aber nur, wenn φ bereits die Nullabbildung ist. □

Bemerkung 78. Im Reellen klappt das übrigens nicht. Z. B. entspricht im \mathbb{R}^2 die Rotation um einen rechten Winkel im Gegenuhrzeigersinn der Multiplikation mit der Matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Als Ergebnis steht dann jeder Vektor senkrecht auf seinem Bild, Gleichung (2.21) ist also erfüllt, obwohl eine Drehung weit davon entfernt ist, die Nullabbildung zu sein.

Man ahnt es auch nach dem Beweis des Satzes: Das künstliche Umschreiben von $(\varphi(\mathbf{u}), \mathbf{w})$ funktioniert im Reellen so nicht. Bei obiger Matrix steht dann auch im Komplexen das Bild nicht senkrecht auf dem Urbild.

Wenn wir den Endomorphismus als selbstadjungiert voraussetzen, klappt es auch im Reellen:

Satz 79. Sei φ ein selbstadjungierter Endomorphismus auf einem Innenproduktraum V und

$$(\varphi(\mathbf{v}), \mathbf{v}) = 0 \quad \forall \mathbf{v} \in V,$$

dann ist φ bereits die Nullabbildung.

Beweis. Für selbstadjungierte Endomorphismen klappt die Umschreiberei auch im Reellen

$$(\varphi(\mathbf{u}), \mathbf{w}) = \frac{(\varphi(\mathbf{u} + \mathbf{w}), \mathbf{u} + \mathbf{w}) - (\varphi(\mathbf{u} - \mathbf{w}), \mathbf{u} - \mathbf{w})}{4}.$$

Wir konnten uns den Imaginärteil sparen. Ab hier ist alles wieder wie beim Beweis von Satz 77. \square

Nun ist selbstadjungiert zu sein eine recht starke Eigenschaft, bedeutet sie doch im Reellen, dass die Abbildungsmatrix symmetrisch zu sein hat. Wir suchen nach geeigneten Abschwächungen, für die sich möglichst viele der Resultate noch retten lassen. Am nützlichsten erwiesen hat sich die folgende:

Definition 80. Ein Endomorphismus auf einem Innenproduktraum, der mit seiner Adjungierten kommutiert, also $\varphi \circ \varphi^* = \varphi^* \circ \varphi$, heißt *normal*. Man spricht auch von einem *normalen Operator*.

Haben wir also mit A eine Darstellung eines Endomorphismus bezüglich einer ONB, so ist er genau dann normal, wenn $AA^* = A^*A$.

Beispiel 81. Ist die Darstellung bzgl. einer ONB, in \mathbb{R}^2 etwa der Standardeinheitsbasis,

$$A = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix},$$

so ist die Abbildung offenbar nicht selbstadjungiert, weil A nicht symmetrisch ist. Wir finden aber

$$\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix},$$

Damit muss der zugrundeliegende Endomorphismus normal sein.

2 Das Innenprodukt kommt ins Spiel

Nun sammeln wir die angenehmen Eigenschaften normaler Endomorphismen ein. Zunächst:

Satz 82. Ein Endomorphismus $\varphi \in \text{hom}(V, V)$ ist genau dann normal, wenn

$$\|\varphi(\mathbf{v})\| = \|\varphi^*(\mathbf{v})\| \quad \forall \mathbf{v} \in V .$$

Beweis. Offenbar ist $\varphi^* \circ \varphi - \varphi \circ \varphi^*$ selbstadjungiert. Mit Satz 79 gilt dann

$$\begin{aligned} \varphi \text{ normal} &\Leftrightarrow \varphi^* \circ \varphi - \varphi \circ \varphi^* = 0 \\ &\Leftrightarrow ((\varphi^* \circ \varphi - \varphi \circ \varphi^*)(\mathbf{v}), \mathbf{v}) = 0 \quad \forall \mathbf{v} \in V \\ &\Leftrightarrow (\varphi^*(\varphi(\mathbf{v})) - \varphi(\varphi^*(\mathbf{v})), \mathbf{v}) = 0 \quad \forall \mathbf{v} \in V \\ &\Leftrightarrow (\varphi(\mathbf{v}), \varphi(\mathbf{v})) = (\varphi^*(\mathbf{v}), \varphi^*(\mathbf{v})) \quad \forall \mathbf{v} \in V \\ &\Leftrightarrow \|\varphi(\mathbf{v})\|^2 = \|\varphi^*(\mathbf{v})\|^2 \quad \forall \mathbf{v} \in V , \end{aligned}$$

und wir sind fertig. □

In der ELA hatten wir in Satz 126 gezeigt, dass diagonalisierbare Matrizen, wenn sie vertauschbar sind, die selben Eigenvektoren haben. Nun ist eine normale Matrix per Definition mit ihrer Adjungierten vertauschbar. Hätten wir bereits die Diagonalisierbarkeit gezeigt, wäre das Folgende einfach ein Korollar dazu. Nun zeigen wir es als Korollar zu Satz 82 und erhalten noch eine Aussage über die Eigenwerte.

Korollar 83. Ist $\varphi \in \text{hom}(V, V)$ ein normaler Endomorphismus auf einem Innenproduktraum und $\mathbf{v} \in V$ ein Eigenvektor von φ zum Eigenwert λ , dann ist es auch ein Eigenvektor der Adjungierten φ^* , allerdings zum Eigenwert $\bar{\lambda}$.

Im Reellen sind also auch die Eigenwerte die gleichen.

Beweis. Offensichtlich ist mit φ auch $\varphi - \lambda \text{id}$ normal, und es gilt $(\varphi - \lambda \text{id})^* = \varphi^* - \bar{\lambda} \text{id}$. Nach Satz 82 ist damit

$$0 = \|(\varphi - \lambda \text{id})(\mathbf{v})\| = \|(\varphi - \lambda \text{id})^*(\mathbf{v})\| = \|(\varphi^* - \bar{\lambda} \text{id})(\mathbf{v})\| .$$

□

Bis jetzt wussten wir, dass Eigenvektoren zu verschiedenen Eigenwerten l. u. sind. Bei normalen Matrizen geht noch mehr:

Satz 84. Sei φ ein normaler Operator auf einem Innenproduktraum. Dann stehen die Eigenvektoren zu verschiedenen Eigenwerten orthogonal aufeinander.

Beweis. Seien $\alpha \neq \beta$ Eigenwerte von φ mit zugehörigen Eigenvektoren \mathbf{u} und \mathbf{v} . Wir verwenden, dass mit Korollar 83 aus $\varphi(\mathbf{v}) = \beta\mathbf{v}$ folgt, dass $\varphi^*(\mathbf{v}) = \bar{\beta}\mathbf{v}$ und erhalten

$$\begin{aligned}(\alpha - \beta)(\mathbf{u}, \mathbf{v}) &= \alpha(\mathbf{u}, \mathbf{v}) - \beta(\mathbf{u}, \mathbf{v}) \\ &= (\alpha\mathbf{u}, \mathbf{v}) - (\mathbf{u}, \bar{\beta}\mathbf{v}) \\ &= (\varphi(\mathbf{u}), \mathbf{v}) - (\mathbf{u}, \varphi^*(\mathbf{v})) \\ &= (\varphi(\mathbf{u}), \mathbf{v}) - (\varphi(\mathbf{u}), \mathbf{v}) \\ &= 0.\end{aligned}$$

Da $\alpha - \beta$ nach Voraussetzung nicht verschwindet, muss das Innenprodukt (\mathbf{u}, \mathbf{v}) verschwinden, was ja gerade die Behauptung war. \square

Bemerkung 85. Die bisherigen Aussagen, die wir über normale Operatoren getroffen haben sowie ein Großteil der Ergebnisse für selbstadjungierte Operatoren konnten wir zeigen, ohne über eine Basis/ONB oder die Dimension des Raums nachzudenken. Diese werden sich später auf allgemeine Innenprodukträume übertragen, wo die Dimension auch unendlich sein darf.

Gehen wir zurück zur Matrix aus Beispiel 81

$$A = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$$

und betrachten den normalen Endomorphismus $\varphi : \mathbb{C}^2 \rightarrow \mathbb{C}^2$; $\mathbf{x} \mapsto A\mathbf{x}$. Dann kennen wir aus Abschnitt 7.1. der ELA bereits eine Basis aus Eigenvektoren, nämlich $(1, i)^T, (1, -i)^T$ mit den zugehörigen Eigenwerten $2 \mp 3i$. Offenbar ist die Basis bezüglich des Standardskalarprodukts bereits orthogonal, muss also, um eine ONB zu erhalten, nur noch normiert werden, indem man $\sqrt{2}/2$ davor multipliziert. Bezüglich dieser Basis ist dann aber die Darstellungsmatrix von φ diagonal, nämlich

$$\begin{pmatrix} 2-3i & 0 \\ 0 & 2+3i \end{pmatrix} = \sqrt{2}/2 \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \cdot A \cdot \sqrt{2}/2 \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}.$$

Wir haben also nicht nur reelle Eigenwerte, aber es gibt eine ONB des \mathbb{C}^2 bezüglich der φ durch eine Diagonalmatrix dargestellt wird.

Dies versuchen wir nun (im Komplexen) für alle normalen Operatoren zu verallgemeinern. Zu beachten ist, dass wir hier mit Orthonormalbasen arbeiten, weshalb dieses Ergebnis zunächst auf unsere vertrauten Unterräume des \mathbb{C}^N beschränkt bleibt.

Satz 86 (Komplexer Spektralsatz). *Sei φ ein Endomorphismus auf einem komplexen Innenproduktraum V . Dann sind äquivalent*

- (a) φ ist normal.
- (b) V besitzt eine ONB aus Eigenvektoren von φ .

2 Das Innenprodukt kommt ins Spiel

(c) φ lässt sich bezüglich einer ONB von V durch eine Diagonalmatrix darstellen.

Beweis. Versuche zunächst aus (c) Bedingung (a) zu folgern. Wir wissen, dass dann $A_{\varphi^*} = \overline{A_{\varphi}^T}$ ist, also ebenfalls eine Diagonalmatrix. Da Diagonalmatrizen immer miteinander kommutieren (also vertauschbar sind), gilt dies auch hier und dementsprechend für die zugrundeliegenden Operatoren. Damit kommutieren φ und φ^* , und sind dementsprechend normal.

Nun versuchen wir von (a) wieder zurück zu (c) zu kommen. Nach dem Satz über die Schursche Normalform 46 bzw. 47 gibt es eine ONB $B = \mathbf{q}_1, \dots, \mathbf{q}_n$ von V bezüglich der φ durch eine obere Dreiecksmatrix dargestellt wird, also

$$A := A_{\varphi, B} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix}. \quad (2.22)$$

Nun würden wir gerne zeigen, dass diese bereits eine Diagonalmatrix sein muss. Dazu verwenden wir Satz 82, nach dem $\|\varphi(\mathbf{v})\| = \|\varphi^*(\mathbf{v})\| \forall \mathbf{v} \in V$. Wir beginnen mit dem ersten Vektor aus der ONB:

$$\begin{aligned} \|\varphi(\mathbf{q}_1)\|^2 &= \|\mathbf{A}\mathbf{e}_1\|^2 = |a_{11}|^2, \\ \|\varphi^*(\mathbf{q}_1)\|^2 &= \|\overline{\mathbf{A}^T}\mathbf{e}_1\|^2 = \sum_{j=1}^n |a_{1j}|^2. \end{aligned}$$

Wenn diese beiden gleich sein sollen, muss $a_{12} = \dots = a_{1n} = 0$ sein. Setzen wir das so in (2.22) ein, setzt sich das Spiel fort:

$$\begin{aligned} \|\varphi(\mathbf{q}_2)\|^2 &= \|\mathbf{A}\mathbf{e}_2\|^2 = |a_{22}|^2, \\ \|\varphi^*(\mathbf{q}_2)\|^2 &= \|\overline{\mathbf{A}^T}\mathbf{e}_2\|^2 = \sum_{j=2}^n |a_{2j}|^2. \end{aligned}$$

Also muss auch gelten $a_{23} = \dots = a_{2n} = 0$. Damit lässt sich zukzessive zeigen, dass alle Außerdiagonaleinträge von A verschwinden, und wir sind fertig.

Da die Äquivalenz von (b) und (c) geschenkt ist, haben wir es also geschafft. \square

Nun wollen wir uns allmählich dem reellen Fall annähern. Im Eindimensionalen haben wir bereits das Problem, dass sich nicht alles mit reellen Zahlen machen lässt. So lässt sich nicht jedes reelle Polynom in (reelle) Linearfaktoren, also solche der Gestalt $(x - \lambda)$, zerlegen. Immerhin gilt:

Lemma 87. Jedes reelle Polynom lässt sich in eindeutiger Weise (bis auf die Reihenfolge der Faktoren) faktorisieren als

$$p(x) = \tilde{c} \cdot \left[\prod_{k=1}^m (x - \lambda_k) \right] \cdot \left[\prod_{j=1}^M (x^2 + b_j x + c_j) \right], \quad (2.23)$$

wobei alle vorkommenden Koeffizienten reell sind, und für jedes j gilt $b_j^2 < 4c_j$.

Beweisskizze. Für \tilde{c} nehmen wir den Koeffizienten vor der höchsten Potenz, etwa n , von x . Im Komplexen ließe sich das Polynom faktorisieren zu

$$p(x) = \tilde{c} \cdot \prod_{k=1}^n (x - \lambda_k). \quad (2.24)$$

Die Nullstellen λ_k , die nicht reell sind, kommen in komplexen konjugierten Paaren vor. Für ein solches ist dann

$$(x - \lambda)(x - \bar{\lambda}) = x^2 - 2\Re(\lambda)x + |\lambda|^2.$$

Nun ist aber für $\lambda \notin \mathbb{R}$

$$4|\lambda|^2 = 4(\Re(\lambda))^2 + 4(\Im(\lambda))^2 > 4(\Re(\lambda))^2 = (2\Re(\lambda))^2.$$

Sortieren wir die Faktoren in (2.24) so passend um und fassen die komplex konjugierten Paare wie gesehen zusammen, haben wir die Behauptung. \square

Für Paare komplex konjugierter Nullstellen haben wir also Faktoren der Form $x^2 + bx + c$ mit $b^2 < 4c$. In dem Fall ist offensichtlich auch

$$x^2 + bx + c = \left(x + \frac{b}{2}\right)^2 + \left(c - \frac{b^2}{4}\right) > 0. \quad (2.25)$$

Damit ist der Ausdruck aber auch für jedes $x \in R$ invertierbar. Nun würden wir gerne wissen, was passiert, wenn wir – in der ELA hatten wir sowas ähnliches schon gemacht – statt x einen Operator φ einsetzen. Damit wir die Invertierbarkeit unter der Bedingung $b^2 < 4c$ auch für Operatoren anstelle von x zeigen können, müssen wir uns auf selbstadjungierte Operatoren beschränken:

Lemma 88. *Ist φ eine selbstadjungierter Endomorphismus auf einem reellen Innenprodukttraum V und $b, c \in \mathbb{R}$ mit $b^2 < 4c$, dann ist der Endomorphismus*

$$\varphi^2 + b \cdot \varphi + c \cdot \text{id} \quad (\text{mit } \varphi^2 := \varphi \circ \varphi) \quad (2.26)$$

invertierbar, also bijektiv (und damit ein Isomorphismus).

Beweis. Wir wenden den Operator (2.26) auf einen nicht verschwindenden Vektor $\mathbf{v} \in V$ an und bilden dann das Innenprodukt mit diesem. Dabei benutzen wir, dass φ selbstadjungiert ist, sowie die Cauchy-Schwarzsche Ungleichung:

$$\begin{aligned} ((\varphi^2 + b \cdot \varphi + c \cdot \text{id})(\mathbf{v}), \mathbf{v}) &= (\varphi(\varphi(\mathbf{v})), \mathbf{v}) + b(\varphi(\mathbf{v}), \mathbf{v}) + c(\mathbf{v}, \mathbf{v}) \\ &= (\varphi(\mathbf{v}), \varphi(\mathbf{v})) + b(\varphi(\mathbf{v}), \mathbf{v}) + c\|\mathbf{v}\|^2 \\ &\geq \|\varphi(\mathbf{v})\|^2 - |b|\|\varphi(\mathbf{v})\|\|\mathbf{v}\| + c\|\mathbf{v}\|^2 \\ &= \left(\|\varphi(\mathbf{v})\| - \frac{|b|\|\mathbf{v}\|}{2}\right)^2 + \left(c - \frac{b^2}{4}\right)\|\mathbf{v}\|^2 \\ &> 0, \end{aligned}$$

2 Das Innenprodukt kommt ins Spiel

wobei wir im letzten Schritt die Voraussetzung an die Koeffizienten verwendet haben. Damit kann $(\varphi^2 + b \cdot \varphi + c \cdot \text{id})(\mathbf{v})$ nicht der Nullvektor sein. Aus der ELA wissen wir, dass damit die Behauptung folgt. \square

Im Komplexen wissen wir, dass jedes Polynom und damit auch das charakteristische Polynom eines Endomorphismus mindestens eine Nullstelle hat. Das hatten wir beim Beweis der Trigonalisierung in Satz 11 benutzt. Nun versuchen wir zu beweisen, dass das im Reellen wenigstens für selbstadjungierte Operatoren gilt.

Satz 89. *Jeder selbstadjungierte Operator $\varphi \in \text{hom}(V, V)$ auf einem nichttrivialen Innenproduktraum hat mindestens einen Eigenwert.*

Beweis. Da im Komplexen eh alles klar ist, beschränken wir uns im weiteren auf reelle Innenprodukträume.

Sei $n = \dim V$ und $\mathbf{v} \in V \setminus \{\mathbf{0}\}$. Dann können die Vektoren

$$\mathbf{v}, \varphi(\mathbf{v}), \varphi^2(\mathbf{v}), \dots, \varphi^n(\mathbf{v})$$

nicht l. u. sein, weil sie einfach einer zu viel sind. Wir können also den Nullvektor als nichttriviale Linearkombinationen von diesen darstellen:

$$\mathbf{0} = \alpha_0 \mathbf{v} + \alpha_1 \varphi(\mathbf{v}) + \alpha_2 \varphi^2(\mathbf{v}) + \dots + \alpha_n \varphi^n(\mathbf{v}) =: p(\varphi)(\mathbf{v})$$

mit dem Polynom $p(x) = \sum_{j=0}^n \alpha_j x^j$. Dieses lässt sich aber nach Lemma 87 faktorisieren zu

$$p(x) = \tilde{c} \cdot \left[\prod_{j=1}^m (x - \lambda_j) \right] \cdot \left[\prod_{j=1}^M (x^2 + b_1 x + c_j) \right],$$

wobei alle vorkommenden Koeffizienten reell sind, $\tilde{c} \neq 0$, und für jedes j gilt $b_j^2 < 4c_j$. Setzen wir nun wieder, d. h. in diese faktorisierte Form, φ ein und wenden es auf unser \mathbf{v} an, so kann aber nach Lemma 88 der Anteil der nichtreellen Nullstellen, also der $(x^2 + b_1 x + c_j)$ nicht Schuld daran sein, dass der Nullvektor rauskommt. Damit ist aber bereits

$$\prod_{j=1}^m (\varphi - \lambda_j \text{id})(\mathbf{v}) = \mathbf{0},$$

weshalb mindestens einer der Faktoren $(\varphi - \lambda_j \text{id})$ nicht injektiv sein kann, und das zugehörige λ_j entsprechend ein Eigenwert sein muss. \square

Bemerkung 90.

- Die Selbstadjungiertheit hatten wir gebraucht, um Lemma 88 in Anschlag bringen zu können.
- Der Beweis benutzt die Dimension des Raums, wird sich also nicht so leicht auf unendlichdimensionale Vektorräume verallgemeinern lassen.

Noch ein kleiner Baustein:

Lemma 91. *Ist $\varphi \in \text{hom}(V, V)$ selbstadjungiert und der Unterraum U von V invariant unter φ , dann gelten*

1. *Auch U^\perp ist invariant unter φ ,*
2. *$\varphi|_U$ ist als Endomorphismus auf U , selbstadjungiert,*
3. *$\varphi|_{U^\perp}$ ist als Endomorphismus auf U^\perp selbstadjungiert.*

Beweis.

1. Sei $\mathbf{u} \in U$ und $\mathbf{v} \in U^\perp$. Weil φ selbstadjungiert ist und U invariant unter φ gilt

$$(\varphi(\mathbf{v}), \mathbf{u}) = (\mathbf{v}, \varphi(\mathbf{u})) = 0.$$

Damit ist auch $\varphi(\mathbf{v}) \in U^\perp$.

2. Seien nun beide Vektoren $\mathbf{u}, \mathbf{v} \in U$. Dann ist

$$(\varphi|_U(\mathbf{u}), \mathbf{v}) = (\varphi(\mathbf{u}), \mathbf{v}) = (\mathbf{u}, \varphi(\mathbf{v})) = (\mathbf{u}, \varphi|_U(\mathbf{v})).$$

Die Selbstadjungiertheit vererbt sich also auf natürliche Weise, weil es ein invarianter Unterraum ist.

3. Folgt direkt aus den beiden vorigen Punkten.

□

Mit diesen Bausteinen haben wir endlich zusammen, was wir wollen: Die (reell) orthogonale Diagonalisierbarkeit selbstadjungierter Operatoren auf reellen Innenprodukträumen:

Satz 92 (Reeller Spektralsatz). *Sei φ ein Endomorphismus eines reellen Innenproduktraums V . Dann sind äquivalent:*

- (a) *φ ist selbstadjungiert.*
- (b) *V besitzt eine ONB aus Eigenvektoren von φ .*
- (c) *Es gibt eine ONB in V bezüglich der φ durch eine Diagonalmatrix dargestellt wird.*

Beweis. Die Äquivalenz von (b) und (c) folgt direkt aus unseren Erkenntnissen aus der ELA.

Um von (a) nach (b) zu kommen, erinnern wir uns an die Beweisstrategie aus dem Beweis zur Existenz einer Trigonalisierung im Komplexen, insbesondere dem Beweis zu Satz 11: Wir machen eine Induktion über die Dimension n des Raums.

Der Induktionsanfang ist einfach, da für $n = 1$ alles trivialerweise gilt.

2 Das Innenprodukt kommt ins Spiel

Nun betrachten wir $n > 1$ und setzen voraus, dass für jeden echten Unterraum bereits gilt, dass Bedingung (a) Bedingung (b) impliziert. Aus (a) folgt mit Satz 89, dass φ mindestens einen Eigenwert hat. Zu diesem nehmen wir uns einen Eigenvektor \mathbf{u} mit Norm Eins. Nenne dessen Erzeugnis $\langle \mathbf{u} \rangle = U$. Dann ist aber nach Lemma 91 die Einschränkung $\varphi|_{U^\perp} \in \text{hom}(U^\perp, U^\perp)$ selbstadjungiert. Damit besitzt nach Induktionsvoraussetzung U^\perp eine ONB aus Eigenvektoren von $\varphi|_{U^\perp}$. Ergänzen wir diese durch \mathbf{u} zu einer ONB von V , haben wir wieder (b). Dabei haben wir verwendet, dass U^\perp unter φ invariant ist, weshalb die Eigenvektoren von $\varphi|_{U^\perp}$ auch welche von φ sind.

Dass aus (c) wieder (a) folgt sieht man daran, dass eine Diagonalmatrix gleich ihrer Transponierten ist, also bzgl. der angesetzten ONB $A_\varphi = A_\varphi^*$. Damit hat also φ^* bezüglich unserer ONB die selbe Darstellung wie φ , muss also mit φ übereinstimmen. \square

Bemerkung 93. Dass wir den Beweis nicht über Darstellungsmatrizen geführt haben, hat uns einerseits unabhängiger vom jeweiligen Innenprodukt gemacht, andererseits die Induktion deutlich erleichtert.

2.6 Positive Operatoren und Isometrien und was man lustiges damit anstellen kann

In diesem Abschnitt soll es darum gehen, wie man Endomorphismen auf Innenprodukträumen in die Hintereinanderausführung eines positiven Operators und einer Isometrie aufspalten kann. Die Idee stammt aus dem Reich der komplexen Zahlen:

$$z = \frac{z}{|z|} |z| = \left(\frac{z}{|z|} \right) \cdot \sqrt{\bar{z}z}. \quad (2.27)$$

Hier haben wir also eine Zahl $z \in \mathbb{C} \setminus \{0\}$ geschrieben als Produkt einer Zahl vom Betrag Eins, nämlich $z/|z|$, mit einer positiven (und damit reellen) Zahl, nämlich $|z| = \sqrt{\bar{z}z}$, einer sogenannten *Polarzerlegung*. Unser Wunsch wäre also, sowas auch für Endomorphismen bzw. quadratische Matrizen zu bekommen. Praktischerweise ist die Entsprechung zu $\bar{z}z$, nämlich $\varphi^* \circ \varphi$ automatisch selbstadjungiert, was ja unsere nächste Entsprechung zu den reellen Zahlen ist. Das äußert sich auch im reellen Spektralsatz. Wir müssen also eine passende Definition für die *Positivität von Endomorphismen* finden und eine für deren *Quadratwurzel*. Außerdem brauchen wir eine Entsprechung zu Zahlen vom Betrag Eins, also *Isometrien*, die den Betrag bzw. die Norm nicht verändern.

Beispiel 94. Zur Illustration des Erkenntnisgewinns, den so eine Aufspaltung liefert, betrachten wir reelle Matrizen der Form

$$\begin{pmatrix} c & -s \\ s & c \end{pmatrix} \quad \text{mit } c = \cos(\theta), s = \sin(\theta),$$

für einen Winkel θ . Diese entsprechen der komplexen Zahl $c + is$ und damit einer Drehung um den Winkel θ , also insbesondere einer Isometrie. Führen wir davor eine Streckung in

2.6 Positive Operatoren und Isometrien und was man lustiges damit anstellen kann

Richtung der Koordinatenachsen durch, etwa durch die Matrix

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix},$$

so lautet das Ergebnis für $\lambda = \sqrt{2}, \mu = \sqrt{2}/2, \theta = \pi/4$

$$\begin{pmatrix} 1 & -1/2 \\ 1 & 1/2 \end{pmatrix}.$$

Dieser Matrix hätte man die Eigenschaft als Komposition von Streckung und Drehung nicht direkt angesehen.

Es gibt natürlich noch andere Isometrien als Drehungen. Schon in der Ebene kennen wir noch die Spiegelung an einer Achse. Aber es gibt noch viel mehr.

Definition 95. Ein Endomorphismus φ auf einem Innenproduktraum V heißt *positiv*, wenn er selbstadjungiert ist, und

$$(\varphi(\mathbf{v}), \mathbf{v}) \geq 0 \quad \forall \mathbf{v} \in V. \quad (2.28)$$

Bemerkung 96.

- Die Bezeichnung *positiv* ist etwas dämlich, weil wir die Null nicht ausgeschlossen haben, hat sich aber so eingebürgert. Genau genommen korrespondieren positive Operatoren nicht mit dem reellen Intervall $(0, \infty)$, sondern mit $[0, \infty)$.
- Im Komplexen lässt sich zeigen, dass aus $(\varphi(\mathbf{v}), \mathbf{v}) \in \mathbb{R} \quad \forall \mathbf{v} \in V$ bereits folgt, dass φ selbstadjungiert ist, sodass man es nicht extra voraussetzen müsste.

Beispiel 97.

1. In Beispiel 16 hatten wir uns ein Innenprodukt auf \mathbb{C}^n bzw. \mathbb{R} gebastelt mittels Diagonalmatrizen $\text{diag}(c_1, \dots, c_n)$ mit Einträgen $c_1, \dots, c_n > 0$. Diese sind offensichtlich positiv bezüglich des Standardskalarprodukts.
2. Ist U ein Unterraum von V , so ist die Orthogonalprojektion P_U positiv (Übung).
3. Bei genauer Betrachtung des Beweises von Lemma 88 stellen wir fest, dass auf reellen Innenprodukträumen Abbildungen der Form $\varphi^2 + b \cdot \varphi + c \cdot \text{id}$ mit $b^2 < 4c$ und selbstadjungiertem φ bereits positiv sind.
4. Betrachten wir den \mathbb{C}^n bzw. \mathbb{R}^n mit dem Standardskalarprodukt sowie eine entsprechende $n \times n$ -Matrix A , dann ist (etwa im Komplexen Fall)

$$\begin{aligned} \varphi : \mathbb{C}^n &\rightarrow \mathbb{C}^n \\ \mathbf{x} &\mapsto A A^* \mathbf{x} \end{aligned}$$

ein positiver Operator. (Übung)

2 Das Innenprodukt kommt ins Spiel

5. Für einen Endomorphismus φ auf einem Innenproduktraum gilt

$$(\varphi^*(\varphi(\mathbf{v})), \mathbf{v}) = (\varphi(\mathbf{v}), \varphi(\mathbf{v})) \geq 0, \quad (2.29)$$

womit wie oben gewünscht $\varphi^* \circ \varphi$ positiv ist.

Definition 98. Ein Operator τ heißt eine *Wurzel* bzw. *Quadratwurzel* eines Endomorphismus φ , wenn $\tau^2 = \varphi$.

Beispiel 99.

1. Da für die Orthogonalprojektion gilt $P_U^2 = P_U$, ist sie ihre eigene Quadratwurzel, genauer: eine der selben, und zwar nach Beispiel 97 eine positive Quadratwurzel.
2. Ist $\varphi : \mathbb{R}^3 \rightarrow \mathbb{R}^3; \mathbf{x} \mapsto (x_3, 0, 0)^T$, so ist $\tau : \mathbb{R}^3 \rightarrow \mathbb{R}^3; \mathbf{x} \mapsto (x_2, x_3, 0)^T$ eine Quadratwurzel von φ . Wie sich das in den Darstellungsmatrizen bzgl. der Standardeinheitsbasis äußert, können Sie sich als Übung überlegen.

Wollten wir nun die nichtnegativen Zahlen $x \in [0, \infty) \subseteq \mathbb{C}$ als Teilmenge der komplexen Zahlen charakterisieren, so kämen wir auf folgende äquivalente Bedingungen:

- (a) $z \in [0, \infty)$.
- (b) $z = \bar{z}$ und $\Re(z) \geq 0$.
- (c) z hat eine nichtnegative Quadratwurzel, d. h., es gibt ein $x \in [0, \infty)$ mit $x^2 = z$.
- (d) Es gibt ein $w = \bar{w} \in \mathbb{R}$ mit $w^2 = z$.
- (e) Es gibt ein $\eta \in \mathbb{C}$ mit $\bar{\eta}\eta = z$.

Da komplexe Zahlen komplexen 1×1 -Matrizen entsprechen und damit Endomorphismen auf \mathbb{C}^1 mit dem Standardskalarprodukt, versuchen wir nun eine Entsprechung für allgemeine Innenprodukträume zu formulieren und zu beweisen.

Satz 100. Sei V ein Innenproduktraum und $\varphi \in \text{hom}(V, V)$. Dann sind äquivalent:

- (a) φ ist positiv.
- (b) φ ist selbstadjungiert, und alle Eigenwerte liegen in $[0, \infty)$.
- (c) φ hat eine positive Quadratwurzel.
- (d) φ hat eine selbstadjungierte Quadratwurzel.
- (e) Es gibt ein $\tau \in \text{hom}(V, V)$ mit $\tau^* \circ \tau = \varphi$.

Beweis. Wir machen einen Ringschluss von oben nach unten und wieder von (e) direkt zu (a):

2.6 Positive Operatoren und Isometrien und was man lustiges damit anstellen kann

(a)⇒(b): Da eine positive Abbildung per definitionem selbstadjungiert ist, müssen wir uns nur noch die Eigenwerte anschauen. Sei also für einen nichtverschwindenden Vektor $\varphi(\mathbf{v}) = \lambda \mathbf{v}$. Dann ist

$$0 \leq (\varphi(\mathbf{v}), \mathbf{v}) = (\lambda \mathbf{v}, \mathbf{v}) = \lambda(\mathbf{v}, \mathbf{v}) = \lambda \|\mathbf{v}\|^2.$$

Das geht aber nur für $\lambda \geq 0$.

(b)⇒(c): Jetzt haben wir also eine selbstadjungierte Abbildung mit nichtnegativen Eigenwerten. Die ist aber nach dem Spektralsatz 92 bzw. 86 orthogonal diagonalisierbar, d. h. es gibt eine ONB $\mathbf{q}_1, \dots, \mathbf{q}_n$ in V aus Eigenvektoren von φ . Nennen wir die zugehörigen Eigenwerte, die ja nach Voraussetzung (b) alle nichtnegativ sind, $\lambda_1, \dots, \lambda_n$, so können wir leicht eine positive Quadratwurzel definieren: Verlange auf der ONB, dass

$$\tau(\mathbf{q}_j) = \sqrt{\lambda_j} \mathbf{q}_j \quad \forall j.$$

Nach Satz 66 aus der ELA ist damit eindeutig ein Endomorphismus gegeben. Wie man leicht sieht, ist

$$\tau(\tau(\mathbf{q}_j)) = \lambda_j \mathbf{q}_j = \varphi(\mathbf{q}_j) \quad \forall j$$

und unter erneuter Anwendung von Satz 66 aus der ELA τ als Quadratwurzel von φ nachgewiesen.

Der Nachweis der Positivität bleibt Ihnen als Übungsaufgabe überlassen.

(c)⇒(d): Dieses ist der einfachste Teil, weil er direkt aus der Definition der Positivität folgt.

(d)⇒(e): Übung.

(e)⇒(a): Ist $\varphi = \tau^* \circ \tau$, so gilt

$$\varphi^* = (\tau^* \circ \tau)^* = \tau^* \circ (\tau^*)^* = \tau^* \circ \tau = \varphi$$

und außerdem

$$(\varphi(\mathbf{v}), \mathbf{v}) = (\tau^*(\tau(\mathbf{v})), \mathbf{v}) = (\tau(\mathbf{v}), \tau(\mathbf{v})) \geq 0 \quad \forall \mathbf{v} \in V.$$

Damit haben wir auch das geschafft.

□

Satz 101. Jeder positive Operator auf einem Innenproduktraum V hat genau eine positive Quadratwurzel.

2 Das Innenprodukt kommt ins Spiel

Beweis. Dass er mindestens eine hat, haben wir gerade eben in Satz 100 gezeigt. Jetzt müssen wir nur noch nachweisen, dass jede positive Quadratwurzel mit der aus obigem Beweis übereinstimmt.

Wir nehmen also einen positiven Operator τ mit $\tau^2 = \varphi$. Dann gibt es nach Satz 100 eine ONB $\mathbf{u}_1, \dots, \mathbf{u}_n$ aus Eigenvektoren von τ mit zugehörigen nichtnegativen reellen Eigenwerten μ_1, \dots, μ_n . Damit ist für jedes in Frage kommende j

$$\varphi(\mathbf{u}_j) = \tau(\tau(\mathbf{u}_j)) = \mu_j^2 \mathbf{u}_j,$$

also μ_j^2 ein Eigenwert von φ zum Eigenvektor \mathbf{u}_j , der damit auch einer von φ ist. Nach eventuellem Umnummerieren finden wir also, dass $\mu_j^2 = \lambda_j \forall j$. Damit ist klar, dass gelten muss

$$\mu_j = \sqrt{\lambda_j} \quad \forall j = 1, \dots, n.$$

Da die $\mathbf{u}_1, \dots, \mathbf{u}_n$ eine Basis bilden, liefert wiederum Satz 66 aus der ELA die Behauptung, d. h. die Eindeutigkeit von τ . \square

Definition 102. Die soeben nachgewiesene eindeutige positive Quadratwurzel eines positiven Operators φ bezeichnen wir mit $\sqrt{\varphi}$.

Insbesondere ist für jeden Endomorphismus φ auf einem Innenproduktraum der Ausdruck $\sqrt{\varphi^* \circ \varphi}$ nach den Sätzen 100 und 101 wohldefiniert und bezeichnet einen positiven Operator.

Nun kommt nach den positiven Operatoren der zweite neue zentrale Begriff, nämlich die Bezeichnung für Endomorphismen, welche die Norm eines Vektors nicht verändern:

Definition 103. Ein Operator $S \in \text{hom}(V, V)$ auf einem Innenproduktraum heißt *Isometrie*, wenn

$$\|S(\mathbf{v})\| = \|\mathbf{v}\| \quad \forall \mathbf{v} \in V. \quad (2.30)$$

Man könnte auch sagen, ein Isomorphismus ist *normerhaltend*.

Beispiel 104. Sind $\lambda_1, \dots, \lambda_n$ Zahlen vom Betrag Eins (egal ob im Reellen oder Komplexen) und $\mathbf{q}_1, \dots, \mathbf{q}_n$ eine ONB, so definiere eine lineare Abbildung durch

$$S(\mathbf{q}_j) = \lambda_j \mathbf{q}_j \quad \forall j.$$

Ist nun $\mathbf{v} \in V$ ein beliebiger Vektor, so können wir ihn ganz leicht als Linearkombination über unserer ONB schreiben:

$$\mathbf{v} = \sum_{j=1}^n (\mathbf{v}, \mathbf{q}_j) \mathbf{q}_j$$

und finden für seine Norm

$$\|\mathbf{v}\|^2 = \sum_{j=1}^n |(\mathbf{v}, \mathbf{q}_j)|^2.$$

2.6 Positive Operatoren und Isometrien und was man lustiges damit anstellen kann

Nun ist wegen der Linearität

$$\begin{aligned} S(\mathbf{v}) &= \sum_{j=1}^n (\mathbf{v}, \mathbf{q}_j) S(\mathbf{q}_j) \\ &= \sum_{j=1}^n (\mathbf{v}, \mathbf{q}_j) \lambda_j \mathbf{q}_j . \end{aligned}$$

Da alle unsere λ_j Betrag Eins haben, gilt dann

$$\begin{aligned} \|S(\mathbf{v})\|^2 &= \sum_{j=1}^n |(\mathbf{v}, \mathbf{q}_j)|^2 |\lambda_j|^2 \\ &= \sum_{j=1}^n |(\mathbf{v}, \mathbf{q}_j)|^2 \\ &= \|\mathbf{v}\|^2 , \end{aligned}$$

womit S als Isometrie nachgewiesen ist.

Dass es im Komplexen keine anderen Isometrien gibt, werden wir noch zeigen. Im Reellen ist die Welt nicht ganz so einfach, weil dort nicht jede Isometrie diagonalisierbar ist.

Daher sammeln wir zunächst Charakterisierungen von Isometrien, die da wie dort gelten:

Satz 105. Für einen Endomorphismus $S \in \text{hom}(V, V)$ auf einem Innenproduktraum sind äquivalent

- (a) S ist eine Isometrie.
- (b) Für alle $\mathbf{u}, \mathbf{v} \in V$ ist $(S(\mathbf{u}), S(\mathbf{v})) = (\mathbf{u}, \mathbf{v})$.
- (c) Bilden die Vektoren $\mathbf{q}_1, \dots, \mathbf{q}_k$ ein Orthonormalsystem, so gilt dies auch für $S(\mathbf{q}_1), \dots, S(\mathbf{q}_k)$.
- (d) Es gibt eine ONB $\mathbf{q}_1, \dots, \mathbf{q}_n$ in V , für die auch $S(\mathbf{q}_1), \dots, S(\mathbf{q}_n)$ eine ONB ist.
- (e) $S^* \circ S = \text{id}$.
- (f) $S \circ S^* = \text{id}$.
- (g) S^* ist eine Isometrie.
- (h) S ist invertierbar mit $S^{-1} = S^*$.

Beweis. Wir machen wieder einen klassischen Ringschluss:

(a) \Rightarrow (b): Sei also S eine Isometrie. Wie wir in der Übung gezeigt haben, lässt sich das Innenprodukt wieder durch die zugehörige Norm ausdrücken (Blatt 2, P1 und H4). Damit kann eine Isometrie auch keine Skalarprodukte verändern. \checkmark

2 Das Innenprodukt kommt ins Spiel

(b)⇒(c): Eine Abbildung, die keine Innenprodukte verändert, verändert auch nicht die Orthogonalität. Ist nämlich $\mathbf{u} \perp \mathbf{v}$, so ist

$$0 = (\mathbf{u}, \mathbf{v}) = (S(\mathbf{u}), S(\mathbf{v})),$$

also auch $S(\mathbf{u}) \perp S(\mathbf{v})$. Da auch die Norm nicht verändert wird, werden automatisch Orthonormalsysteme auf Orthonormalsysteme abgebildet. ✓

(c)⇒(d): Und damit werden natürlich erst recht Orthonormalbasen auf Orthonormalbasen abgebildet. ✓

(d)⇒(e): Sei nun $\mathbf{q}_1, \dots, \mathbf{q}_n$ eine ONB in V , für die auch $S(\mathbf{q}_1), \dots, S(\mathbf{q}_n)$ eine ONB ist. Dann ist für jedes verfügbare Indexpaar j, k

$$(S^*(S(\mathbf{q}_j)), \mathbf{q}_k) = (S(\mathbf{q}_j), S(\mathbf{q}_k)) = \delta_{jk} = (\mathbf{q}_j, \mathbf{q}_k).$$

Da die \mathbf{q}_j eine Basis bilden, gilt also bei gegebenem k für jedes $\mathbf{v} \in V$

$$(S^*(S(\mathbf{v})), \mathbf{q}_k) = (\mathbf{v}, \mathbf{q}_k).$$

Analog sehen wir, dass für alle $\mathbf{u}, \mathbf{v} \in V$

$$(S^*(S(\mathbf{v})), \mathbf{u}) = (\mathbf{v}, \mathbf{u}).$$

Das geht aber nur, wenn $S^* \circ S$ bereits die Identität ist. ✓

(e)⇒(f): Folgt aus unseren Erkenntnissen in der ELA. ✓

(f)⇒(g): Mit $S \circ S^* = \text{id}$ gilt dann für jedes $\mathbf{v} \in V$

$$\|S^*(\mathbf{v})\|^2 = (S^*(\mathbf{v}), S^*(\mathbf{v})) = (S(S^*(\mathbf{v})), \mathbf{v}) = (\mathbf{v}, \mathbf{v}) = \|\mathbf{v}\|^2. \quad \checkmark$$

(g)⇒(h): Da wir es geschafft hatten, von (a) nach (e) bzw. (f) zu kommen, können wir, das selbe auch für S^* tun. Da $(S^*)^* = S$, sind wir dann aber auch schon fertig. ✓

(h)⇒(a): Wir müssen jetzt also zeigen, dass eine Abbildung mit $S^* \circ S = S \circ S^* = \text{id}$ bereits eine Isometrie ist. Das machen wir wie bei der Implikation von (f) nach (g), nur mit den Rollen von S und der Adjungierten S^* vertauscht:

$$\|S(\mathbf{v})\|^2 = (S(\mathbf{v}), S(\mathbf{v})) = (S^*(S(\mathbf{v})), \mathbf{v}) = (\mathbf{v}, \mathbf{v}) = \|\mathbf{v}\|^2. \quad \checkmark$$

Damit sind wir einmal komplett im Kreis rum und somit fertig. □

Wir sehen leicht das folgende

Korollar 106. *Jede Isometrie ist normal*

2.6 Positive Operatoren und Isometrien und was man lustiges damit anstellen kann

Beweis. Wir konnten zeigen, dass für jede Isometrie $\text{id} = S^* \circ S = S \circ S^*$. □

Damit kriegen wir dann auch die schöne Charakterisierung im Komplexen:

Satz 107. Für einen Endomorphismus $S \in \text{hom}(V, V)$ auf einem komplexen Innenproduktraum sind äquivalent:

- S ist eine Isometrie.
- Es gibt eine ONB $\mathbf{q}_1, \dots, \mathbf{q}_n$ aus Eigenvektoren von S , und die zugehörigen Eigenwerte $\lambda_1, \dots, \lambda_n$ haben alle Betrag Eins.

Beweis. Da nach Korollar 106 eine Isometrie normal ist, können wir uns auf den komplexen Spektralsatz 86 berufen, nach dem es die gewünschte ONB aus Eigenvektoren von S gibt. Betrachten wir nun einen solchen Eigenvektor, etwa \mathbf{q}_j , dann gilt, weil S eine Isometrie ist

$$1 = \|\mathbf{q}_j\| = \|S(\mathbf{q}_j)\| = \|\lambda_j \mathbf{q}_j\| = |\lambda_j| \|\mathbf{q}_j\| = |\lambda_j|. \quad \checkmark$$

Die andere Richtung des Beweises, ist gerade unser Beispiel 104, sodass wir bereits fertig sind. □

Im Reellen klappt das nicht ganz. Während bei der komplexen Spektralzerlegung ein konjugiert komplexes Paar $a \pm ib$ einem Kästchen der Form $\begin{pmatrix} a+ib & 0 \\ 0 & a-ib \end{pmatrix}$ entspricht, müssen wir uns im Reellen dafür mit einem Kästchen der Form $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ begnügen, das offenbar nicht diagonal ist. Es ist überhaupt nicht reell diagonalisierbar. Hat eine komplexe Zahl den Betrag Eins, so hat sie die Gestalt $\cos(\theta) \pm i \sin(\theta)$ mit einem passenden Winkel θ . Das entspricht im Reellen einem Kästchen $\begin{pmatrix} c & -s \\ s & c \end{pmatrix}$, wobei wie in Beispiel 94 die Buchstaben c und s für den Cosinus und den Sinus von θ stehen. Daher ist es naheliegend, dass wir nur das folgende bekommen

Satz 108. Für einen Endomorphismus $S \in \text{hom}(V, V)$ auf einem reellen Innenproduktraum sind äquivalent:

- S ist eine Isometrie.
- Es gibt eine ONB $\mathbf{q}_1, \dots, \mathbf{q}_n$ von V bezüglich der S sich durch eine Blockdiagonalmatrix darstellen lässt, bei der jeder reelle Eigenwert durch einen 1×1 -Block mit einer 1 oder -1 dargestellt wird, die anderen durch einen 2×2 -Block der Gestalt

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad \text{mit } \theta \in [0, \pi).$$

Beweis. Siehe Literatur. □

2 Das Innenprodukt kommt ins Spiel

Nun haben wir genügend Information über positive Operatoren und Isometrien zusammengetragen, um unsere Polarzerlegung eines Operators φ zu basteln. Die sollte ja die Verallgemeinerung der Polarzerlegung

$$z = \left(\frac{z}{|z|} \right) \cdot \sqrt{\bar{z}z}$$

einer komplexen Zahl sein und dementsprechend mit einer passenden Isometrie S die Gestalt $\varphi = S \circ \sqrt{\varphi^* \circ \varphi}$ besitzen. Das klappt tatsächlich. Um den Beweis übersichtlicher zu machen, ziehen wir einen Teil davon in zwei Lemmata vorne raus:

Lemma 109. *Sei V ein Innenproduktraum und $\varphi \in \text{hom}(V, V)$. Dann haben φ und $\sqrt{\varphi^* \circ \varphi}$ gleichen Rang.*

Beweis. Nach Satz 70 haben in einem endlichdimensionalen Innenproduktraum φ und φ^* den gleichen Rang, und der Kern der einen ist das orthogonale Komplement des Bilds der anderen. Dementsprechend haben auch $\varphi^* \circ \varphi$ und damit auch $\sqrt{\varphi^* \circ \varphi}$ den selben Rang. \square

Lemma 110. *Sei V ein Innenproduktraum, U ein Unterraum von V und $\mathbf{q}_1, \dots, \mathbf{q}_n$ eine ONB von U . Dann gilt für $S \in \text{hom}(V, V)$:*

1. Ist sowohl $S_1 := S|_U$ als auch $S_2 := S|_{U^\perp}$ eine Isometrie, so ist auch S eine Isometrie.
2. Gilt $\|S(\mathbf{q}_j)\| = 1$ für jedes j , so ist S bereits eine Isometrie.

Beweis.

1. Da sich mit Satz 54 jedes \mathbf{v} eindeutig zerlegen lässt als $\mathbf{v} = \mathbf{u} + \mathbf{w}$ mit $\mathbf{u} \in U$ und $\mathbf{w} \in U^\perp$, gilt mit Satz 105 $S(\mathbf{u}) \perp S(\mathbf{w})$ und nach Pythagoras

$$\begin{aligned} \|S(\mathbf{v})\|^2 &= \|S_1(\mathbf{u}) + S_2(\mathbf{w})\|^2 = \|S_1(\mathbf{u})\|^2 + \|S_2(\mathbf{w})\|^2 \\ &= \|\mathbf{u}\|^2 + \|\mathbf{w}\|^2 = \|\mathbf{u} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2. \quad \checkmark \end{aligned}$$

2. Folgt analog, wenn man schreibt

$$V = \bigoplus_{j=1}^n \langle \mathbf{q}_j \rangle$$

und Pythagoras wiederholt anwendet. \square

Nun kommt der ersehnte

Satz 111 (Polarzerlegung). *Zu jedem Endomorphismus $\varphi \in \text{hom}(V, V)$ auf einem Innenproduktraum gibt es eine Isometrie $S \in \text{hom}(V, V)$, sodass*

$$\varphi = S \circ \sqrt{\varphi^* \circ \varphi}. \quad (2.31)$$

2.6 Positive Operatoren und Isometrien und was man lustiges damit anstellen kann

Beweis. Nach den Sätzen 100 und 101 ist $\sqrt{\varphi^* \circ \varphi}$ ein wohldefinierter positiver Operator. Nun versuchen wir rauszufinden, in welchem Verhältnis die Normen von φ und $\sqrt{\varphi^* \circ \varphi}$ stehen. Für ein beliebiges $\mathbf{v} \in V$ gilt

$$\begin{aligned} \|\varphi(\mathbf{v})\|^2 &= (\varphi(\mathbf{v}), \varphi(\mathbf{v})) \\ &= ((\varphi^* \circ \varphi)(\mathbf{v}), \mathbf{v}) \\ &= ((\sqrt{\varphi^* \circ \varphi} \circ \sqrt{\varphi^* \circ \varphi})(\mathbf{v}), \mathbf{v}) \\ &= ((\sqrt{\varphi^* \circ \varphi})(\mathbf{v}), (\sqrt{\varphi^* \circ \varphi})(\mathbf{v})) \\ &= \|(\sqrt{\varphi^* \circ \varphi})(\mathbf{v})\|^2, \end{aligned}$$

also

$$\|\varphi(\mathbf{v})\| = \|(\sqrt{\varphi^* \circ \varphi})(\mathbf{v})\|. \quad (2.32)$$

Wenn es also eine Lineare Abbildung S mit

$$\varphi(\mathbf{v}) = (S \circ \sqrt{\varphi^* \circ \varphi})(\mathbf{v}) \quad \forall \mathbf{v} \in V$$

gibt, muss mindestens ihre Einschränkung auf das Bild von $\sqrt{\varphi^* \circ \varphi}$ eine Isometrie sein.

Schritt 1: Bastle eine Isometrie $S_1 : \text{im}(\sqrt{\varphi^* \circ \varphi}) \rightarrow \text{im}(\varphi)$ mit

$$S_1((\sqrt{\varphi^* \circ \varphi})(\mathbf{v})) = \varphi(\mathbf{v}) \quad \forall \mathbf{v} \in V. \quad (2.33)$$

Wir probieren, ob durch (2.33) bereits eine lineare Abbildung definiert ist. Dazu wählen wir uns $\mathbf{v}_1, \mathbf{v}_2 \in V$ mit $(\sqrt{\varphi^* \circ \varphi})(\mathbf{v}_1) = (\sqrt{\varphi^* \circ \varphi})(\mathbf{v}_2)$ und verwenden Gleichung (2.32):

$$\begin{aligned} \|\varphi(\mathbf{v}_1) - \varphi(\mathbf{v}_2)\| &= \|\varphi(\mathbf{v}_1 - \mathbf{v}_2)\| \\ &= \|(\sqrt{\varphi^* \circ \varphi})(\mathbf{v}_1 - \mathbf{v}_2)\| \\ &= \|(\sqrt{\varphi^* \circ \varphi})(\mathbf{v}_1) - (\sqrt{\varphi^* \circ \varphi})(\mathbf{v}_2)\| \\ &= \|\mathbf{0}\| = 0. \end{aligned}$$

Also muss bereits gelten $\varphi(\mathbf{v}_1) = \varphi(\mathbf{v}_2)$, womit Gleichung (2.33) bereits eine lineare Abbildung definiert, die wegen (2.32) auf ihrem Definitionsbereich, nämlich dem Bild von $\sqrt{\varphi^* \circ \varphi}$ bereits eine Isometrie ist. Damit ist sie natürlich auch bijektiv, weil sie ja nur den Nullvektor im Kern haben kann.

Wäre φ bijektiv, so wäre nach Satz 70 auch φ^* und damit $\sqrt{\varphi^* \circ \varphi}$ bijektiv, und wir wären fertig, weil S_1 schon auf ganz V definiert wäre. Nun hatten wir φ nicht als bijektiv vorausgesetzt und müssen noch schauen, wie wir uns im anderen Fall helfen:

Schritt 2: Bastle eine Isometrie S auf V mit $S|_{\text{im}(\sqrt{\varphi^* \circ \varphi})} = S_1$.

Nach Lemma 109 haben $\text{im}(\sqrt{\varphi^* \circ \varphi})$ und $\text{im}(\varphi)$ gleiche Dimension, etwa k . Nun wählen wir uns in $\text{im}(\sqrt{\varphi^* \circ \varphi})$ eine ONB $\mathbf{q}_1, \dots, \mathbf{q}_k$ und in $\text{im}(\varphi)$ entsprechend eine

2 Das Innenprodukt kommt ins Spiel

ONB $\mathbf{u}_1, \dots, \mathbf{u}_k$. Ergänze nun beide Orthonormalbasen durch ONBs der jeweiligen orthogonalen Komplemente zu Orthonormalbasen $\mathbf{q}_1, \dots, \mathbf{q}_n$ bzw. $\mathbf{u}_1, \dots, \mathbf{u}_n$ von ganz V . Verlangen wir nun

$$S(\mathbf{q}_j) = \begin{cases} S_1(\mathbf{q}_j) & \text{für } j = 1, \dots, k, \\ \mathbf{u}_j & \text{für } j = k + 1, \dots, n, \end{cases} \quad (2.34)$$

so haben wir eine lineare Abbildung, deren Einschränkungen auf $\text{im}(\sqrt{\varphi^* \circ \varphi})$ gerade S_1 ist, und deren Einschränkung S_2 auf $\text{im}(\sqrt{\varphi^* \circ \varphi})^\perp$ ebenfalls eine Isometrie ist (Teil (d) von Satz 105). Nach Lemma 110 ist dann auch S bereits eine Isometrie.

Weiter ist für beliebige Vektoren $\mathbf{v} \in V$

$$S((\sqrt{\varphi^* \circ \varphi})(\mathbf{v})) = S_1((\sqrt{\varphi^* \circ \varphi})(\mathbf{v})) = \varphi(\mathbf{v}),$$

womit wir an unserem Ziel angelangt sind. □

Eine Nutzenanwendung der Polarzerlegung ist die äußerst angenehme *Singulärwertzerlegung*.

- Bei der Polarzerlegung hatten wir zwei Faktoren gefunden, die besondere Eigenschaften haben. Insbesondere ist $\sqrt{\varphi^* \circ \varphi}$ positiv und somit *orthogonal diagonalisierbar mit nichtnegativen Eigenwerten*.
- Dass die nachgeschaltete Abbildung – die in der Formel deshalb davor steht – eine *Isometrie ist, die selber wieder orthogonal diagonalisierbar ist*, lässt uns auf mehr hoffen. Z. B., dass es neben der ONB $B = \mathbf{q}_1, \dots, \mathbf{q}_n$ bzgl. der $\sqrt{\varphi^* \circ \varphi}$ durch eine Diagonalmatrix dargestellt wird, noch eine weitere ONB $C = \mathbf{u}_1, \dots, \mathbf{u}_n$ gibt, die dafür sorgt, dass die Abbildungsmatrix $A_{\varphi; B, C}$ Diagonalgestalt hat mit den Eigenwerten von $\sqrt{\varphi^* \circ \varphi}$ auf der Diagonalen.
- Mit dem reellen Spektralsatz 92 ist *im reellen Fall* die Basis B reell. Außerdem wissen wir, dass dann S eine *reelle Isometrie* ist. Damit können wir auch erwarten, dass die ONB C auch wieder reell ist.

Weil die Eigenwerte von $\sqrt{\varphi^* \circ \varphi}$ im allgemeinen nicht mit denen von φ selber übereinstimmen, bekommen sie einen eigenen Namen:

Definition 112 (Singulärwerte: Version 1). Ist $\varphi \in \text{hom}(V, V)$ ein Endomorphismus auf einem Innenproduktraum, so heißen die Eigenwerte von $\sqrt{\varphi^* \circ \varphi}$ die *Singulärwerte* von φ .

Eine Eigenschaft der Singulärwerte haben wir eigentlich schon gezeigt, nämlich in den Beweisen der Sätze 100 und 101. Dort müssen wir nur jeweils φ durch $\varphi^* \circ \varphi$ ersetzen:

2.6 Positive Operatoren und Isometrien und was man lustiges damit anstellen kann

Satz 113. Die Singulärwerte von φ sind gerade die (nichtnegativen) Quadratwurzeln der Eigenwerte von $\varphi^* \circ \varphi$.

Das erlaubt die Berechnung der Singulärwerte bzw. deren Approximation (siehe Numerik II), ohne die Wurzel $\sqrt{\varphi^* \circ \varphi}$ ausrechnen zu müssen.

Nun beweisen wir, dass das mit der Singulärwertzerlegung tatsächlich klappt. Zunächst:

Satz 114. Sind $\sigma_1, \dots, \sigma_n$ die Singulärwerte unseres Operators φ , dann gibt es in unserem Innenproduktraum V Orthonormalbasen $B = \mathbf{q}_1, \dots, \mathbf{q}_n$ und $C = \mathbf{u}_1, \dots, \mathbf{u}_n$, sodass für jedes $\mathbf{v} \in V$ gilt

$$\varphi(\mathbf{v}) = \sum_{j=1}^n \sigma_j(\mathbf{v}, \mathbf{q}_j) \mathbf{u}_j. \quad (2.35)$$

Beweis. Nach dem Spektralsatz existiert eine ONB $B = \mathbf{q}_1, \dots, \mathbf{q}_n$, sodass

$$(\sqrt{\varphi^* \circ \varphi})(\mathbf{q}_j) = \sigma_j \mathbf{q}_j \quad \forall j.$$

Über dieser ONB lässt sich nun ein Vektor $\mathbf{v} \in V$ schreiben als

$$\mathbf{v} = \sum_{j=1}^n (\mathbf{v}, \mathbf{q}_j) \mathbf{q}_j.$$

Wenden wir $\sqrt{\varphi^* \circ \varphi}$ darauf an, so finden wir

$$(\sqrt{\varphi^* \circ \varphi})(\mathbf{v}) = \sum_{j=1}^n \sigma_j(\mathbf{v}, \mathbf{q}_j) \mathbf{q}_j.$$

Jetzt wenden wir noch die Isometrie S aus unserer Polarzerlegung darauf an. Damit ist dann wegen der Linearität

$$\varphi(\mathbf{v}) = S((\sqrt{\varphi^* \circ \varphi})(\mathbf{v})) = \sum_{j=1}^n \sigma_j(\mathbf{v}, \mathbf{q}_j) S(\mathbf{q}_j). \quad (2.36)$$

Da aber Isometrien Orthonormalbasen auf Orthonormalbasen abbilden (Satz 105), bekommen wir mit $\mathbf{u}_j := S(\mathbf{q}_j)$ unsere gewünschte Orthonormalbasis C , und Gleichung (2.36) wird zu

$$\varphi(\mathbf{v}) = \sum_{j=1}^n \sigma_j(\mathbf{v}, \mathbf{q}_j) \mathbf{u}_j.$$

Da \mathbf{v} beliebig war, sind wir fertig. (Und die \mathbf{u}_j sind im reellen Fall tatsächlich reell.) \square

Korollar 115. Mit den Bezeichnungen von Satz 114 gilt

$$\mathbf{A}_{\varphi; B, C} = \text{diag}(\sigma_1, \dots, \sigma_n). \quad (2.37)$$

2 Das Innenprodukt kommt ins Spiel

Beweis. Setzen wir in Gleichung (2.35) eines der \mathbf{q}_j ein, so ergibt sich

$$\varphi(\mathbf{q}_j) = \sigma_j \mathbf{u}_j. \quad (2.38)$$

Da dies für jedes verfügbare j gilt, folgt die Behauptung. \square

Wir können Gleichung (2.38) auch für eine alternative Definition der Singulärwerte verwenden:

Definition 116 (Singulärwerte: Version 2). Ist V ein Innenproduktraum, $\varphi \in \text{hom}(V, V)$, und existieren in V Orthonormalbasen $\mathbf{q}_1, \dots, \mathbf{q}_n$ und $\mathbf{u}_1, \dots, \mathbf{u}_n$, sodass mit nichtnegativen reellen Zahlen $\sigma_1, \dots, \sigma_n$ gilt

$$\varphi(\mathbf{q}_j) = \sigma_j \mathbf{u}_j \quad \forall j,$$

so heißen die σ_j die *Singulärwerte* von φ , die \mathbf{q}_j die entsprechenden *Rechts-Singulärvektoren* und die \mathbf{u}_j die *Links-Singulärvektoren* von φ .

Bemerkung 117.

- Während viele reelle Endomorphismen bzw. Matrizen noch nicht einmal reell diagonalisierbar sind, besitzen sie alle eine reelle Singulärwertzerlegung.
- Bei positiven Operatoren stimmen die Diagonalisierung und die Singulärwertzerlegung bereits überein. (Dort stimmt sie dann auch mit der Schur-Normalform überein.)
- Bei allgemeinen Selbstadjungierten Operatoren sind die Singulärwerte gerade die Beträge der Eigenwerte, und die Singulärvektoren lassen sich ebenfalls aus den Eigenvektoren generieren.
- Üblicherweise nummeriert man die Singulärwerte so, dass

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n.$$

Die größeren Singulärwerte haben auch größeren Anteil an der „Wirkung“ des Operators. Sie beschreiben via Gleichung (2.38) die Transformationen mit der größten Energie. (Anwendungen z. B. in der Physik der Atmosphäre.)

- Die Singulärwertzerlegung lässt sich auch auf Homomorphismen $\varphi \in \text{hom}(V, W)$ zwischen verschiedenen Innenprodukträumen verallgemeinern, insbesondere auf solche verschiedener Dimension.

2.7 Wie sich das Ganze auf Matrizen überträgt

Nachdem es uns die Beweise in Abschnitt 2.6 sehr erleichtert hat, dass wir nicht in Matrizen, sondern in Operatoren gedacht hatten, wollen wir aber am Ende dennoch wissen, wie sich das Ganze (wieder) auf Matrizen (zurück)übersetzen lässt. Wir wollen also wissen, was passiert, wenn wir unsere Vektoren und unseren Operator bereits bezüglich einer bestimmten Basis dargestellt haben.

2.7.1 Die Sache mit den orthogonalen und unitären Matrizen und weshalb das Standardskalarprodukt die Hauptrolle spielt

Zunächst halten wir fest, dass nach Satz 71 für die Matrixdarstellung von Operatoren und ihren Adjungierten über einer Orthonormalbasis gilt

$$A_{\varphi^*} = A_{\varphi}^* = \overline{A_{\varphi}^T} \quad (= A_{\varphi}^H), \quad (2.39)$$

weshalb (siehe Korollar 73) für die Darstellung eines selbstadjungierten Operators gilt:

$$A^* = A. \quad (2.40)$$

Dabei setzen wir natürlich immer voraus, dass wir den Operator und seine Adjungierte bezüglich der selben ONB darstellen. Damit bedeutet Satz 75, dass eine selbstadjungierte Matrix, im Reellen also eine symmetrische, nur reelle Eigenvektoren hat. Wir vergeben für die selbstadjungierten reellen Matrizen den (bereits bekannten) Begriff *symmetrische Matrizen* und für die komplexen selbstadjungierten den Begriff *hermitesche Matrizen*.

Für die Darstellungsmatrix eines normalen Operators gilt analog

$$A^*A = AA^*, \quad (2.41)$$

weshalb wir in diesem Fall auch von *normalen Matrizen* sprechen. Entsprechend Satz 82 ist also eine Matrix A genau dann normal, wenn

$$\|Ax\| = \|A^*x\| \quad \forall x \in \mathbb{R}^n \quad (\text{bzw. } \mathbb{C}^n). \quad (2.42)$$

Bemerkenswert ist, dass für die Darstellungsmatrix Q einer Isometrie über einer ONB nach Satz 105 gelten muss

$$Q^*Q = QQ^* = E$$

mit der Einheitsmatrix E . Das bedeutet aber, dass sowohl die Zeilen als auch die Spalten bezüglich des Standardskalarprodukts eine Orthonormalbasis bilden. Es verstecken sich in so einer Matrix im allgemeinen zwei ONBs des \mathbb{R}^n bzw. \mathbb{C}^n . Daher vergeben wir für solche Matrizen einen eigenen Namen:

Definition 118. Eine quadratische Matrix mit

$$Q^*Q = QQ^* = E \quad (2.43)$$

heißt im reellen Fall *orthogonal*, bzw. eine *Orthogonalmatrix*, im komplexen Fall *unitär*.

Bemerkung 119. Manchmal werden entsprechend reelle Isometrien *orthogonale Abbildungen* genannt, komplexe Isometrien entsprechend *unitäre Abbildungen*.

Dass bei der Darstellung einer Isometrie über einer ONB automatisch das Standardskalarprodukt ins Spiel kommt, weckt den Verdacht, dass dieses bei der Koordinatendarstellung bezüglich einer ONB allgemein eine besondere Rolle spielt. Tatsächlich gilt

2 Das Innenprodukt kommt ins Spiel

Satz 120. Sei V ein n -dimensionaler Innenproduktraum und $B = \mathbf{q}_1, \dots, \mathbf{q}_n$ eine ONB in V . Sind $\mathbf{u}, \mathbf{v} \in V$ mit

$$\mathbf{u} = \sum_{j=1}^n w_j \mathbf{q}_j, \quad \mathbf{v} = \sum_{j=1}^n z_j \mathbf{q}_j,$$

so gilt

$$(\mathbf{u}, \mathbf{v}) = \mathbf{w} \cdot \mathbf{z}$$

mit den Koordinatenvektoren $\mathbf{w} = (w_1, \dots, w_n)^T$ und $\mathbf{z} = (z_1, \dots, z_n)^T$ und dem Standardskalarprodukt \cdot .

Beweis. Wir setzen einfach ein:

$$(\mathbf{u}, \mathbf{v}) = \left(\sum_{j=1}^n w_j \mathbf{q}_j, \sum_{j=1}^n z_j \mathbf{q}_j \right).$$

Der Rest ergibt sich aus der Definition des Innenprodukts und der Tatsache, dass B eine Orthonormalbasis ist (siehe Übung). \square

Damit können wir uns im Folgenden auf das Standardskalarprodukt beschränken. Dies wird auch der Fall sein, wenn Sie etwa im Kurs Numerik I/II numerische lineare Algebra betreiben. Anders wird das Ganze, wenn Sie etwa numerische Quadraturformeln vom Gauß-Typ betrachten, bei denen Orthogonalpolynome ins Spiel kommen. Dort muss man sich tatsächlich mit den Innenprodukten von dem Typ befassen, wie es schon in der Übung für Polynome dran war (Gauß-Legendre Quadratur) bzw. Varianten davon. Auch bei sogenannten Finite-Element-Verfahren (FEM), etwa für Probleme der Strukturmechanik muss man sich mit speziellen Innenprodukten befassen, bevor man eine Matrixdarstellung seines Problems bekommt.

Die Sache mit den orthogonalen bzw. unitären Matrizen motiviert auch die folgende Spezialisierung der Definition für die Ähnlichkeit zweier Matrizen:

Definition 121.

- Zwei quadratische reelle Matrizen A und B heißen *orthogonal ähnlich*, wenn es eine Orthogonalmatrix Q gibt, sodass

$$B = Q^T A Q.$$

- Zwei quadratische komplexe Matrizen A und B heißen *unitär ähnlich*, wenn es eine unitäre Matrix U gibt, sodass

$$B = U^T A U.$$

Der komplexe Spektralsatz [86](#) wird damit zu

Satz 122 (komplexer Spektralsatz für Matrizen). *Jede normale komplexe Matrix ist unitär ähnlich zu einer Diagonalmatrix. Es gibt keine anderen komplexen Matrizen, für die das gilt.*

Der reelle Spektralsatz 92 lautet entsprechend

Satz 123 (reeller Spektralsatz für Matrizen). *Jede symmetrische reelle Matrix ist orthogonal ähnlich zu einer Diagonalmatrix. Es gibt keine anderen reellen Matrizen, für die das gilt.*

Auch der Satz von Schur 46 bzw. 47 lässt sich entsprechend formulieren:

Satz 124 (Satz von Schur für Matrizen). *Jede (reelle oder) komplexe quadratische Matrix ist unitär ähnlich zu einer oberen Dreiecksmatrix.*

Dabei werden reelle Matrizen einfach als komplexe Matrizen betrachtet.

In Satz 105 hatten wir ganz allgemein Isometrien durch äquivalente Bedingungen charakterisiert. Wie wir schon gesehen hatten, müssen wir hier bei Matrizen im reellen Fall von Orthogonalen Matrizen reden, im komplexen von Unitären Matrizen. Dabei hatten wir einen Teil der äquivalenten Bedingungen schon reingepackt. Den Rest fassen wir für den reellen und komplexen Fall getrennt zusammen. Zunächst reell:

Satz 125. *Für eine Matrix $Q \in \mathbb{R}^{n \times n}$ sind äquivalent*

- Q ist orthogonal.
- Q^T ist orthogonal.
- Für jedes Paar x, y von Vektoren in \mathbb{R}^n gilt

$$(Qx) \cdot (Qy) = x \cdot y .$$

- Ist $S \in \mathbb{R}^{n \times n}$ ebenfalls orthogonal, so sind auch SQ und QS orthogonal.

Im Komplexen können wir wegen Satz 107 noch eine weitere Bedingung anfügen:

Satz 126. *Für eine Matrix $U \in \mathbb{C}^{n \times n}$ sind äquivalent*

- U ist unitär.
- U^T ist unitär.
- Für jedes Paar w, z von Vektoren in \mathbb{C}^n gilt

$$(Uw) \cdot (Uz) = w \cdot z .$$

- Ist $S \in \mathbb{C}^{n \times n}$ ebenfalls unitär, so sind auch SU und US unitär.
- U ist unitär ähnlich zu einer Diagonalmatrix, deren Diagonaleinträge sämtlich Betrag Eins haben.

Anders ausgedrückt ist U unitär diagonalisierbar mit Eigenwerten vom Betrag Eins.

2.7.2 Positiv definite Matrizen

Bei den positiven Operatoren müssen wir bezüglich der Begriffe vorsichtig sein, denn der Begriff „positive Matrix“ ist bereits anderweitig belegt. Wir legen uns neue Begriffe zurecht:

Definition 127. Eine selbstadjungierte Matrix $A \in \mathbb{C}^{n \times n}$ bzw. $\in \mathbb{R}^{n \times n}$ heißt, wenn für jedes $\mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$ bzw. $\in \mathbb{R}^n \setminus \{\mathbf{0}\}$ gilt

$$\begin{aligned} (\mathbf{Ax}) \cdot \mathbf{x} &\geq 0, && \text{positiv semidefinit,} \\ (\mathbf{Ax}) \cdot \mathbf{x} &> 0, && \text{positiv definit,} \\ (\mathbf{Ax}) \cdot \mathbf{x} &\leq 0, && \text{negativ semidefinit,} \\ (\mathbf{Ax}) \cdot \mathbf{x} &< 0, && \text{negativ definit,} \\ &\text{ansonsten} && \text{indefinit.} \end{aligned}$$

Die Voraussetzung, dass A selbstadjungiert ist, sorgt wie im Fall der Operatoren dafür, dass das Skalarprodukt $(\mathbf{Ax}) \cdot \mathbf{x}$ tatsächlich reell ist. Den positiven Operatoren entsprechen also gerade die positiv semidefiniten Matrizen.

Offenbar gilt ganz allgemein

Lemma 128.

- Genau dann ist A positiv semidefinit, wenn $-A$ negativ semidefinit ist.
- Genau dann ist A positiv definit, wenn $-A$ negativ definit ist.
- Genau dann ist A indefinit, wenn $-A$ indefinit ist.

Die Übertragung von Satz 100 auf Matrizen lautet dann

Satz 129. Für eine Matrix $A \in \mathbb{C}^{n \times n}$ bzw. $\in \mathbb{R}^{n \times n}$ sind äquivalent:

- (a) A ist positiv semidefinit.
- (b) A ist selbstadjungiert, und alle Eigenwerte liegen in $[0, \infty)$.
- (c) Es gibt eine positiv semidefinite Matrix B mit $B^2 = A$.
- (d) Es gibt eine selbstadjungierte Matrix B mit $B^2 = A$.
- (e) Es gibt eine Matrix $M \in \mathbb{C}^{n \times n}$ bzw. $\in \mathbb{R}^{n \times n}$ mit $M^*M = A$.

Wir müssen nichts mehr beweisen, weil wir das für Satz 100 bereits getan haben. Aber wir können sofort eine erste Charakterisierung für positiv definite Matrizen daraus ableiten:

Korollar 130. Für eine Matrix $A \in \mathbb{C}^{n \times n}$ bzw. $\in \mathbb{R}^{n \times n}$ sind äquivalent:

- (a) A ist positiv definit.

- (b) A ist selbstadjungiert, und alle Eigenwerte sind positiv.
- (c) Es gibt eine positiv definite Matrix B mit $B^2 = A$.
- (d) Es gibt eine reguläre selbstadjungierte Matrix B mit $B^2 = A$.
- (e) Es gibt eine reguläre Matrix $M \in \mathbb{C}^{n \times n}$ bzw. $\in \mathbb{R}^{n \times n}$ mit $M^*M = A$.

Beweis. Es ist leicht, zu sehen, dass der Beweis von Satz 100 für den Fall, dass man statt „positiv“ so etwas wie „streng positiv“ mit „... $(\varphi(\mathbf{v}), \mathbf{v}) > 0 \quad \forall \mathbf{v} \in V \setminus \{\mathbf{0}\}$ “ statt „... $(\varphi(\mathbf{v}), \mathbf{v}) \geq 0 \quad \forall \mathbf{v} \in V$ “ benutzt hätte, die vorkommenden Operatoren einfach nur als bijektiv hätten angenommen werden müssen. Dann hätte der Beweis auch wieder funktioniert. (Übung) \square

Ein besonders interessantes Kriterium dafür, dass eine Matrix positiv definit ist, ist das folgende:

Satz 131. Für eine Matrix $A \in \mathbb{C}^{n \times n}$ bzw. $\in \mathbb{R}^{n \times n}$ sind äquivalent:

1. A ist positiv definit.
2. Es gibt eine reguläre rechte obere Dreiecksmatrix C mit $C^*C = A$.
3. Es gibt eine reguläre linke untere Dreiecksmatrix L mit $LL^* = A$.

Beweis.

- Dass die beiden letzten Bedingungen äquivalent sind, ist offensichtlich: man wähle $L = C^*$.
- Dass aus diesen die positive Definitheit folgt sehen wir, wenn wir in Korollar 130 die Matrix M durch C bzw. L^* ersetzen.
- Um zu zeigen, dass aus der positiven Definitheit von A die Existenz der Zerlegung im zweiten Punkt folgt, schreiben wir entsprechend Korollar 130 $A = M^*M$.

Laut Bemerkung 41 zur Gram-Schmidt-Orthogonalisierung können wir die reguläre Matrix M schreiben als

$$M = QR$$

mit einer (im reellen Fall) orthogonalen bzw. (im komplexen Fall) unitären Matrix Q und einer oberen Dreiecksmatrix R . Diese muss aber vollen Rang haben, also regulär sein, damit M selber regulär ist. Nun können wir schreiben

$$A = M^*M = (QR)^*QR = R^*Q^*QR = R^*R$$

und sind mit $C := R$ am Ziel. \square

2 Das Innenprodukt kommt ins Spiel

Bemerkung 132 (Cholesky-Zerlegung). Attraktiv an diesem Ergebnis ist, dass es uns eine Möglichkeit eröffnet, bei einer Matrix, von der wir es nicht aus anderen Gründen wissen, zu überprüfen, ob sie positiv definit ist:

- In Bemerkung 83 im ELA-Skript finden wir die Erkenntnis, dass man eine Matrix, für die der Gauß-Algorithmus ohne Zeilentausch funktioniert, faktorisieren kann zu $A = LR$, wobei R die rechte obere Dreiecksmatrix ist, die uns der Gauß-Algorithmus liefert und L eine linke untere Dreiecksmatrix mit Einsen auf der Diagonalen, die sich ansonsten aus den Zeilenoperationen ergibt, die man beim Gauß-Algorithmus angewandt hat. Der Beweis dafür beruht auf den *Elementarmatrizen*, die wir erst später in Abschnitt 6.2 eingeführt hatten. benutzt wird eigentlich nur der erste Fall aus dem dortigen Lemma 94.
- Da R regulär ist, können wir es schreiben als $R = D\tilde{R}$, wobei D eine Diagonalmatrix mit den Diagonaleinträgen von R ist, und \tilde{R} dementsprechend nur noch Einsen auf der Diagonalen hat.
- Wenn $A = LD\tilde{R}$ selbstadjungiert ist, muss dementsprechend gelten $\tilde{R} = L^*$, und D muss selbstadjungiert, also reell sein.
- Damit A positiv definit ist, muss dann D positiv definit sein, sonst gäbe es wegen der Regularität von L ein $x \in \mathbb{R}^n$ bzw. \mathbb{C}^n mit

$$(Ax) \cdot x = (D(L^*x)) \cdot (L^*x) \leq 0.$$

Daher müssen dann, wie man leicht sieht, alle Diagonaleinträge positiv sein, sodass wir die Matrix

$$\sqrt{D} := \text{diag}(\sqrt{d_{1,1}}, \dots, \sqrt{d_{n,n}})$$

bilden können und damit $C = \sqrt{D}L^*$. Damit ist dann wie gewünscht $A = C^*C$.

- Alternativ hätten wir auch L durch $L\sqrt{D}$ ersetzen können.
- Eine Möglichkeit zur Berechnung wäre etwa

```

for i=1 to n
  cii =  $\left( a_{ii} - \sum_{k=1}^{i-1} |c_{ki}|^2 \right)^{1/2}$  ;
  for k=i+1 to n
    cik =  $\left( a_{ik} - \sum_{j=1}^{i-1} \overline{c_{jk}} c_{ji} \right) / c_{ii}$  ;
  end;
end

```

- Funktioniert der so abgewandelte Gauß-Algorithmus, ohne dass man eine Null oder schlimmeres auf der Diagonalen bekommt, ist die Matrix A positiv definit, ansonsten eben nicht.
- Die so gewonnene Zerlegung heißt *Cholesky-Zerlegung*.

Bemerkung 133. Es gibt auch hinreichende Kriterien, die nicht notwendig sind. Eines davon geht auch wieder auf die Sache mit dem Gauß-Algorithmus zurück.

- Erfüllt eine $n \times n$ -Matrix A das *starke Zeilensummenkriterium*

$$|a_{j,j}| > \sum_{\substack{k=1 \\ k \neq j}} |a_{j,k}|, \quad (2.44)$$

so lässt sich leicht zeigen (Übung), dass diese Eigenschaft durch keinen Schritt der Gauß-Elimination zerstört wird. (Dabei wird die *Gauß-Elimination im strengen Sinn* verwendet, also nur von oben nach unten und nur mit Abziehen des Vielfachen einer Zeile von einer anderen.)

- Dabei wechseln die Diagonaleinträge der Matrix zwar ihren Betrag, nicht aber ihr Vorzeichen.
- Ist unser A nun *selbstadjungiert mit positiven Diagonaleinträgen*, so sind wir wieder in der selben Situation wie bei der Konstruktion der Cholesky-Zerlegung und wissen daher, dass diese existiert, also unsere Matrix positiv definit ist.
- Das Kriterium (2.44) lässt sich dahingehend abschwächen, dass man „>“ nur für eine Zeile verlangt, für die anderen lediglich „≥“ und dafür zusätzliche Voraussetzungen an die Besetzungsstruktur der Matrix macht, also wo Nullen stehen dürfen und wo nicht (siehe z. B. Numerik I). Ein solches Kriterium wird dann auch von der symmetrischen Matrix aus Beispiel 109 im ELA-Skript erfüllt, von der wir bereits wissen, dass alle Eigenwerte positiv sind, und sie daher nach Korollar 130 positiv definit ist. (Dass die dort angegebenen Eigenvektoren eine ONB bilden, lässt sich auch leicht nachrechnen.)
- Das Kriterium ist wirklich nur hinreichend und nicht notwendig, wie man etwa an der positiv definiten Matrix

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 100 \end{pmatrix}$$

sieht. Diese erfüllt auch nicht die abgeschwächte Form des Zeilensummenkriteriums.

Bevor wir eine weitere äquivalente Bedingung für die positive Definitheit betrachten machen wir uns das folgende kleine Lemma bewusst:

Lemma 134. *Hat eine $n \times n$ -Matrix A die (eventuell komplexen) Eigenwerte $\lambda_1, \dots, \lambda_n$, so ist*

$$\det(A) = \prod_{j=1}^n \lambda_j.$$

2 Das Innenprodukt kommt ins Spiel

Beweis. Wir müssen uns nur bewusst machen, dass

$$\det(\mathbf{A}) = \chi_{\mathbf{A}}(0),$$

wobei $\chi_{\mathbf{A}}$ das charakteristische Polynom von \mathbf{A} ist. □

Und ein noch offensichtlicheres:

Lemma 135. Die Determinante einer positiv definiten Matrix ist positiv.

Beweis. Das können wir wahlweise aus dem vorigen Lemma folgern oder aus der Cholesky-Zerlegung. □

Damit sich das Ganze leichter formulieren lässt, noch ein neuer Begriff:

Definition 136. Ist \mathbf{A} eine $n \times n$ -Matrix, so bekommen wir für eine natürliche Zahl $k \leq n$ den k -ten *führenden Hauptminor*, indem wir aus \mathbf{A} alles streichen außer den ersten k Zeilen und Spalten, in Formeln:

$$\mathbf{A}^{(k)} := (a_{i,j})_{i,j=1}^k,$$

und davon die Determinante bilden.

Beispiel 137. Wir betrachten folgende Matrix \mathbf{A} mit ihren führenden Untermatrizen:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad \mathbf{A}^{(1)} = (1), \quad \mathbf{A}^{(2)} = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}, \quad \mathbf{A}^{(3)} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix},$$

Damit bekommen wir die drei führenden Hauptminoren

$$\det(\mathbf{A}^{(1)}) = 1,$$

$$\det(\mathbf{A}^{(2)}) = 1 \cdot 5 - 4 \cdot 2 = -3,$$

$$\det(\mathbf{A}^{(3)}) = 1 \cdot 5 \cdot 9 + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 - 7 \cdot 5 \cdot 3 - 8 \cdot 6 \cdot 1 - 9 \cdot 4 \cdot 2 = 0.$$

Nun kommt das ersehnte *Hauptminorenkriterium*:

Satz 138 (Hauptminorenkriterium). Eine reelle oder komplexe selbstadjungierte $n \times n$ -Matrix \mathbf{A} ist genau dann positiv definit, wenn ihre führenden Hauptminoren positiv sind.

Beweis.

\Rightarrow : Sei also \mathbf{A} positiv definit. Das heißt aber, dass für jeden Vektor $\mathbf{y} \in \mathbb{C}^n$ bzw. \mathbb{R}^n gelten muss

$$\mathbf{y}^T \mathbf{A} \bar{\mathbf{y}} > 0,$$

auch für $\mathbf{y} = (y_1, \dots, y_k, 0, \dots, 0)^T =: \begin{pmatrix} \tilde{\mathbf{y}}_k \\ \mathbf{0} \end{pmatrix}$. Damit ist aber

$$0 < \mathbf{y}^T \mathbf{A} \bar{\mathbf{y}} = (\tilde{\mathbf{y}}_k^T, \mathbf{0}^T) \begin{pmatrix} \mathbf{A}^{(k)} & * \\ * & * \end{pmatrix} \begin{pmatrix} \overline{\tilde{\mathbf{y}}_k} \\ \mathbf{0} \end{pmatrix} = \tilde{\mathbf{y}}_k^T \mathbf{A}^{(k)} \overline{\tilde{\mathbf{y}}_k},$$

weshalb alle führenden Untermatrizen positiv definit sein müssen und nach Lemma 135 eine positive Determinante haben.

⇐: Da die führenden Hauptminoren positiv sind, sind auch alle führenden Untermatrizen regulär. Damit kann man sukzessive zeigen, dass sie alle eine LR-Zerlegung $A^{(k)} = L^{(k)}R^{(k)}$ besitzen (Übung). Aus der Überlegung

$$\begin{pmatrix} L^{(k)} & * \\ \mathbf{0} & * \end{pmatrix} \begin{pmatrix} R^{(k)} & \mathbf{0} \\ * & * \end{pmatrix} \begin{pmatrix} L^{(k)}R^{(k)} & * \\ * & * \end{pmatrix} = \begin{pmatrix} A^{(k)} & * \\ * & * \end{pmatrix}$$

sieht man leicht, dass die $L^{(k)}$ tatsächlich die führenden Untermatrizen von L und die $R^{(k)}$ entsprechend von R sind. Nun gilt offensichtlich für jedes k

$$0 < \det(A^{(k)}) = \det(R^{(k)}) = \prod_{j=1}^k r_{j,j}.$$

Damit diese Produkte alle positiv sind, müssen alle Faktoren, also die Diagonaleinträge von R positiv sein, womit wir wieder bei der Konstruktion der Cholesky-Zerlegung in Bemerkung 132 sind. Da die Cholesky-Zerlegung existiert, muss dann auch die Matrix A selber positiv definit sein.

□

Bemerkung 139. Welches der genannten Kriterien man am besten benutzt, um zu überprüfen, ob eine gegebene selbstadjungierte Matrix positiv definit ist, richtet sich nach der Situation:

- Gilt das starke Zeilensummenkriterium, so hat man bereits gewonnen. Dieses ist ja am einfachsten zu überprüfen.
- Ist die Matrix groß, und man möchte einen Rechner zum Test verwenden, so probiert man es mit der Cholesky-Zerlegung. Die braucht weniger Rechenoperationen als die Berechnung der Hauptminoren.
- Hat man eine kleine Matrix, und rechnet von Hand, so reicht meist das Hauptminorenkriterium. Beispiele für diesen Fall werden Sie in der Analysis 2 sehen, wenn es darum geht, zu entscheiden, ob ein kritischer Punkt ein Minimum ist (Hessematrix positiv definit), ein Maximum (Hessematrix negativ definit) oder ein Sattelpunkt.

Wie man das Hauptminorenkriterium umformuliert, damit man die negative Definitheit nachweisen kann, ohne das negative der Matrix auf positive Definitheit zu prüfen, können Sie sich als Übung gerne selber überlegen.

2.7.3 Polarzerlegung und Singulärwertzerlegung von Matrizen

Die Übertragung von Satz 111 auf Matrizen ist recht schnell erzählt. Wir müssen nur beachten, dass Isometrien im Reellen durch orthogonale, im Komplexen durch unitäre Matrizen dargestellt werden:

2 Das Innenprodukt kommt ins Spiel

Satz 140 (Polarzerlegung von quadratischen Matrizen).

- Zu jeder Matrix $A \in \mathbb{R}^{n \times n}$ gibt es eine orthogonale Matrix $Q \in \mathbb{R}^{n \times n}$ und eine eindeutige positiv semidefinite Matrix $\sqrt{A^T A}$ mit $\sqrt{A^T A}^2 = A^T A$, sodass

$$A = Q \sqrt{A^T A}.$$

- Zu jeder Matrix $A \in \mathbb{C}^{n \times n}$ gibt es eine unitäre Matrix $U \in \mathbb{C}^{n \times n}$ und eine positiv semidefinite Matrix $\sqrt{A^T A}$ mit $\sqrt{A^T A}^2 = A^T A$, sodass

$$A = U \sqrt{A^T A}.$$

Bei den Singulärwerten können wir die zweite Form der Definition bequem für Matrizen aufschreiben:

Definition 141 (Singulärwerte und Singulärvektoren von reellen Matrizen). Sei $A \in \mathbb{R}^{n \times n}$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$ und U, V Orthogonalmatrizen mit

$$AV = U\Sigma, \quad (\text{wobei } \Sigma = \text{diag}(\sigma_1, \dots, \sigma_n),) \quad (2.45)$$

so heißen die σ_j die *Singulärwerte* von A , die Spalten v_1, \dots, v_n von V die entsprechenden *Rechts-Singulärvektoren* und die Spalten u_j von U die *Links-Singulärvektoren* von A .

Bemerkung 142.

- Gleichung (2.45) bedeutet nichts anderes als

$$Av_j = \sigma_j u_j \quad \forall j = 1, \dots, n,$$

was genau unserer zweiten Definition bei den Endomorphismen war (mit q_j statt v_j). Immerhin wird in Gleichung (2.45) schnell deutlich, woher die Vorsätze „Rechts“ und „Links“ vor den Singulärvektoren kommen.

- Am wichtigsten ist die Singulärwertzerlegung für reelle Matrizen.
- Möchte man sie für komplexe Matrizen definieren, so muss man eigentlich nur $\mathbb{R}^{n \times n}$ durch $\mathbb{C}^{n \times n}$ und „Orthogonalmatrizen“ durch „unitäre Matrizen“ ersetzen.
- Multipliziert man Gleichung (2.45) von rechts mit V^T (im Reellen) bzw. $\overline{V^T}$ (im Komplexen), so bekommt man

$$A = U\Sigma V^*, \quad (2.46)$$

woher die Singulärwertzerlegung den Namensteil „Zerlegung“ hat.

Damit wird Satz 114 bzw. Korollar 115 zu

Satz 143 (Singulärwertzerlegung von Matrizen). *Jede quadratische reelle oder komplexe Matrix besitzt eine Singulärwertzerlegung.*

Bemerkung 144.

- Wie schon in Abschnitt 2.6 erwähnt, lässt sich die Existenz der Singulärwertzerlegung auch für nichtquadratische Matrizen beweisen. Die hier angegebene Definition reicht auch dafür aus.
- Wir werden uns in Kapitel 3 mit der Jordan-Normalform befassen, die sich bei der theoretischen Betrachtung von gewöhnlichen Differentialgleichungen (z. T. auch partiellen Differentialgleichungen) als nützlich erwiesen hat. Dort liegt der Fokus auf der wiederholten bzw. kontinuierlichen Anwendung der selben Matrix/Transformation, was insbesondere bei linearen Differentialgleichungen wichtig ist. Ist die Differentialgleichung nicht linear, so kann es zur Untersuchung, etwa der Stabilität, wichtiger sein, in einem bestimmten Zustand die Linearisierung und die zugehörige Singulärwertzerlegung zu betrachten. Schließlich ändert sich diese Linearisierung mit der Lösung, wird also nicht wiederholt angewandt, und es ist wichtiger die momentan stärksten Anteile der Transformation zu identifizieren.
- Eine weitere wichtige Anwendung ist die Hauptkomponentenanalyse in der Statistik. Hier sind die Matrizen im allgemeinen weit davon entfernt, quadratisch zu sein, da für viele Merkmalsträger eine begrenzte Anzahl von Merkmalen aufgenommen wird. Es geht also um multivariate Statistik und Datenanalyse.

3 Die Jordansche Normalform

3.1 Warum wir überhaupt sowas suchen

Da die wichtigste Verwendung der *Jordan-Normalform* die Behandlung gewöhnlicher Differentialgleichungen ist, insbesondere das Finden expliziter Lösungen für solche, wollen wir uns noch einmal an Abschnitt 7.4 aus der Elementaren Linearen Algebra erinnern. Dort hatten wir die Dualität zwischen dem Eigenwertproblem und dem Nullstellenproblem für Polynome entdeckt. Ausgangspunkt war die lineare Differentialgleichung

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_3y^{(3)} + a_2\ddot{y} + a_1\dot{y} + a_0y = 0$$

mit $a_0, \dots, a_{n-1} \in \mathbb{R}$ bzw. \mathbb{C} und $y = y(t)$.

Setzt man in diese den Ansatz $y(t) = \exp(\lambda t)$ ein, so führt dies auf

$$0 = e^{\lambda t} \left(\lambda^n + \sum_{k=1}^n a_{n-k} \lambda^{n-k} \right) =: e^{\lambda t} \chi(\lambda). \tag{3.1}$$

Da die Exponentialfunktion keine Nullstelle hat, führt es also auf das Nullstellenproblem für das charakteristische Polynom χ .

Den Zusammenhang mit Matrizen – und somit auch Endomorphismen – hatten wir hergestellt durch die Idee, die Differentialgleichung in ein System von solchen umzuschreiben, wobei nur noch erste Ableitungen auftauchen:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= x_3, \\ &\vdots \\ \dot{x}_{n-1} &= x_n, \\ \dot{x}_n &= -a_0x_1 - a_1x_2 - \dots - a_{n-1}x_n, \end{aligned}$$

mit Vektoren und Matrizen geschrieben

$$\dot{x} = Ax \quad \text{mit} \quad A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ -a_0 & -a_1 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}.$$

3 Die Jordansche Normalform

Da die Matrix A mit diesen Einträgen das selbe charakteristische Polynom hat wie die zugrundeliegende skalare Gleichung mit den höheren Ableitungen, nennt man sie auch die *Begleitmatrix* zum Polynom aus Gleichung (3.1).

Weil die Lösungen der einfachen Differentialgleichung

$$\dot{y} = \lambda y$$

die Gestalt

$$y(t) = y(0)e^{\lambda t}$$

haben, erwarten wir für ein System

$$\dot{\mathbf{y}}(t) = \mathbf{A}\mathbf{y}(t), \quad \mathbf{y}(t) \in \mathbb{R}^n, \quad \mathbf{A} \in \mathbb{R}^{n \times n}$$

Lösungen der Form

$$\mathbf{y}(t) = e^{t\mathbf{A}}\mathbf{y}(0). \quad (3.2)$$

Dabei machen wir uns zu nutze, dass

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Da das ein Grenzwert von Polynomen mit steigendem Grad ist, und wir quadratische Matrizen ungestraft in Polynome einsetzen dürfen, ist Formel (3.2) wohldefiniert. (Diejenigen, die erst im kommenden Wintersemester mit der Analysis beginnen, nehmen dies einfach mal als Spoiler zur Kenntnis.) Bemerkenswert ist außerdem, dass dies alles auch dann noch funktioniert, wenn wir die reellen Zahlen \mathbb{R} durch die komplexen \mathbb{C} ersetzen.

Das Schöne ist, dass wir im Komplexen bereits gezeigt haben, dass jede quadratische Matrix ähnlich zu einer oberen Dreiecksmatrix ist (Satz 10). Ist also

$$\mathbf{A} = \mathbf{S}\mathbf{T}\mathbf{S}^{-1}$$

mit einer oberen Dreiecksmatrix \mathbf{T} , so gilt für jedes Polynom p

$$p(\mathbf{A}) = \mathbf{S}p(\mathbf{T})\mathbf{S}^{-1}$$

und dementsprechend

$$e^{t\mathbf{A}} = \mathbf{S}e^{t\mathbf{T}}\mathbf{S}^{-1}. \quad (3.3)$$

Wäre nun $\mathbf{T} = \mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$ bereits eine Diagonalmatrix, so hätten wir

$$e^{t\mathbf{T}} = \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t}),$$

also wieder eine Diagonalmatrix. Da auch die Potenzen von oberen Dreiecksmatrizen wieder solche sind, stellt sich die Frage, wie nahe man an eine Diagonalmatrix kommen kann, wenn man mit einer nicht diagonalisierbaren Matrix startet. Das Ziel ist es, auf eine Besetzung oberhalb der Hauptdiagonalen zu kommen, die beim Einsetzen in die Exponentialfunktion möglichst wenig zusätzliche Nichtnullen bekommt. In diesem Sinne optimal ist die *Jordansche Normalform*, der wir in diesem Kapitel nachjagen.

Bemerkung 145.

- Da wir im Reellen noch nicht einmal die Garantie haben, dass unsere jeweilige Matrix trigonalisierbar ist, lassen wir es uns egal sein, wenn nicht alles, was im folgenden kommt, auch im Reellen klappt. Es reicht uns auch im Komplexen. Da die Exponentialfunktion auch im Komplexen funktioniert, ist das keine echte Einschränkung.
- Es gibt natürlich auch mathematische Fragestellungen, wo wir uns nicht ins Komplexe zurückziehen können. Von denen werden Ihnen aber während des Studiums selten welche begegnen, sodass es reicht, wenn Sie dann schauen, was man ersatzweise tun kann.
- Bei oberen Dreiecksmatrizen kann man die Eigenwerte an der Hauptdiagonalen ablesen. Dabei ist es eine bemerkenswerte Besonderheit, wenn eine Zahl dort mehrfach auftaucht. Aber nur, wenn das passiert, kann es auch sein, dass die Matrix nicht diagonalisierbar ist (Korollar 121 aus dem ELA-Skript). Sobald wir ein wenig an der Matrix „wackeln“, ist sie wieder diagonalisierbar.
- Das bedeutet vor allem, dass unsere folgenden Betrachtungen exakte Rechnung voraussetzen, und wir nicht mit numerisch generierten Näherungen arbeiten können¹. Da haben wir aber nur bis zum Fall 4×4 die Garantie, dass wir das charakteristische Polynom auflösen können. Mit steigender Dimension sinkt die Chance, dieses tun zu können. Daher werden auch in der Praxis die Beispielmatrizen nicht besonders groß sein und, wenn doch, recht angenehm strukturiert. (Ansonsten wird es schwierig.)
- Fehlende Diagonalisierbarkeit steht oft mit Resonanzphänomenen im Zusammenhang.

3.2 Vorüberlegungen zu wünschenswerten Matrixstrukturen

In Kapitel 1 hatten wir sukzessive wachsende invariante Unterräume verwendet, um die Darstellung eines Operators durch eine rechte obere Dreiecksmatrix zu beweisen. Wir hatten auch gesehen, dass, wenn sich der Gesamtraum als direkte Summe der Eigenräume schreiben lässt, bereits eine Diagonalisierung möglich ist. Nun hätten wir gerne für den Fall eines nichtdiagonalisierbaren Endomorphismus zu jedem der Eigenwerte $\lambda_1, \dots, \lambda_k$ einen invarianten Unterraum $H_{\lambda_j} \supseteq V_{\lambda_j}$, einen sogenannten *Hauptraum* gefunden, sodass sich der Gesamtraum als deren direkte Summe schreiben lässt, also im Fall $V = \mathbb{C}^n$ so etwas wie

$$\mathbb{C}^n = \bigoplus_{j=1}^k H_{\lambda_j} .$$

¹Muss man das trotzdem, ist es eine ziemlich heikle Angelegenheit.

3 Die Jordansche Normalform

Das gibt uns die Hoffnung, zu jedem Eigenwert einen Block $J_{\lambda_j} \in \mathbb{C}^{\dim H_{\lambda_j} \times \dim H_{\lambda_j}}$ zu finden, dessen Diagonaleinträge alle λ_j sind, der obere Dreiecksgestalt hat, und aus denen sich eine Darstellungsmatrix der Gestalt

$$\begin{pmatrix} J_{\lambda_1} & \mathbf{0} & \dots & \dots & \mathbf{0} \\ \mathbf{0} & J_{\lambda_1} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \dots & \mathbf{0} & J_{\lambda_k} \end{pmatrix}, \quad \in \mathbb{C}^{n \times n},$$

einer sogenannten *Blockdiagonalgestalt* ergibt. Deshalb wollen wir zunächst diese Bausteine näher betrachten, also Blockdiagonalmatrizen und Matrizen der Gestalt

$$\lambda E + \begin{pmatrix} 0 & * & \dots & * \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & * \\ 0 & \dots & \dots & 0 \end{pmatrix} =: \lambda E + N. \quad (3.4)$$

Matrizen der Gestalt N , also *strikte/echte obere Dreiecksmatrizen*, hatten wir in der ELA bereits in der Übung betrachtet, und dabei festgestellt, dass mit jeder Potenz oberhalb der Hauptdiagonalen eine weitere Diagonale zu Null wird, also irgendwann die Nullmatrix rauskommt. Diese Eigenschaft soll einen eigenen Namen bekommen:

Definition 146.

- Eine Matrix $A \in \mathbb{C}^{n \times n}$ (oder $\mathbb{R}^{n \times n}$) heißt *nilpotent*, wenn es ein $p \in \mathbb{N}$ gibt mit

$$A^p = \mathbf{0}.$$

- Sei V ein Vektorraum. Ein Endomorphismus $\varphi \in \text{hom}(V, V)$ heißt *nilpotent*, wenn es ein $p \in \mathbb{N}$, sodass φ^p die Nullabbildung ist.

Lemma 147. In Definition 146 reicht immer $p = n$ bzw. im zweiten Fall $p = \dim V$.

Beweis. Die Dimension des Bilds muss mit jeder Potenz der Matrix bzw. des Operators fallen, sonst bleibt sie stehen. (Das ist jetzt extra knapp formuliert, damit Sie was zum Nachdenken haben.) \square

Wir finden auch, dass die Form N der nilpotenten Matrix in Formel (3.4) reicht, um alle nilpotenten Endomorphismen zu beschreiben:

Lemma 148. Jeder nilpotente Endomorphismus besitzt bezüglich einer geeigneten Basis eine Matrixdarstellung der Form N aus Formel (3.4).

Beweis. Nennen wir unseren Endomorphismus ϕ . Dann wissen wir aufgrund der Nilpotenz, dass er einen nichttrivialen Kern hat, der mit jeder Potenz von ϕ größer wird.

3.2 Vorüberlegungen zu wünschenswerten Matrixstrukturen

- Wir nehmen eine Basis dieses Kerns, etwa $\mathbf{v}_1, \dots, \mathbf{v}_{k_1}$ und nehmen diese als Anfang unserer gesuchten Basis. Da alle diese auf den Nullvektor abgebildet werden, sind die ersten k_1 Spalten der Darstellungsmatrix Null.
- Nun ergänzen wir diese Basisvektoren zu einer Basis des Kerns von ϕ^2 , also um die Vektoren $\mathbf{v}_{k_1+1}, \dots, \mathbf{v}_{k_2}$. Diese werden offenbar auf Vektoren aus dem Kern von ϕ abgebildet, also auf Linearkombinationen der $\mathbf{v}_1, \dots, \mathbf{v}_{k_1}$. Infolgedessen haben die Spalten $k_1 + 1$ bis k_2 der Darstellungsmatrix von ϕ nur Einträge in den ersten k_1 Zeilen.
- Ergänzen wir die Basis weiter zu einer von $\ker \phi^3$, so bekommen wir analog in den Spalten $k_2 + 1$ bis k_3 nur Einträge in den ersten k_2 Zeilen.
- In dieser Weise fortfahrend erhalten wir am Ende eine Darstellungsmatrix, bei der alle nichtverschwindenden Einträge oberhalb der Hauptdiagonalen liegen.

□

Das ist jetzt nicht der Einfachste Beweis – einfacher wäre gewesen, festzustellen, dass ein Nilpotenter Operator keinen von Null verschiedenen Eigenwert haben kann, sonst wäre er nicht nilpotent; der Rest folgt dann mit der Trigonalisierung in Satz 11 –, aber er hat den Vorteil, den Gedankengang aus dem Beweis von Satz 11 nochmal aufnimmt.

Wir wollen diesen Gedankengang und damit das Ergebnis noch weiter verschärfen, denn es gefällt uns nicht so recht, dass die Einträge zwar immer weiter von der Diagonalen wegrücken, aber irgendwie unkontrollierbar.

Lemma 149. *Jeder nilpotente Endomorphismus besitzt bezüglich einer geeigneten Basis eine Matrixdarstellung der Form*

$$\begin{pmatrix} 0 & * & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & * \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}, \quad (3.5)$$

wobei die Sternchen $*$ für Einsen oder Nullen stehen.

Beweis. Wir passen den Beweis des vorigen Lemmas ein wenig an und betrachten zunächst den Fall, dass der Kern eindimensional ist.

- Als Basis des Kerns reicht hier ein einzelner Eigenvektor \mathbf{v}_1 zum Eigenwert Null. Das gibt uns eine einzelne Nullspalte am Anfang der Matrix.
- Wähle nun als zweiten Basisvektor einen Vektor $\mathbf{v}_2 \in (\ker \phi^2 \setminus \ker \phi) \cap \phi^{-1}(\{\mathbf{v}_1\})$. Dieser gibt uns als zweite Spalte den ersten Standardeinheitsvektor \mathbf{e}_1 , wird er doch auf \mathbf{v}_1 abgebildet.

3 Die Jordansche Normalform

- Als dritten Basisvektor nehmen wir nun ein $\mathbf{v}_3 \in (\ker \varphi^3 \setminus \ker \varphi^2) \cap \varphi^{-1}(\{\mathbf{v}_2\})$. Da dieser auf \mathbf{v}_2 abgebildet wird, bekommen wir als dritte Matrixspalte \mathbf{e}_2 .
- Fahre sinngemäß fort, bis ich die Dimension des Raums ausgeschöpft habe.

Das gibt uns eine Darstellungsmatrix der Form

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & 0 \end{pmatrix}, \quad (3.6)$$

Ist der Kern zwei- oder höherdimensional, wird der beschriebene Vorgang schon abbrechen müssen, ohne dass die Dimension des Raums ausgeschöpft ist, einfach weil sich kein Vektor in

$$(\ker \varphi^k \setminus \ker \varphi^{k-1}) \cap \varphi^{-1}(\{\mathbf{v}_{k-1}\})$$

mehr findet. Dann starten wir den Vorgang von neuem mit dem nächsten Basisvektor des Kerns, der dann entsprechend in der k -ten Spalte eine Nullspalte macht. Dann folgen wieder Einheitsvektoren \mathbf{e}_k, \dots , bis auch da der Vorgang wieder abbricht. Bei einem zweidimensionalen Kern, weil die Dimension voll ist, bei höherdimensionalem Kern anderweitig, sodass wir mit dem dritten Basisvektor des Kerns wieder einsteigen müssen, etc. \square

Bemerkung 150. Die so entstandene Matrixdarstellung hat die Blockdiagonalform $(\mathbf{J}_1, \dots, \mathbf{J}_m)$, wobei m die Dimension des Kerns von φ ist, und die \mathbf{J}_j die Gestalt (3.6) haben. Man nennt die \mathbf{J}_j dann *Jordanblöcke* und das Ganze, also die Matrix $(\mathbf{J}_1, \dots, \mathbf{J}_m)$ die *Jordan-Normalform* von φ .

Diese wollen wir später auf allgemeine komplexe Matrizen übertragen, wobei wir dort die Diagonalen der Jordanblöcke durch solche der Form $\lambda \mathbf{E}$ mit einem Eigenwert λ des Endomorphismus bzw. der Matrix ersetzen wollen.

Am Ende dieses Abschnitts noch ein kleines Ergebnis zu Blockdiagonalmatrizen:

Lemma 151. Sind $\mathbf{A} = \text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_k), \mathbf{B} = \text{diag}(\mathbf{B}_1, \dots, \mathbf{B}_k) \in \mathbb{C}^{n \times n}$ mit $\mathbf{A}_j, \mathbf{B}_j \in \mathbb{C}^{n_j \times n_j} \forall j$, dann ist

$$\mathbf{AB} = \text{diag}(\mathbf{A}_1 \mathbf{B}_1, \dots, \mathbf{A}_k \mathbf{B}_k).$$

(Das selbe gilt natürlich auch mit \mathbb{R} statt \mathbb{C} .)

Beweis. Übung. \square

Auf alle diese Dinge werden wir später nochmals zurückgreifen. Insbesondere werden wir das letzte Lemma immer im Hinterkopf behalten, auch wenn wir es nicht extra zitieren.

3.3 Die Sache mit den Haupträumen ...

Um die Besetzungsstruktur der Darstellungsmatrix eines Endomorphismus genauer kontrollieren zu können, als wir das für die Trigonalisierung in Kapitel 1 getan haben, müssen wir die invarianten Unterräume genauer untersuchen. Die wichtigsten davon sind die sogenannten *Haupträume*, die wir uns im folgenden anschauen.

Wie wir schon gesehen hatten, sind die Eigenvektoren einer Matrix A zum Eigenwert λ gerade die nichttrivialen Lösungen des homogenen LGS $(A - \lambda E)x = \mathbf{0}$. Damit ist

$$V_\lambda = \ker(A - \lambda E).$$

Nun ist aber

$$V_\lambda = \ker(A - \lambda E) \subseteq \ker((A - \lambda E)^2) \subseteq \ker((A - \lambda E)^3) \subseteq \dots \quad (3.7)$$

Gilt an einer Stelle Gleichheit, gilt sie offenbar danach auch weiter. Dass dies irgendwann passiert, ist klar, da bei Ungleichheit jedesmal die Dimension um mindestens Eins größer wird. Das geht aber nicht beliebig oft. Schließlich befinden wir uns im Endlichdimensionalen. Für eine $(n \times n)$ -Matrix ist also spätestens beim Exponenten n Schluss. Offenbar sind alle diese Räume invariant unter der von A induzierten linearen Abbildung. So gilt etwa für $x \in \ker((A - \lambda E)^2)$

$$\begin{aligned} (A - \lambda E)^2(Ax) &= ((A - \lambda E)^2 [(A - \lambda E) + \lambda E])x \\ &= [(A - \lambda E)^3 + \lambda(A - \lambda E)^2]x \\ &= (A - \lambda E)^3x + \lambda(A - \lambda E)^2x. \end{aligned}$$

Da aber nach Voraussetzung x im Kern von $(A - \lambda E)^2$ und damit erst recht im Kern von $(A - \lambda E)^3$ liegt, verschwinden beide Summanden. Damit liegt aber auch Ax im Kern von $(A - \lambda E)^2$. Wie man leicht sieht, hätten wir anstelle der Exponenten zwei und drei auch beliebige r und $r + 1$ hinschreiben können. Formel (3.7) liefert uns also eine endliche Folge von invarianten Unterräumen mit aufsteigender Dimension.

Diese Überlegungen führen uns zur folgenden Definition:

Definition 152. Ist $A \in \mathbb{C}^{n \times n}$ (bzw. $\in \mathbb{R}^{n \times n}$) und $\lambda \in \mathbb{C}$ (bzw. $\in \mathbb{R}$), dann nennen wir

$$H_\lambda := \ker((A - \lambda E)^n)$$

den *Hauptraum* von A bzgl. λ . Die Elemente von H_λ heißen die *Hauptvektoren* von A bzgl. λ . Ist

$$v \in \ker((A - \lambda E)^p) \setminus \ker((A - \lambda E)^{p-1}),$$

so nennen wir v einen *Hauptvektor der Stufe* p .

Damit sind insbesondere die Eigenvektoren Hauptvektoren der Stufe 1. Aus der Definition folgt direkt

3 Die Jordansche Normalform

Lemma 153. Die Einschränkung $(A - \lambda E)|_{H_\lambda}$ ist nilpotent.

Damit wissen wir bereits, dass diese Einschränkung eine Matrixdarstellung der Blockdiagonalform (J_1, \dots, J_m) hat, wobei die Blöcke J_j alle die Gestalt (3.6) haben, also gelegentliche Einsen auf der ersten oberen Nebendiagonalen, ansonsten Nullen.

Ist λ kein Eigenwert von A , dann ist der Hauptraum trivial. Er ist also nur spannend für Eigenwerte der Matrix. Wir erhoffen uns, dass wir mit dem Übergang von Eigenräumen zu den evtl. größeren Haupträumen auch bei nichtdiagonalisierbaren Matrizen den Gesamtraum, also \mathbb{C}^n bzw. \mathbb{R}^n als direkte Summe invarianter Unterräume schreiben können. Dazu müssen wir zeigen, dass für verschiedene Eigenwerte die Haupträume trivialen Schnitt haben:

Satz 154. Seien $\lambda \neq \mu$ Eigenwerte von A , so gilt

$$H_\lambda \cap H_\mu = \{0\}.$$

Beweis. Wir zeigen induktiv, dass für $k + m = N$ gilt

$$\ker((A - \lambda E)^k) \cap \ker((A - \mu E)^m) = \{0\} \quad (3.8)$$

und beginnen mit $N = 0$.

- Dann ist $k = m = 0$, und wir schneiden den Kern der Identität mit sich selber. Das ist aber schlicht der gewünschte Nullraum.
- Sei nun (3.8) für N bewiesen und $k + m = N + 1$.
 - Falls $k = 0$ oder $m = 0$ ist wieder alles klar, weil die Identität den Nullraum als Kern hat.
 - Ansonsten wähle ein beliebiges

$$v \in \ker((A - \lambda E)^k) \cap \ker((A - \mu E)^m)$$

und betrachte den Vektor $u = (A - \lambda E)v$. Für diesen gelten dann

$$\begin{aligned} (A - \lambda E)^{k-1}u &= (A - \lambda E)^k v = 0, \\ (A - \mu E)^m u &= (A - \lambda E)(A - \mu E)^m v = 0. \end{aligned}$$

Damit liegt u in $\ker((A - \lambda E)^{k-1}) \cap \ker((A - \mu E)^m)$ und muss nach Induktionsvoraussetzung selbst der Nullvektor sein. Damit ist aber

$$v \in \ker(A - \lambda E) = V_\lambda.$$

Betrachten wir analog den Vektor $w = (A - \mu E)v$, so finden wir auch $v \in V_\mu$. Da aber der Schnitt von Eigenräumen zu verschiedenen Eigenwerten trivial ist, muss v selber der Nullvektor sein, was wir ja zeigen wollten.

□

Eine zu (3.7) analoge Formel lässt sich allgemein für quadratische Matrizen aufstellen:

$$\mathbb{C}^n = \ker E = \ker A^0 \subseteq \ker A \subseteq \ker A^2 \subseteq \ker A^3 \subseteq \dots \quad (3.9)$$

Dass ab dort, wo das erste mal Gleichheit gilt, immer Gleichheit gilt, ist natürlich auch hier der Fall. Ist nämlich $\ker A^k = \ker A^{k+1}$, so gilt für $\mathbf{v} \in \ker A^{k+2}$, dass $A\mathbf{v} \in \ker A^{k+1} = \ker A^k$, weshalb \mathbf{v} dann selber wieder im Kern von A^{k+1} liegen muss. Da aber die Dimension der Kerne in (3.9) nicht größer als die Dimension des Raums werden kann, haben wir folgendes Lemma:

Lemma 155. Sei $A \in \mathbb{C}^{n \times n}$, dann gelten für jedes $k \in \mathbb{N}$

$$\ker A^{n+k} = \ker A^n, \quad (3.10)$$

$$\operatorname{im} A^{n+k} = \operatorname{im} A^n \quad (3.11)$$

und

$$\mathbb{C}^n = \operatorname{im} A^n \oplus \ker A^n. \quad (3.12)$$

(Im Reellen alles analog.)

Beweis. Die Formeln (3.10) und (3.11) folgen direkt aus unserer Vorüberlegung.

Für (3.12) genügt es zu zeigen, dass die Summe direkt ist, also der Schnitt zwischen Bild und Kern trivial. Der Rest folgt dann aus dem Dimensionssatz.

Sei also $\mathbf{v} \in \operatorname{im} A^n \cap \ker A^n$. Weil es im Bild liegt, gibt es ein \mathbf{u} mit $A^n \mathbf{u} = \mathbf{v}$. Weil es im Kern liegt, gilt aber auch $A^n \mathbf{v} = \mathbf{0}$. Also ist

$$A^{2n} \mathbf{u} = \mathbf{0} \quad \Rightarrow \quad \mathbf{u} \in \ker A^n \quad \Rightarrow \quad \mathbf{v} = A^n \mathbf{u} = \mathbf{0}.$$

Damit ist die Summe wie gewünscht direkt. □

Offenbar kommutieren die Matrizen $(A - \lambda E)$, d. h. es ist $(A - \lambda E)(A - \mu E) = (A - \mu E)(A - \lambda E)$ unabhängig von den Werten für λ und μ . Für Matrizen, die miteinander kommutieren, können wir aber folgendes zeigen:

Lemma 156. Sind die Matrizen $A_1, \dots, A_k \in \mathbb{C}^{n \times n}$ (bzw. $\mathbb{R}^{n \times n}$) paarweise vertauschbar, dann gilt

$$\ker(A_1^n \cdots A_k^n) = \sum_{j=1}^k \ker(A_j^n).$$

Für Matrizen der Gestalt $(A - \lambda_j E)$ stünde also rechts die Summe der Haupträume zu den λ_j .

Beweis.

3 Die Jordansche Normalform

- Gilt es für zwei je zwei Matrizen, dann ist der Rest einfach:

Schreibe $A_1^n \cdots A_k^n$ einfach als $(A_1^n \cdots A_{k-1}^n)A_k^n$, also als Produkt von zwei Matrizen, wobei wir beachten, dass $A_1^n \cdots A_{k-1}^n = (A_1 \cdots A_{k-1})^n$. Damit können wir also die Erkenntnis für zwei Matrizen anwenden, und finden

$$\begin{aligned} \ker(A_1^n \cdots A_k^n) &= \ker(A_1^n \cdots A_{k-1}^n) + \ker(A_k^n) \\ &= \left[\sum_{j=1}^{k-1} \ker(A_j^n) \right] + \ker(A_k^n). \end{aligned}$$

- Für zwei Matrizen verwenden wir das Wissen, dass

$$\ker(A^n) \subseteq \ker(A^n B^n) \quad \text{und} \quad \ker(B^n) \subseteq \ker(A^n B^n).$$

Damit ist auch

$$\ker(A^n) + \ker(B^n) \subseteq \ker(A^n B^n),$$

und wir können die zugehörigen Endomorphismen als Endomorphismen auf diesem Kern des Produkts, etwa dem Vektorraum U , betrachten. Auf diesem ist aber $\varphi_{A^n B^n} = \varphi_{A^n} \circ \varphi_{B^n}$ die Nullabbildung und damit

$$\text{im } \varphi_{B^n} \subseteq \ker \varphi_{A^n}.$$

Mit Lemma 155 gilt dann

$$\begin{aligned} U &= \ker \varphi_{B^n} + \text{im } \varphi_{B^n} \\ &\subseteq \ker \varphi_{B^n} + \ker \varphi_{A^n} \\ &\subseteq U. \end{aligned}$$

Damit muss aber überall Gleichheit gelten, und, weil wir $U = \ker(A^n B^n)$ gesetzt hatten folgt die Behauptung auch für zwei kommutierende Matrizen.

□

Nun kommt endlich der schon in der ELA angedeutete Sachverhalt, dass zumindest im Komplexen jede quadratische Matrix von ihrem charakteristischen Polynom annulliert wird. Im Grunde hätten wir ihn schon in Kapitel 1 beweisen können. Aber jetzt werden wir ihn gleich brauchen, um zu zeigen, dass die direkte Summe der Haupträume wirklich den ganzen Raum ergibt. Bei einer diagonalisierbaren Matrix ist das trivial, da für jeden Eigenwert λ mit $A\mathbf{r} = \lambda\mathbf{r}$ auch gilt

$$\chi_A(A)\mathbf{r} = \chi_A(\lambda)\mathbf{r} = 0 \cdot \mathbf{r} = \mathbf{0}.$$

Da sich in diesem Fall eine Basis aus Eigenvektoren finden lässt, ist also bereits $\chi_A(A)$ eine Matrixdarstellung der Nullabbildung und damit die Nullmatrix.

Mithilfe von Analysis hätte man nun sehr einfach zeigen können, dass dies auch für nicht diagonalisierbare Matrizen gilt:

Ist A nicht diagonalisierbar, so hat sie mehrfache Eigenwerte und damit die Schursche Normalform (oder eine andere Trigonalisierung) mehrfache Diagonaleinträge. Nun kann ich diese beliebig gut durch eine obere Dreiecksmatrix mit paarweise verschiedenen Diagonaleinträgen approximieren. Den Rest bekommt man dann durch Stetigkeitsargumente.

Es geht aber auch ohne Analysis. Zunächst für obere Dreiecksmatrizen, dann allgemein:

Proposition 157. Eine obere Dreiecksmatrix $T \in \mathbb{C}^{n \times n}$ wird von ihrem charakteristischen Polynom χ_T annulliert.

Beweis. Schreibe die Dreiecksmatrix als

$$T = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

Damit ist dann

$$\chi_T(x) = \prod_{j=1}^n (x - \lambda_j).$$

Wir gehen induktiv und versuchen für $j = 1, \dots, n$ zu zeigen, dass

$$(T - \lambda_1 E) \cdots (T - \lambda_j E) e_j = \mathbf{0},$$

wobei e_j den j -ten Standardbasisvektor bezeichnet. Wichtig ist dabei zu beachten, dass die Matrixfaktoren miteinander vertauschbar sind.

- Für $j = 1$ ist das trivial, da der erste Faktor im Produkt in der ersten Spalte den Nullvektor hat.
- Gelte es nun bis j . Versuche zu zeigen, dass es auch für $j + 1$ klappt:

Der Vektor $(T - \lambda_{j+1} E) e_{j+1}$ ist ja gerade die $j + 1$ -te Spalte der Matrix $(T - \lambda_{j+1} E)$, hat also ab der $j + 1$ -ten Komponente nur Nullen und kann daher als Linearkombination der e_1, \dots, e_j geschrieben werden. Für die ist aber schon alles gezeigt (nach Induktionsvoraussetzung), und wir sind fertig.

Da aber mit der Vertauschbarkeit der Matrixfaktoren für jedes j

$$\chi_T(T) e_j = \left(\prod_{k=j+1}^n (T - \lambda_k E) \right) \left(\prod_{k=1}^j (T - \lambda_k E) \right) e_j = \left(\prod_{k=j+1}^n (T - \lambda_k E) \right) \mathbf{0} = \mathbf{0},$$

verschwindet $\chi_T(T)$ auf einer Basis des Raums, stellt also die Nullabbildung dar und ist demnach die Nullmatrix. □

3 Die Jordansche Normalform

Satz 158 (Cayley-Hamilton). Sei $A \in \mathbb{C}^{n \times n}$ und χ_A das zugehörige charakteristische Polynom. Dann gilt

$$\chi_A(A) = \mathbf{0}.$$

Beweis. Wegen Satz 11 wissen wir, dass A trigonalisierbar ist, also mit einer geeigneten Transformation gilt

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} =: T.$$

Damit ist dann

$$\chi_A(A) = \chi_T(A) = \chi_T(STS^{-1}) = S\chi_T(T)S^{-1},$$

weshalb wir für unsere Betrachtungen A o.B.d.A. als obere Dreiecksmatrix voraussetzen können. (Ist ja klar, weil es ja eigentlich um den dargestellten Endomorphismus geht und nicht um die konkrete Matrixdarstellung.) Damit sind wir aber in der Situation von Proposition 157 und bereits fertig. \square

Bemerkung 159. Da eine reelle Matrix auch eine komplexe Matrix ist, gilt das Ganze auch für Matrizen $A \in \mathbb{R}^{n \times n}$.

Satz 160 (Hauptraumzerlegung). Sei $A \in \mathbb{C}^{n \times n}$ (bzw. $\mathbb{R}^{n \times n}$) mit dem charakteristischen Polynom

$$\chi_A(z) = \prod_{j=1}^k (z - \lambda_j)^{n_j}, \quad (3.13)$$

wobei λ_j die paarweise verschiedenen Eigenwerte sind mit ihren Vielfachheiten n_j . Dann ist

$$\mathbb{C}^n = \bigoplus_{j=1}^k H_{\lambda_j}, \quad (3.14)$$

also der Gesamttraum die direkte Summe der Haupträume unserer Matrix. (Im reellen Fall müssen wir natürlich \mathbb{R}^n statt \mathbb{C}^n schreiben, und die Forderung, dass das charakteristische Polynom die Gestalt (3.13) hat, ist eine zusätzliche Einschränkung.)

Beweis. Nach dem Satz von Cayley-Hamilton 158 ist

$$\mathbf{0} = \chi_A(A) = \prod_{j=1}^k (A - \lambda_j E)^{n_j} =: \prod_{j=1}^k A_j^{n_j}.$$

(Wir bezeichnen also $(A - \lambda_j E)$ kurzerhand als A_j .) Wir multiplizieren das Ganze mit einem geeigneten Polynom und erhalten

$$\mathbf{0} = \left(\prod_{j=1}^k A_j^{n-n_j} \right) \cdot \left(\prod_{j=1}^k A_j^{n_j} \right) = \prod_{j=1}^k A_j^n,$$

was bedeutet, dass der Kern des Produkts ganz rechts der gesamte Raum ist, und die Kerne der A_j^n sind ja gerade die Haupträume. Die haben aber wieder paarweise einen trivialen Schnitt, sodass ihre Summen immer direkt sind. Dieses zusammen mit Lemma 156 liefert uns

$$\mathbb{C}^n = \ker \left(\prod_{j=1}^k A_j^n \right) = \sum_{j=1}^k \ker A_j^n = \bigoplus_{j=1}^k \ker A_j^n = \bigoplus_{j=1}^k H_{\lambda_j},$$

wobei wir im reellen Fall natürlich \mathbb{R}^n statt \mathbb{C}^n schreiben. □

Korollar 161. *Es gilt $\dim H_{\lambda_j} = n_j$.*

Beweis. Mit dem Satz und der Tatsache, dass das charakteristische Polynom zum zugrundeliegenden Endomorphismus gehört, gilt

$$\chi_A = \chi_{\varphi_A} = \prod_{j=1}^k \chi_{\varphi_A|_{H_{\lambda_j}}},$$

wobei wir bei der letzten Gleichheit die Invarianz der Haupträume verwendet haben. Aber damit folgt auch bereits die Behauptung. □

Die Dimension ist also gleich der Vielfachheit des Eigenwerts als Nullstelle des charakteristischen Polynoms. Man nennt diese auch die *algebraische Vielfachheit* des Eigenwerts.

Satz 162. *Sei $A \in \mathbb{C}^{n \times n}$ (bzw. $\mathbb{R}^{n \times n}$) mit dem charakteristischen Polynom*

$$\chi_A(z) = \prod_{j=1}^k (z - \lambda_j)^{n_j},$$

wobei λ_j die paarweise verschiedenen Eigenwerte sind mit ihren Vielfachheiten n_j . Dann gibt es eine Basis aus Hauptvektoren von A , bezüglich der der Endomorphismus φ_A durch eine Blockdiagonalmatrix $\text{diag}(T_1, \dots, T_k)$ mit Dreiecksmatrizen der Gestalt

$$T_j = \begin{pmatrix} \lambda_j & * & \dots & * \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \lambda_j \end{pmatrix} \in \mathbb{C}^{n_j \times n_j} \tag{3.15}$$

dargestellt wird. In anderen Worten ist A ähnlich zu $\text{diag}(T_1, \dots, T_k)$.

Da die Gefahr besteht, dass ein vollständiger formaler Beweis die Zusammenhänge mehr verdunkelt als erhellt, hier nur eine Beweisskizze.

Beweisskizze.

3 Die Jordansche Normalform

- In Satz 9 hatten wir benutzt, dass die Eigenschaft der linearen Hüllen der einzelnen Eigenvektoren als eindimensionale invariante Unterräume zu einer Blockdiagonalgestalt mit 1×1 -Blöcken, also zur Diagonalgestalt führt. In der selben Weise ergibt sich hier also eine Blockdiagonalmatrix, bei der zu jedem Hauptraum einer der Blöcke auf der Diagonalen gehört.
- Analog zu Satz 10 (Trigonalisierung) können wir nun in den Haupträumen die Basis so wählen, dass sich am Ende eine obere Dreiecksmatrix ergibt.

Insgesamt bleibt der Darstellungsmatrix bezüglich dieser Basis nichts anderes übrig, als die oben behauptete Form (3.15) zu haben. \square

3.4 ... und wie wir uns daraus unsere Jordan-Normalform bauen

Nun hätten wir die Form (3.15) gerne noch weiter vereinfacht. Sie hat ja die Grundform $\lambda_j E + N_j$, wobei N_j eine strenge obere Dreiecksmatrix und damit nilpotent ist. Insbesondere ist $N_j^{n_j} = \mathbf{0}$. Wir machen nun Gebrauch von der Tatsache, dass für eine Ähnlichkeitstransformation von $\lambda_j E + N_j$ gilt

$$S^{-1}(\lambda_j E + N_j)S = S^{-1}(\lambda_j E)S + S^{-1}N_j S = \lambda_j E + S^{-1}N_j S.$$

Mit Lemma 149 und der zugehörigen Bemerkung 150 können wir also ohne weitere Beschränkung davon ausgehen, dass

$$\lambda_j E + N_j = \text{diag}(J_1, \dots, J_{k_j}) \quad \text{mit Blöcken} \quad J_\ell = \begin{pmatrix} \lambda_j & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda_j \end{pmatrix},$$

die sich auch zu 1×1 -Blöcken der Gestalt (λ_j) reduzieren können.

Entsprechend der Konstruktion im Beweis von Lemma 149 ist k_j die Dimension des Eigenraums V_{λ_j} , die *geometrische Vielfachheit* des Eigenwerts λ_j . Die Länge des größten *Jordanblocks* J_ℓ gibt nun an, ab welcher Potenz N_j^ℓ verschwindet. Nennen wir diese einfach m_j , so finden wir leicht:

Satz 163. Sei $A \in \mathbb{C}^{n \times n}$ (bzw. $\mathbb{R}^{n \times n}$) mit dem charakteristischen Polynom

$$\chi_A(z) = \prod_{j=1}^k (z - \lambda_j)^{n_j},$$

3.4 ... und wie wir uns daraus unsere Jordan-Normalform bauen

wobei λ_j die paarweise verschiedenen Eigenwerte sind mit ihren Vielfachheiten n_j . Dann annulliert auch das sogenannte Minimalpolynom

$$M_A(z) = \prod_{j=1}^k (z - \lambda_j)^{m_j} \quad (3.16)$$

mit dem soeben definierten m_j , also der Länge des längsten zugehörigen Jordanblocks, die Matrix A .

Es gibt kein Polynom mit kleinerem Grad, das die Matrix A annullieren würde. (Daher der Name.)

Beweis. Das ergibt sich direkt aus unseren Vorüberlegungen zusammen mit Lemma 151. \square

Außerdem ist jetzt klar, wie wir uns im konkreten Fall unsere *Jordan-Normalform* basteln können: Wir wenden sukzessive das Vorgehen aus Lemma 149 an. Zunächst am Beispiel:

Beispiel 164. Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 1 & 6 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 2 \end{pmatrix}.$$

- Wir berechnen nun das charakteristische Polynom. Da der obere linke Zweierblock ober Dreiecksgestalt hat, bekommen wir

$$\chi_A(\lambda) = \det(A - \lambda E) = (1 - \lambda)^2 \begin{vmatrix} 1 - \lambda & 0 \\ -2 & 2 - \lambda \end{vmatrix} = (1 - \lambda)^3 (2 - \lambda).$$

Damit hat der Eigenwert $\lambda_1 = 1$ die algebraische Vielfachheit 3, der Eigenwert $\lambda_2 = 2$ die algebraische Vielfachheit 1.

- Um den Eigenraum zu $\lambda_1 = 1$ zu bestimmen lösen wir das homogene LGS mit der Matrix

$$A - E = \begin{pmatrix} 0 & 1 & 6 & -2 \\ 0 & 0 & -3 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 1 \end{pmatrix}.$$

Um diese auf Treppenform zu bringen, tauschen wir die Zeilen zwei bis vier im Kreis

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

3 Die Jordansche Normalform

und ziehen das anderthalbfache der jetzigen zweiten von der dritten ab:

$$\begin{pmatrix} 0 & 1 & 6 & -2 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Diesen Vorgang merken wir uns, weil wir vermeiden wollen, höhere Potenzen von $(A - E)$ zu berechnen, und lieber inhomogene Gleichungssysteme mit der selben Matrix und einer geschickten rechten Seite lösen wollen.

Es ist leicht zu sehen, dass die Matrix Rang Drei hat und damit der Eigenraum Dimension Eins. Wir erhalten als Eigenvektor den ersten Standard Einheitsvektor $\mathbf{v}_1 = \mathbf{e}_1$. Wir brauchen also, um den Hauptraum aufzuspinnen, noch einen Hauptvektor zweiter und einen dritter Stufe. Dazu lösen wir als nächstes das inhomogene LGS $(A - E)\mathbf{x} = \mathbf{v}_1 = \mathbf{e}_1$. Die Lösungen liegen automatisch in $(\ker(A - E)^2 \setminus \ker(A - E)) \cap \varphi_{A-\lambda E}^{-1}\{\mathbf{v}_1\}$. Die obigen Zeilenoperationen haben keinen Einfluss auf die rechte Seite. Damit ist zu lösen

$$\left(\begin{array}{cccc|c} 0 & 1 & 6 & -2 & 1 \\ 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right)$$

Aus der zweiten und dritten Zeile schließen wir, dass die letzten beiden Komponenten verschwinden müssen. Damit muss nach der ersten Zeile $x_2 = 1$ gelten, während x_1 beliebig ist. Damit können wir für \mathbf{v}_2 den zweiten Standard Einheitsvektor \mathbf{e}_2 wählen.

Zur Berechnung des zugehörigen Hauptvektors dritter Stufe \mathbf{v}_3 nehmen wir nun $\mathbf{v}_2 = \mathbf{e}_2$ als rechte Seite und lösen $(A - E)\mathbf{x} = \mathbf{e}_2$. Die Lösungen liegen automatisch in $(\ker(A - E)^3 \setminus \ker(A - E)^2) \cap \varphi_{A-\lambda E}^{-1}\{\mathbf{v}_2\}$. Da wir auch auf diesen die obigen Zeilenoperationen anwenden müssen, lautet das zu lösende LGS

$$\left(\begin{array}{cccc|c} 0 & 1 & 6 & -2 & 0 \\ 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1/2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Nach der dritten Zeile ist $x_4 = 2$. Eingesetzt in die zweite Zeile bekommen wir damit $x_3 = 1$, danach aus der ersten $x_2 = -2$. Die erste Komponente ist wieder frei. Wir wählen also $\mathbf{v}_3 = (0, -2, 1, 2)^T$. Damit haben wir unseren dreidimensionalen Eigenraum beisammen. Man sieht auch leicht, dass das LGS $(A - E)\mathbf{x} = \mathbf{v}_3$ keine Lösung mehr besitzt.

- Um den Eigenraum zu $\lambda_2 = 2$ zu bestimmen lösen wir das homogene LGS mit der

Matrix

$$A - 2E = \begin{pmatrix} -1 & 1 & 6 & -2 \\ 0 & -1 & -3 & 2 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -2 & 0 \end{pmatrix}.$$

Wir erhalten als Lösung die lineare Hülle von $\mathbf{v}_4 = (0, 2, 0, 1)^T$. Da die algebraische Vielfachheit Eins war, ist nicht mehr zu erwarten.

- Bilden wir nun die Matrix $V = (\mathbf{v}_1 | \mathbf{v}_2 | \mathbf{v}_3 | \mathbf{v}_4)$, so wissen wir bereits, dass gilt

$$\begin{aligned} V^{-1}AV &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 6 & -2 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}. \end{aligned}$$

Das Berechnen der Inversen von V und das konkrete Ausmultiplizieren, kann man dann vielleicht noch zur Probe machen.

Das Allgemeine Vorgehen lautet also:

- Berechne das charakteristische Polynom und dessen Zerlegung in Linearfaktoren. Damit wissen wir bereits wie groß die Haupträume sind.
- Beginne mit dem ersten Eigenwert und berechne eine Basis des Eigenraums.
 - Ist die geometrische Vielfachheit des Eigenwerts λ gleich der algebraischen haben wir den Eigenwert damit abgearbeitet.
 - Ist sie kleiner, beginne mit dem ersten Eigenvektor \mathbf{v}_1 aus der Basis und suche nach den zugehörigen Hauptvektoren höherer Stufe:

Berechne die Lösung von $(A - \lambda E)\mathbf{x} = \mathbf{v}_1$. Diese ist automatisch eine Teilmenge von $(\ker(A - \lambda E)^2 \setminus \ker(A - \lambda E)) \cap \varphi_{A - \lambda E}^{-1}\{\mathbf{v}_1\}$.

Ist die Lösungsmenge leer, ist der Eigenvektor \mathbf{v}_1 abgearbeitet, und wir beginnen das selbe Spiel mit dem nächsten Eigenvektor.

Ansonsten wähle aus der Lösungsmenge einen beliebigen Vektor \mathbf{v}_2 aus.

Berechne nun die Lösung von $(A - \lambda E)\mathbf{x} = \mathbf{v}_2$, wiederum automatisch eine Teilmenge von $(\ker(A - \lambda E)^3 \setminus \ker(A - \lambda E)^2) \cap \varphi_{A - \lambda E}^{-1}\{\mathbf{v}_2\}$ und fahre sinngemäß fort.

3 Die Jordansche Normalform

- Haben wir damit noch keine Basis des Hauptraums, muss es noch einen weiteren Basisvektor zum Eigenraum geben, mit dem wir das Spiel ebenfalls treiben. Usw.
- Mache mit dem nächsten Eigenwert sinngemäß weiter. Achte darauf, die gewonnenen Eigen- bzw. Hauptvektoren durchlaufend zu nummerieren.
- Am Ende kann ich an der Rechnung sowohl die Transformationsmatrix für die Ähnlichkeitstransformation als auch die Jordan-Normalform selber ablesen.

Bemerkung 165. Bei einer Kette aus Eigenvektor (= Hauptvektor erster Stufe), zugehöriger Hauptvektor zweiter Stufe u.s.w. spricht man auch von einer *Jordan-Kette*. Damit ist offensichtlich die Länge des größten Jordanblocks zum Eigenwert λ_j auch die Länge der längsten Jordankette zu λ_j . Dass sich diese aus der Vielfachheit von λ_j als Nullstelle des Minimalpolynoms M_A ablesen lässt, hatten wir in Satz 163 schon gesehen.

4 Mathematische Grundbegriffe und algebraische Strukturen

Die Ziele dieses Kapitels sind:

- Sie sollen das Konzept der Verknüpfung als Verallgemeinerung der uns bisher bekannten Verknüpfungen verstehen, sodass Sie in der Lage sind, es an den verschiedensten Stellen wiederzuentdecken.
- Dabei sollen Sie auch verstehen, wie man durch die Betrachtung strukturhaltender Abbildungen aus einer algebraischen Struktur weitere bilden kann, indem man nach einer solchen faktorisiert, was zu einer kleineren Menge führt, oder auf den strukturhaltenden Abbildungen wieder algebraische Strukturen findet, womit man also auf einer größeren Menge unterwegs ist.

Dabei versuchen wir auch ein wenig, den Bogen zu den ganzen Zahlen zu schlagen, deren Verständnis gerade für Lehramtskandidat(inn)en sehr hilfreich ist.

4.1 Warum wir noch mehr formalisieren wollen

Im Rahmen der Übung haben wir bereits gesehen, dass es noch andere Dinge gibt, die wir gerne als Vektorraum betrachten würden, etwa die Polynome bis zu einem bestimmten Grad oder, noch besser, die reellen oder komplexen Polynome überhaupt. Wer parallel den Analysiskurs besucht, hat solche Dinge wie *Linearität* auch schon bei Grenzwerten von Folgen, Ableitungen und Integralen gesehen. Also würden wir auch gerne *Vektorräume von Folgen oder Funktionen* betrachten. Für den Zweck wäre es günstig, einen allgemeineren Begriff von Vektorraum zu haben als nur Unterräume von \mathbb{R}^n oder \mathbb{C}^n , vor allem einen, der auch *unendliche Dimension* erlaubt.

Interessant ist auch, dass wir mit \mathbb{R} und \mathbb{C} bereits zwei Objekte haben, in denen algebraisch gesehen die gleichen Regeln gelten. Der einzige Unterschied ist, dass in \mathbb{C} jedes nichtkonstante Polynom auch tatsächlich eine Nullstelle hat. Die Regeln sind aber auch die selben wie bei den rationalen Zahlen \mathbb{Q} , über denen wir noch gar keine Vektorräume betrachtet haben. Bei den ganzen Zahlen \mathbb{Z} dagegen fehlen uns ein paar dieser Regeln, bei den natürlichen $\mathbb{N} = \{1, 2, 3, \dots\}$ noch mehr. Die Frage stellt sich, wie man das alles formalisieren und in passende Klassen einteilen kann, sodass man nur noch sagen muss, um welche es sich handelt. Dabei werden wir für Dinge, die den Regeln von \mathbb{Q} , \mathbb{R} , \mathbb{C}

folgen den Namen *Körper* verteilen, für die „Zwillinge“ der ganzen Zahlen den Namen *Ring*. Wir werden noch ein paar weitere Namen einführen müssen, um alle spannenden Fälle abzuhandeln.

Nun ist es rein mit der Einteilung in algebraische Strukturen noch nicht getan. Beim Integrieren stellen wir z. B. fest, dass es außer der Nullfunktion noch weitere gibt mit $\|f\|_1 := \int_a^b |f(x)| dx = 0$ (für $a < b$. Möchte man das als Verallgemeinerung der Betragsummennorm (2.2) auf integrierbare Funktionen verwenden, so klappt es mit der Definitheit nicht). Man möchte nun gerne alle Funktionen f, g mit

$$\int_a^b |f(x) - g(x)| dx = 0$$

in ein Objekt zusammenfassen, das sich dann wieder als Element eines Vektorraums betrachten lässt. Dazu brauchen wir den Begriff der *Äquivalenzklassen*, der sich wiederum aus der Theorie der *Relationen* ergibt.

Doch zurück zu den Rechengesetzen. Wir kennen bereits aus der Behandlung der Zahlen in der Schule

das Assoziativgesetz

$$(a + b) + c = a + (b + c) \quad \text{bzw.} \quad (a \cdot b) \cdot c = a \cdot (b \cdot c);$$

das Neutralelement Null für die Addition sowie Eins für die Multiplikation, d. h., es gilt immer

$$a + 0 = 0 + a = a \quad \text{und} \quad 1 \cdot a = a \cdot 1 = a;$$

das inverse Element bei der Multiplikation $a^{-1} = 1/a$ mit $a^{-1} \cdot a = a \cdot a^{-1} = 1$ für jede reelle, komplexe, rationale Zahl a außer der Null sowie bei der Addition einfach $-a$;

das Kommutativgesetz $a + b = b + a$ bzw. $a \cdot b = b \cdot a$.

Außerdem kennen wir ein Gesetz, das Addition und Multiplikation miteinander verbindet, nämlich das *Distributivgesetz*

$$a \cdot (b + c) = a \cdot b + a \cdot c, \quad (a + b) \cdot c = a \cdot c + b \cdot c, \quad (4.1)$$

wobei die beiden Formeln nur gleichbedeutend sind, wenn das Kommutativgesetz gilt.

In Satz 2 der ELA konnten wir für den \mathbb{R}^n zunächst nur eine Addition definieren, die aber alle wünschenswerten Eigenschaften, Assoziativität, Existenz von neutralem Element und inversem Element sowie die Kommutativität gelten. Alles weitere ergab sich aus der Idee, an die Vektoren auch reelle Zahlen dran zu multiplizieren. Diese Idee nennt man *äußere Verknüpfung* damit wollen wir uns dann in Abschnitt 4.4 näher befassen.

4.2 Die mathematischen Grundbegriffe, die wir dafür brauchen

4.2.1 Mengen und Relationen

4.2.1.1 Allgemeines zu Mengen

Zu Mengen selber müssen wir nicht mehr viel sagen, weil Sie im ersten Semester bereits angefangen haben, damit umzugehen. Wir wollen aber nochmals die Regeln festhalten, die für Vereinigung und Schnitt gelten:

Assoziativgesetze:

$$(A \cap B) \cap C = A \cap (B \cap C) ,$$
$$(A \cup B) \cup C = A \cup (B \cup C) ,$$

Kommutativgesetze:

$$A \cap B = B \cap A ,$$
$$A \cup B = B \cup A ,$$

Distributivgesetze:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) ,$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) ,$$

Absorptionsgesetze:

$$A \cap (A \cup B) = A ,$$
$$A \cup (A \cap B) = A ,$$

wobei die letzteren eine einfache Folge der Distributivgesetze sind. Bemerkenswert ist, dass alles symmetrisch auftaucht, d. h. jedes Gesetz eine Entsprechung hat, bei der einfach Vereinigung und Schnitt vertauscht sind.

Definiert waren der *Schnitt* als

$$x \in A \cap B \iff x \in A \text{ und } x \in B$$

und die *Vereinigung* als

$$x \in A \cup B \iff x \in A \text{ oder } x \in B ,$$

wobei das logische „oder“ kein ausschließliches ist. Damit ließen sich alle Gesetze auf Gesetze der Logik zurückführen. Dies war der Anlass, dass irgendwann für das logische und

4 Mathematische Grundbegriffe und algebraische Strukturen

das Symbol \wedge und für das logische oder das Symbol \vee als eckige Varianten der Symbole für Schnitt und Vereinigung eingeführt wurden.

Neben Vereinigung und Schnitt kann man auch die *Differenz* bilden:

$$A \setminus B = \{x \mid (x \in A) \wedge (\neg(x \in B))\},$$

wobei das Symbol \neg die Negation bezeichnet, und die *symmetrische Differenz*

$$A \Delta B := (A \cup B) \setminus (A \cap B),$$

wobei letztere im Gegensatz zur ersteren kommutativ ist und noch weitere nette algebraische Eigenschaften hat, worauf wir noch zurückkommen werden (siehe auch Übung).

4.2.1.2 Wenn die Grundmenge ins Spiel kommt

Oft ist es nützlich, festzulegen, welches die *Grundmenge* sein soll, in der sich alles abspielt. Dann spielen auch nur noch ihre Teilmengen mit. Nennen wir die Grundmenge Ω , so können wir unserem Spielfeld den Namen *Potenzmenge* von Ω geben und schreiben

$$\mathcal{P}(\Omega) := \{A \mid A \subseteq \Omega\}.$$

Dann lässt sich auch das *Komplement* einer Menge sinnvoll erklären, nämlich

$$\bar{A} := A^c := \Omega \setminus A,$$

wobei wir uns zwei alternative Schreibweisen zurechtgelegt haben für den Fall, dass eine davon – meist die erste – wegen Mehrfachbelegung zu Missverständnissen führen könnte. Hierfür gibt es nun auch wieder bekannte und leicht zu beweisende Rechenregeln, nämlich

$$\overline{\bar{A}} = A \quad \forall A \in \mathcal{P}(\Omega)$$

und die einfachen *Regeln von de Morgan*

$$\overline{A \cap B} = \bar{A} \cup \bar{B} \quad \forall A, B \in \mathcal{P}(\Omega),$$

$$\overline{A \cup B} = \bar{A} \cap \bar{B} \quad \forall A, B \in \mathcal{P}(\Omega),$$

bei denen sich auch wieder die Dualität von Schnitt und Vereinigung zeigt. Die Gesetze gelten auch, wenn man mehr als zwei Mengen hat: Ist I eine beliebige Indexmenge (darf jede beliebige unendliche Größe haben) und $A_i \in \mathcal{P}(\Omega)$ für jeden Index $i \in I$, so gelten

$$\overline{\bigcap_{i \in I} A_i} = \bigcup_{i \in I} \bar{A}_i \quad \text{und} \quad \overline{\bigcup_{i \in I} A_i} = \bigcap_{i \in I} \bar{A}_i.$$

4.2.1.3 Kreuzprodukte und Relationen

Nun kann man aus einer oder mehreren Mengen auch neue Mengen basteln. Beim \mathbb{R}^n hatten wir ja schon sowas gemacht. Wir hatten n -mal die reellen Zahlen \mathbb{R} genommen und daraus Vektoren gebaut. Allgemeiner bildet man das *Kreuzprodukt* bzw. *kartesische Produkt* zweier Mengen durch

$$A \times B := \{(a, b) \mid a \in A, b \in B\}.$$

Für größere Ansammlungen von Mengen geht das analog, etwa

$$A_1 \times A_2 \times \cdots \times A_n := \{(a_1, a_2, \dots, a_n) \mid a_i \in A_i \text{ für } i = 1, 2, \dots, n\}.$$

Der Ursprung des Namens liegt in den kartesischen Koordinaten, bei denen ein Punkt in der Ebene oder im Raum auch durch ein *Tupel* von Koordinaten beschrieben wird. Man kann natürlich auch wieder das Spiel mit der Indexmenge spielen, was wir hier aber nicht vertiefen wollen. Möchte man für beliebige Indexmengen sicher stellen, dass das Kreuzprodukt von nichtleeren Mengen wieder nicht leer ist, so muss man das extra als Axiom einbauen, das sogenannte *Auswahlaxiom*.

Handelt es sich jedes mal um die gleich Menge, so vereinfachen wir uns die Notation

$$A \times A =: A^2, \quad A \times A \times A =: A^3, \dots,$$

wie wir das ja bereits bei \mathbb{R}^n und \mathbb{C}^n gemacht haben.

Was uns hier am meisten interessiert, ist der Fall von zwei Mengen.

Definition 166. Sind A und B Mengen, so heißt eine Menge $R \subseteq A \times B$ eine *Relation zwischen A und B* . Sind beide Mengen gleich, also $R \subseteq A^2$, so sprechen wir von einer *Relation in A* . Für $(a, b) \in R$ schreiben wir auch $a R b$, wobei wir für die meisten konkreten Fälle das R durch ein sinnfälliges Symbol ersetzen.

So lässt sich etwa ein Kreis durch die Relation für seine Koordinaten beschreiben:

$$R = \{(x, y) \in \mathbb{R}^2 \mid (x - x_0)^2 + (y - y_0)^2 = r^2\},$$

wobei (x_0, y_0) den Mittelpunkt und r den Radius des Kreises beschreibt.

Weitere Beispiele sind auch

Beispiel 167.

- Hat man eine Menge und ihre Potenzmenge, so beschreibt \in eine Relation zwischen diesen.
- In den natürlichen Zahlen ist die Teilbarkeitsrelation \mid definiert durch

$$a \mid b :\Leftrightarrow \exists m \in \mathbb{N} : a \cdot m = b.$$

4 Mathematische Grundbegriffe und algebraische Strukturen

- Auch Vergleiche wie \leq oder \subseteq sind im Grunde Relationen, im ersten Fall in \mathbb{R} oder einer Teilmenge davon, im zweiten Fall in der Potenzmenge einer Menge.
- Auch Parallelität \parallel zwischen Geradenpaaren oder Vektoren sowie Orthogonalität \perp , etwa in einem Innenproduktraum bilden Relationen.
- Die Menge aller Paare $(a, b) \in \mathbb{Z} \times \mathbb{Z}$ mit $a \equiv b \pmod{7}$, zu lesen als „ a ist kongruent zu b modulo 7.“, womit gemeint ist, dass a und b bezüglich 7 den selben Teilungsrest besitzen, ist die *Kongruenzrelation* modulo 7 in der Menge der ganzen Zahlen.

Im Folgenden beschränken wir uns auf Relationen *in* einer Menge. Um etwas Ordnung zu schaffen, geben wir besonderen Eigenschaften Namen:

Definition 168. Sei R eine Relation in A . Diese heißt

- *reflexiv*, wenn für jedes $a \in A$ gilt $a R a$,
- *antireflexiv* wenn kein $a \in A$ in Relation zu sich selber steht,
- *symmetrisch* wenn mit $a R b$ automatisch auch $b R a$,
- *antisymmetrisch* oder *identitiv*, wenn aus $a R b$ und $b R a$ automatisch folgt, dass $a = b$,
- *transitiv*, wenn mit $a R b$ und $b R c$ bereits folgt, dass $a R c$.

Zu beachten ist, dass der Wortbestandteil „Anti“ gerade nicht das logische Gegenteil bedeutet. Damit lassen sich die Relationen in Klassen einteilen, die eine bestimmte Auswahl dieser Eigenschaften erfüllen:

Definition 169. Ist eine Relation in einer Menge

- reflexiv, symmetrisch und transitiv, so nennen wir sie eine *Äquivalenzrelation*,
- reflexiv, antisymmetrisch und transitiv, so nennen wir sie eine *Ordnungsrelation*,
- antireflexiv und transitiv, so nennen wir sie eine *strenge Ordnungsrelation*,

Beispiel 170.

- \leq und \subseteq sind Ordnungsrelationen, $<$ und \subsetneq strenge Ordnungsrelationen. Es wäre zwar konsistenter gewesen für die echte Teilmenge das Symbol \subset zu verwenden. Leider wird das von manchen synonym zu \subseteq verwendet und ist damit „vergiftet“. Am besten verwendet man es wegen der unklaren Bedeutung gar nicht.
- Das selbe gilt für die Teilbarkeit und die echte Teilbarkeit.
- Die Parallelität ist eine Äquivalenzrelation.

4.2 Die mathematischen Grundbegriffe, die wir dafür brauchen

- Auch die oben beschriebene Kongruenzrelation modulo 7 in der Menge der ganzen Zahlen ist eine Äquivalenzrelation.
- Da die Orthogonalität nicht transitiv ist, ist sie weder eine Äquivalenz- noch eine Ordnungsrelation. Da der Nullvektor als einziger senkrecht auf sich selber steht, ist sie – außer im Nullraum – weder reflexiv noch antireflexiv. Außerdem ist sie (außer im Nullraum) weder symmetrisch noch antisymmetrisch.

Für die Algebra am wichtigsten sind die Äquivalenzrelationen. Für diese definieren wir die sogenannten *Äquivalenzklassen*

$$[a]_R := \{x \in A \mid x R a\} .$$

Wir nennen a einen *Vertreter* bzw. *Repräsentanten* der Klasse. Die Sprechweise erweist sich als sinnvoll wegen

Satz 171. *Ist A eine nichtleere Menge und R eine Äquivalenzrelation in A , dann gilt für $a, b \in A$:*

1. $[a]_R \neq \emptyset$.
2. $[a]_R = [b]_R$ genau dann, wenn $a R b$.
3. $[a]_R \cap [b]_R = \emptyset$ genau dann, wenn nicht $a R b$.
4. Es gilt $\bigcup_{a \in A} [a]_R = A$.

Beweis. Folgt recht schnell aus Reflexivität, Symmetrie und Transitivität (Übung). \square

Mittels einer Äquivalenzrelation lässt sich also eine Menge in disjunkte Klassen zerlegen. Das wird sich im Folgenden als sehr praktisch erweisen. Umgekehrt definiert natürlich jede disjunkte Zerlegung einer Menge eine Äquivalenzrelation auf dieser.

Beispiel 172. In $\mathbb{N}_0 \times \mathbb{N}_0$ lässt sich folgende Äquivalenzrelation definieren:

$$(a_1, a_2) \sim (b_1, b_2) :\Leftrightarrow a_1 + b_2 = a_2 + b_1 .$$

Bei genauer Betrachtung stellt man fest, dass es sich um Klassen differenzgleicher Paare handelt. Diese entsprechen wieder gerade den ganzen Zahlen, sodass wir auf diese Weise die ganzen aus den natürlichen zzgl. der Null konstruieren konnten.

Beispiel 173. Analog kann man in $\mathbb{N} \times \mathbb{N}$ Äquivalenzklassen quotientengleicher Paare finden, aus denen man die Bruchzahlen herleitet. Die zugehörige Äquivalenzrelation ist

$$(a_1, a_2) \sim (b_1, b_2) :\Leftrightarrow a_1 \cdot b_2 = a_2 \cdot b_1 .$$

Beispiel 174. In der ebenen Geometrie nennt man zwei Figuren (Mengen von Punkten) in der Ebene *kongruent*, wenn man die eine durch eine Bewegung (Spiegelung, Verschiebung, Drehung, Schubspiegelung) auf die andere abbilden kann. Die Kongruenz ist eine Äquivalenzrelation in der Menge aller Figuren der Ebene.

4.2.2 Abbildungen

Wir haben Abbildungen schon häufig verwendet, ohne sie extra formal zu definieren. Für uns war eine Abbildung $f : X \rightarrow Y$; $x \mapsto f(x)$ einfach etwas, was einem $x \in X$ genau ein $f(x) \in Y$ zuordnet. Im Rahmen von Relationen könnte man sie auch als *linkstotale*, *rechtseindeutige* Relation definieren, d. h. als eine Relation $f \in X \times Y$, bei der für jedes $x \in X$ die Menge $\{y \in Y \mid x f y\}$ genau ein Element enthält.

Auch die Begriffe *injektiv*, *surjektiv* und *bijektiv* hatten wir schon verwendet:

Definition 175. Eine Abbildung $f : X \rightarrow Y$ heißt

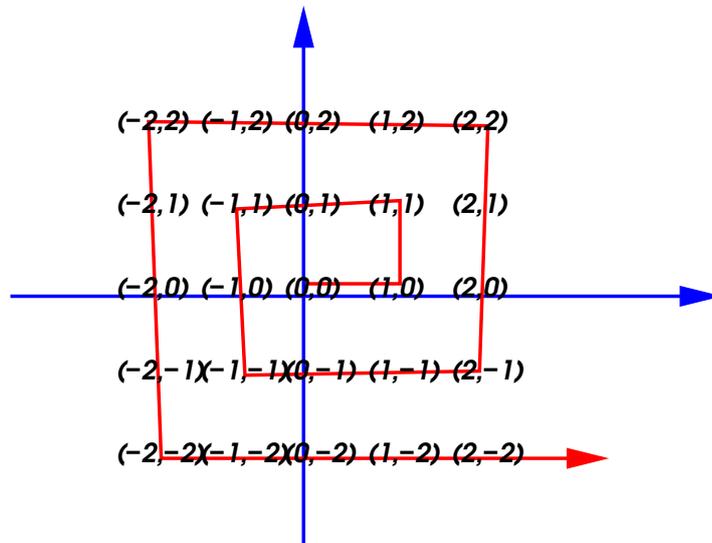
- *injektiv*, wenn sie als Relation auch linkseindeutig ist, d. h. je jedem $y \in Y$ gibt es *höchstens* ein $x \in X$ mit $f(x) = y$,
- *surjektiv*, wenn sie als Relation auch rechtstotal ist, d. h. je jedem $y \in Y$ gibt es *mindestens* ein $x \in X$ mit $f(x) = y$,
- *bijektiv*, wenn beides der Fall ist.

Wie man leicht sieht, lässt sich bei bijektiven Abbildungen auch eine Umkehrabbildung f^{-1} angeben. Im Allgemeinen schreibt man auch für das *Urbild* einer Menge $U \subseteq Y$

$$f^{-1}(U) = \{x \in X \mid f(x) \in U\},$$

Enthalten alle Urbilder einelementiger Teilmengen von Y genau ein Element, so ist die Abbildung bijektiv, bei „mindestens ein Element“ surjektiv und bei „höchstens ein Element“ injektiv.

Damit ist auch klar, dass eine bijektive Abbildung zwischen endlichen Mengen nur existieren kann, wenn beide gleich groß sind. Bei unendlichen Mengen kann man die Existenz bijektiver Abbildungen verwenden, um verschiedene unendliche Kardinalitäten zu definieren. Z. B. kann man die ganzen Zahlen leicht nummerieren, etwa $0, 1, -1, 2, -2, 3, -3, \dots$, was einer bijektiven Abbildung zwischen \mathbb{N} und \mathbb{Z} entspricht. Zeichnet man sich die Paare aus $\mathbb{Z} \times \mathbb{Z}$ ein ein Koordinatensystem, so kann man sie etwa (grob) spiralförmig nummerieren (s. u.), was bedeutet, dass sie auch nicht mehr sind als die natürlichen Zahlen und daher genau so viele wie die rationalen Zahlen. Bei den reellen Zahlen kann man zeigen, dass das nicht mehr klappt.



Beispiel 176. Bei einer endlichen Menge entspricht das Umsortieren der Elemente gerade einer bijektiven Selbstabbildung der Menge. Man nennt diese auch *Permutationen*. Der Bequemlichkeit halber nummerieren wir die Elemente unserer Menge einfach durch, sodass wir uns auf Mengen der Form $\{1, 2, \dots, n\}$ beschränken können. Dann können wir eine Permutation einfach durch die nummerierten Bilder angeben, etwa

$$(1, 3, 5, 4, 2) \quad \text{statt} \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 5 & 4 & 2 \end{pmatrix},$$

wo das Bild jeweils unter seinem Urbild steht. Permutationen, bei denen lediglich zwei Elemente vertauscht wurden, nennt man *Transpositionen*. Wie man sich leicht überlegt, kann man jede Permutation einer endlichen Menge als Hintereinanderausführung von Transpositionen schreiben. Braucht man dabei eine gerade Anzahl von Transpositionen, heißt die Permutation *gerade*, bei einer ungeraden Anzahl *ungerade*. Es lässt sich zeigen, dass sich jede Permutation eindeutig so einordnen lässt.

Da für das Bild der 1 alle n Elemente zur Auswahl stehen, danach für das der 2 nur noch $n-1$ u. s. w., gibt es genau $n! = 1 \cdot 2 \cdot \dots \cdot n$ Permutationen einer n -elementigen Menge.

Wir bezeichnen die Menge der Permutationen von $\{1, 2, \dots, n\}$ mit \mathcal{S}_n .

Beispiel 177. Nummeriert man bei einem Polygon die Ecken, so entsprechen seine kongruenten Selbstabbildungen gerade Permutationen der Ecken. Bei einem gleichseitigen Dreieck gibt es genau drei Symmetrieachsen, also drei Spiegelungen, und drei Winkel, um die man kongruent drehen kann, nämlich Null, ein Drittelkreis und zwei Drittelkreise. Das sind sechs Abbildungen, genau so viele, wie es Permutationen von $\{1, 2, 3\}$ gibt. Bei einem sonstigen gleichschenkligen Dreieck gibt es nur noch zwei, nämlich die Identität und eine Spiegelung. Ist das Dreieck nicht mal gleichschenklige, bleibt nur noch die Identität.

Es ist leicht zu sehen, dass sich bei mehr als drei Ecken niemals alle Permutationen der Ecken erreichen lassen, denn benachbarte Ecken müssen ja benachbart bleiben.

Beispiel 178. Bildet man von einer endliche Menge, etwa $\{1, 2, 3, \dots, n\}$ nach einer Menge A ab, so entspricht jede Abbildung gerade einem n -Tupel

$$(a_1, a_2, a_3, \dots, a_n) \in A^n$$

und damit die Menge

$$\{f \mid f : \{1, 2, 3, \dots, n\} \rightarrow A\}$$

gerade dem mehrfachen Kreuzprodukt A^n selber. Ersetzen wir die endliche Menge $\{1, 2, 3, \dots, n\}$ durch \mathbb{N} , so bekommen wir unendlich lange Tupel, sogenannte *Folgen* und schreiben entsprechend

$$\{f \mid f : \mathbb{N} \rightarrow A\} =: A^{\mathbb{N}}.$$

Ganz allgemein schreiben wir für die Menge der Abbildungen von B nach A einfach A^B .

Besonders wichtige Abbildungen sind die Verknüpfungen, zu denen wir im folgenden Abschnitt kommen.

Doch zuvor betrachten wir noch eine Strategie mit der man eine beliebige Abbildung als Hintereinanderausführung einer surjektiven und einer injektiven Abbildung darstellen kann:

Sei $f : X \rightarrow A$ eine Abbildung. Definiere nun in X eine Relation \sim durch

$$x \sim y :\Leftrightarrow f(x) = f(y).$$

Wie man leicht sieht, ist das eine Äquivalenzrelation. Schreibe nun $X/\sim := \{[x]_{\sim} \mid x \in X\}$, die Menge der so entstandenen Äquivalenzklassen, die *Faktorisierung von X nach \sim* bzw. der *Quotient/die Faktormenge von X bzgl. \sim* . Dann ist die *kanonische Abbildung*, die jedem $x \in X$ seine Äquivalenzklasse zuordnet, surjektiv. Aufgrund der Definition von \sim , ist aber die Abbildung

$$\begin{aligned} \bar{f} : X/\sim &\rightarrow A, \\ [x]_{\sim} &\mapsto f(x) \end{aligned}$$

zum einen wohldefiniert, schließlich wird jeder Äquivalenzklasse genau ein Bild zugeordnet, nämlich das, durch welches sie definiert ist. Zum anderen ist sie automatisch injektiv, da ja jedes $a \in f(X)$ genau eine Äquivalenzklasse als Urbild hat. Bezeichnen wir nun die kanonische Abbildung mit κ , so ist

$$f = \bar{f} \circ \kappa,$$

also die Hintereinanderausführung einer surjektiven und einer injektiven Abbildung.

Unser Ziel in den nächsten Abschnitten wird es unter anderem sein, zu zeigen, dass bei Strukturerthaltenden Abbildungen, wie wir sie ja von den linearen Abbildungen für Vektorräume kennen, sowohl κ als auch \bar{f} ebenso strukturerthaltend sind, und X/\sim die entsprechende algebraische Struktur von X erbt. Wir werden dabei auch auf Fälle stoßen, wo es mangels Definitheit auf X keine Norm bzw. kein Innenprodukt gibt, wo es aber eine einfache strukturerthaltende Abbildung gibt bzgl. der X/\sim plötzlich die Definitheit liefert.

Bemerkung 179. Manchmal schreibt man statt \sim auch den Namen der Abbildung oder etwas anderes, das die Äquivalenzrelation definiert unter den Schrägstrich, in unserem Fall etwa X/f .

Bemerkung 180. Hat man zuerst die Äquivalenzrelation gegeben, so kann man sich leicht eine passende Abbildung basteln bzgl. der man wieder die Äquivalenzrelation zurückbekommt.

4.3 Was es mit den Verknüpfungen auf sich hat

4.3.1 Eine einzelne Verknüpfung

Das typische an den Verknüpfungen $+$, \cdot , \cup , \cap , Δ , \circ , ... ist, dass sie aus zwei Zahlen, zwei Mengen, zwei Abbildungen etc. eine neue machen. Das lässt sich gut mit einer Abbildung beschreiben:

Definition 181. Sei M eine Menge. Eine Abbildung

$$\begin{aligned} \nu : M \times M &\rightarrow M \\ (a, b) &\mapsto a \nu b \end{aligned}$$

heißt (zweistellige) Verknüpfung in M .

- Ist (M, ν) eine Menge mit einer (zweistelligen inneren) Verknüpfung, so nennt man dieses ein *Magma* oder *Verknüpfungsgebilde*.
- Gilt in diesem Magma das Assoziativgesetz, so nennt man es eine *Halbgruppe*.
- Besitzt diese Halbgruppe ein neutrales Element, etwa $e \in M$ mit

$$e \nu m = m \nu e = m \quad \forall m \in M,$$

so nennt man sie ein *Monoid*.

- Besitzt in diesem Monoid jedes $m \in M$ ein inverses Element m^{-1} , d. h.

$$m \nu m^{-1} = m^{-1} \nu m = e,$$

so nennt man es eine *Gruppe*.

- Gilt in einer der vorgenannten algebraischen Strukturen das Kommutativgesetz, so nennt man sie *kommutativ*. Eine kommutative Gruppe nennt man auch *abelsche Gruppe*. (Nach Niels Henrik Abel, 1802–1829.)

Bemerkung 182. In einer endlichen Menge kann man eine Verknüpfung auch durch eine Verknüpfungstafel angeben, etwa

♠	a	π	♠
a	π	a	♠
π	♠	π	a
♠	a	π	♠

Damit liest man aus der ersten Zeile ab $a♠a = \pi$, $a♠\pi = a$, $a♠♠ = \pi$. Aus den anderen Zeilen analog.

Bemerkung 183.

- Üblicherweise sucht man sich für die Verknüpfung schönere Symbole als Kleinbuchstaben. Ist nur eine Verknüpfung im Spiel, und besteht keine Verwechslungsgefahr, lässt man das Verknüpfungssymbol einfach weg.
- In Anlehnung an unsere Notation bei der Multiplikation von Zahlen schreibt man (wobei wir das Verknüpfungszeichen gleich mal weglassen) in Halbgruppen

$$\underbrace{a a \dots a}_{n \text{ mal}} =: a^n .$$

Ist (M, ν) keine Halbgruppe, gilt also das Assoziativgesetz nicht, so ist der Ausdruck auf der linken Seite im allgemeinen nicht eindeutig.

Aber Vorsicht: Bei den natürlichen Zahlen mit der Addition als Verknüpfung, entspräche das genau $n \cdot a$, was wir nicht gerne als a^n schreiben wollen, weil wir die Schreibweise bei den Zahlen für die Multiplikation reserviert haben.

Wir können damit auch schon die ersten Sätze formulieren:

Satz 184. Sei $(M, *)$ eine Halbgruppe. Dann gelten

$$x^m * x^n = x^{m+n} \quad \forall x \in M, m, n \in \mathbb{N}, \quad (4.2)$$

$$(x^m)^n = x^{mn} \quad \forall x \in M, m, n \in \mathbb{N}. \quad (4.3)$$

Beweis. Jeweils mit Induktion über n (Übung), wobei zu beachten ist, dass wir keine Kommutativität vorausgesetzt haben. □

Satz 185. Ist $(M, *)$ eine kommutative Halbgruppe, dann gilt

$$x^n * y^n = (x * y)^n \quad \forall x, y \in M, n \in \mathbb{N}. \quad (4.4)$$

Beweis. Einfache Anwendung von Assoziativ- und Kommutativgesetz. □

Dass man dazu die Kommutativität braucht, ist der Grund, weshalb es bei der Multiplikation quadratischer Matrizen im Allgemeinen nicht klappt. Da musste man für sowas immer die Vertauschbarkeit der involvierten Matrizen voraussetzen.

Bemerkung 186. In $(\mathbb{R}, +)$ bedeutet Satz 185

$$nx + ny = n(x + y),$$

was bereits einen Spezialfall des Distributivgesetzes darstellt.

Beispiel 187. In der Menge $\mathcal{S}_3 = \{\text{id}, \sigma_1, \sigma_2, \tau_1, \tau_2, \tau_3\}$ mit¹

$$\begin{aligned} \text{id} &= \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}, & \sigma_1 &= \begin{pmatrix} 2 & 3 & 1 \end{pmatrix}, & \sigma_2 &= \begin{pmatrix} 3 & 1 & 2 \end{pmatrix}, \\ \tau_1 &= \begin{pmatrix} 1 & 3 & 2 \end{pmatrix}, & \tau_2 &= \begin{pmatrix} 3 & 2 & 1 \end{pmatrix}, & \tau_3 &= \begin{pmatrix} 2 & 1 & 3 \end{pmatrix}, \end{aligned}$$

der Permutationen gilt das Kommutativgesetz nicht. So ist etwa $\tau_1 \circ \tau_2 = \sigma_1$, aber $\tau_2 \circ \tau_1 = \sigma_2$. Daher klappt das in (\mathcal{S}_3, \circ) nicht mit Formel (4.4). Denn offenbar ist das Quadrat jeder Transposition die Identität, aber $(\tau_1 \circ \tau_2)^2 = \sigma_1^2 = \sigma_2$.

Beispiel 188. Wie wir in der ELA bereits festgestellt hatten, sind die Magmen $(\mathbb{R}^n, +)$, $(\mathbb{C}^n, +)$, $(\mathbb{R}^{n \times n}, +)$, $(\mathbb{C}^{n \times n}, +)$ und $(\text{hom}(V, V), +)$ (mit einem reellen oder komplexen Vektorraum V und der elementweisen Addition) abelsche Gruppen. Das selbe gilt natürlich auch für ihre Untervektorräume.

Außerdem sind $(\mathbb{R}^{n \times n}, \cdot)$, $(\mathbb{C}^{n \times n}, \cdot)$, $(\text{hom}(V, V), \circ)$ Monoide mit der Einheitsmatrix E bzw. der identischen Abbildung id als Neutralelement. Genau genommen ist $(\text{hom}(V, V), \circ)$ ein Untermonoid von (V^V, \circ) . Beschränken wir uns jeweils auf die invertierbaren Matrizen oder Abbildungen (als jeweiliges Untermonoid), so haben wir mit den inversen Matrizen bzw. Abbildungen jeweils eine Gruppe. Diese ist aber, wie wir schon in der ELA gesehen hatten, nicht kommutativ.

Beispiel 189. Ist M eine Menge, so bildet $(\mathcal{P}(M), \Delta)$ eine abelsche Gruppe, $(\mathcal{P}(M), \cap)$ ein kommutatives Monoid. (Siehe Übung.)

Genau genommen bildet auch $(\mathcal{P}(M), \cup)$ ein kommutatives Monoid. Aber das mit dem Schnitt ist im folgenden wichtiger.

Beispiel 190.

- Bei den Zahlbereichen finden wir $(\mathbb{N}, +)$ als kommutative Halbgruppe, $(\mathbb{N}_0, +)$ als kommutatives Monoid. Außerdem bilden (\mathbb{N}, \cdot) und (\mathbb{N}_0, \cdot) kommutative Monoide.
- Betrachten wir die ganzen Zahlen, so ist $(\mathbb{Z}, +)$ bereits eine abelsche Gruppe, (\mathbb{Z}, \cdot) lediglich ein kommutatives Monoid. Die einzigen Elemente, für die wir ein multiplikatives Inverses finden, sind das Neutralelement 1 und sein additives Inverses -1 .
- Bei \mathbb{Q} , \mathbb{R} und \mathbb{C} sieht es schon besser aus. Da kommen im Vergleich zu den ganzen Zahlen für alle multiplikative Inverse dazu außer für die Null. Nehmen wir die jeweils raus, haben wir auch bezüglich der Multiplikation eine abelsche Gruppe.

¹Die Buchstaben σ und τ stehen lautmalerisch für „Shift“ und „Transposition“, wobei bei den Verschiebungen der Index angibt, um wieviel im Kreis verschoben wird und bei den Transpositionen, welches Element stehen bleibt.

Beispiel 191. Wir können auch die *Polynome* über \mathbb{Q} , \mathbb{R} oder \mathbb{C} betrachten, etwa für \mathbb{R} :

$$\mathbb{R}[x] := \left\{ p = \sum_{k=0}^n a_k x^k \mid \text{alle } a_i \in \mathbb{R}, n \in \mathbb{N} \right\}.$$

Dabei darf im Gegensatz zu unseren bisherigen Übungsbeispielen der Grad des Polynoms beliebig hoch sein.

- Bei den hier genannten Beispielen entspricht jedem Polynom auch genau eine Polynomfunktion, im Reellen also

$$\begin{aligned} f_p : \mathbb{R} &\rightarrow \mathbb{R}, \\ x &\mapsto p(x), \end{aligned}$$

wovon wir bereits Gebrauch gemacht haben. Wir wollen aber unter einem Polynom den formalen Ausdruck verstehen, um uns offen zu halten, Polynome auch über Strukturen definieren zu können, bei denen diese Entsprechung nicht eins zu eins ist.

- Wenn wir hier wieder die üblichen Rechenoperationen ansetzen und mit x wie mit einer Zahl umgehen, bekommen wir $(\mathbb{R}[x], +)$ als abelsche Gruppe, $(\mathbb{R}[x], \cdot)$ als kommutatives Monoid. Für $\mathbb{Q}[x]$ und $\mathbb{C}[x]$ analog. Genau genommen klappt es sogar für $\mathbb{Z}[x]$.
- Beschränken wir allerdings künstlich den Grad nach oben, indem wir etwa maximal quadratische Polynome erlauben, so klappt das zwar noch mit der Addition, aber bzgl. der Multiplikation bekommen wir kein Untermonoid von $(\mathbb{R}[x], \cdot)$, weil schon das Quadrat eines quadratischen Polynoms einen zu hohen Grad hat. Die Verknüpfung führt uns aus der Menge heraus.

Beispiel 192. Sei $(M, \#)$ ein Magma. Dann kann man, wie wir das vom \mathbb{R}^n und der Addition kennen, auch M^n durch komponentenweise Anwendung der Verknüpfung zu einem Magma machen, für $n = 3$ also durch

$$(x_1, x_2, x_3) \# (y_1, y_2, y_3) := (x_1 \# y_1, x_2 \# y_2, x_3 \# y_3).$$

Ist $(M, \#)$ eine Halbgruppe, so überträgt sich das automatisch auf $(M^n, \#)$ bei Monoiden und Gruppen ebenso. Auch die Kommutativität überträgt sich auf natürliche Weise.

Wir können das Spiel sogar noch weitertreiben und statt M^n gleich allgemein M^A mit einer beliebigen Menge A betrachten. Dann ist eben $\#$ auf M^A definiert durch

$$(f \# g)(x) = f(x) \# g(x) \quad \forall x \in M.$$

Das Spiel funktioniert immer noch. Das macht man sich zu Nutze in der Analysis, wo man auf diese Weise die Funktionen auf einem Intervall $[a, b]$, also $\mathbb{R}^{[a,b]}$ durch punktweise Addition zu einer abelschen Gruppe und durch punktweise Multiplikation zu einem (kommutativen) Monoid macht. Die stetigen Funktionen $C[a, b]$ auf dem Intervall sind dann einfach bezüglich der Addition eine Untergruppe von $(\mathbb{R}^{[a,b]}, +)$ und bezüglich der Multiplikation ein Untermonoid von $(\mathbb{R}^{[a,b]}, \cdot)$.

Satz 193. *Gibt es in (M, ν) ein Neutralelement, so ist dieses eindeutig.*

Beweis. Seien e_1, e_2 Neutralelemente. Dann gilt

$$e_1 e_2 = e_2 e_1 = e_2 \quad \text{aber auch} \quad e_2 e_1 = e_1 e_2 = e_1 .$$

Damit müssen beide gleich sein. □

Satz 194. *Auch das inverse Element ist jeweils eindeutig.*

Beweis. Geht so ähnlich (Übung). □

Außerdem können wir schon strukturerhaltende Abbildungen definieren:

Definition 195. Seien (M, ν) und $(X, *)$ Magmen. Eine Abbildung $\varphi : M \rightarrow X$ heißt *Morphismus*, falls gilt

$$\varphi(a \nu b) = \varphi(a) * \varphi(b) \quad \forall a, b \in M .$$

Ab Halbgruppe aufwärts nennt man die Morphismen auch *Homomorphismen*.

Satz 196. *Ein Morphismus bildet Neutralelemente auf Neutralelemente und Inverse auf Inverse ab.*

Beweis. Übung. □

Satz 197. *Für Morphismen von Magmen gilt:*

1. *Die Komposition von Morphismen ist ein Morphismus.*
2. *Ist ein Morphismus bijektiv, so ist auch die Umkehrabbildung ein Morphismus. Man spricht dann von einem Isomorphismus.*
3. *Das Bild einer Halbgruppe unter einem Morphismus ist eine Halbgruppe, das Bild eines Monoids ein Monoid und das einer Gruppe eine Gruppe. Auch die Kommutativität vererbt sich auf diese Weise.*
4. *Ist (M, ν) ein Magma, so bildet die Menge aller Morphismen von (M, ν) nach (M, ν) bezüglich der Verkettung \circ ein Monoid.*
5. *Beschränkt man sich dabei auf die Isomorphismen, als Selbstabbildungen auch Automorphismen genannt, so bilden diese bereits eine Gruppe.*

Beweis. Übung. (Vieles überträgt sich einfach von den Beweisen aus Abschnitt 4.2 des ELA-Skripts.) □

Auch den Kern können wir definieren, brauchen dafür aber mindestens ein Monoid.

Definition 198. Seien (M, ν) und $(M', \#)$ Monoide mit Neutralelementen e bzw. e' sowie $\varphi : M \rightarrow M'$ ein Homomorphismus. Dann heißt

$$\ker \varphi := \varphi^{-1}(\{e'\})$$

der Kern von φ .

Im \mathbb{R}^n hatten wir für e und e' jeweils den Nullvektor.

Satz 199. Seien (M, ν) und $(M', \#)$ Monoide mit Neutralelementen e bzw. e' sowie $\varphi : M \rightarrow M'$ ein Homomorphismus. Dann ist $\ker \varphi$ ein Untermonoid von (M, ν) . Handelt es sich bei (M, ν) und $(M', \#)$ sogar um Gruppen, so ist der Kern eine Untergruppe von (M, ν) .

Beweis. Lässt sich direkt aus dem Beweis von Lemma 62 im ELA-Skript abschreiben. \square

Auch der Beweis von Lemma 63 aus der ELA lässt sich verwenden, allerdings nur, wenn es sich um einen Gruppenhomomorphismus handelt, schließlich benutzt der Beweis in den additiven Gruppen $V \subseteq \mathbb{R}^n$ und $W \subseteq \mathbb{R}^m$ additive Inverse:

Satz 200. Ein Gruppenhomomorphismus ist genau dann injektiv, wenn sein Kern trivial ist, d. h. mit den Bezeichnungen aus Definition 198 nur aus dem Neutralelement $e \in M$ besteht.

Jetzt bleibt uns noch, zu schauen, was passiert, wenn man nach einem Morphismus faktorisiert. Hier beginnen wir mit einem Beispiel:

Beispiel 201. Definiert man für ein $m \in \mathbb{N}$ die Menge $\mathbb{Z}_m := \{0, 1, 2, \dots, m-1\}$, so kann man damit den Teilungsrest einer ganzen Zahl z bzgl. des Moduls² m definieren. Es gibt dann nämlich eine eindeutige ganzzahlige Zerlegung der Form

$$z = d \cdot m + r \quad \text{mit} \quad d \in \mathbb{Z}, r \in \mathbb{Z}_m. \quad (4.5)$$

Haben zwei Zahlen $z, w \in \mathbb{Z}$ den selben Teilungsrest bzgl. des Moduls m , so schreibt man auch

$$z \equiv w \pmod{m}.$$

Wie man leicht sieht, ist „hat den selben Teilungsrest bzgl. m wie“ bzw. „ist modulo m kongruent zu“ eine Äquivalenzrelation in \mathbb{Z} .

- Nun kann man auf \mathbb{Z}_m eine Addition und eine Multiplikation dadurch definieren, dass man die Addition bzw. Multiplikation zunächst in \mathbb{Z} durchführt und das Ergebnis dann wieder auf seinen Teilungsrest bzgl. des Moduls m abbildet. Wir erhalten dadurch die Magmen $(\mathbb{Z}_m, +)$ und (\mathbb{Z}_m, \cdot) . Wie man leicht nachprüft, geht dabei keine der Rechenregeln aus den ganzen Zahlen verloren, wir haben also bezüglich der Addition eine abelsche Gruppe, bezüglich der Multiplikation ein kommutatives Monoid.

²Der Modul, die Moduln; kurzes betontes o, wie beim Wort Consul.

- Hier gewinnen wir sogar etwas dazu, denn durch unseren Kunstgriff, alles wieder auf den Teilungsrest zurückzuwerfen, haben wir jetzt mehr Elemente mit einem multiplikativen Inversen, für den Fall, dass m eine Primzahl ist, sogar alle außer der Null.³
- Definieren wir nun die Abbildung

$$\begin{aligned}\varphi : \mathbb{Z} &\rightarrow \mathbb{Z}_m \\ z &\mapsto \text{Teilungsrest von } z \text{ bzgl. } m,\end{aligned}$$

so finden wir leicht, dass dies sowohl bezüglich der Addition als auch bezüglich der Multiplikation ein Morphismus ist. Außerdem finden wir, dass

$$\mathbb{Z}/\varphi = \{[s]_m \mid s \in \mathbb{Z}_m\},$$

wobei die $[s]_m$ gerade die Äquivalenzklassen zur Kongruenzrelation modulo m sind. Da die zugehörige injektive Abbildung $\bar{\varphi} : \mathbb{Z}/\varphi \rightarrow \mathbb{Z}_m$ auf surjektiv ist, ist sie bijektiv und definiert auf \mathbb{Z}/φ eine Addition bzw. Multiplikation von Äquivalenzklassen.

Satz 202. Seien $(M, *)$ und $(M', \#)$ Magmen, $\varphi : M \rightarrow M'$ ein Morphismus und $\sim_\varphi = \{(x_1, x_2) \in M \times M \mid \varphi(x_1) = \varphi(x_2)\}$. Dann gelten:

1. \sim_φ ist eine Äquivalenzrelation in M .
2. $(\sim_\varphi, *)$ ist ein Untermagma von $(M \times M, *)$.
3. Auf M/φ ist durch $[x]_{\sim_\varphi} * [y]_{\sim_\varphi} = [x * y]_{\sim_\varphi}$ eine Verknüpfung definiert.

Beweis.

1. Wissen wir bereits aus Abschnitt 4.2.2.
2. Dass $M \times M$ selber ein Magma ist, hatten wir uns in Beispiel 192 überlegt. Seien also $(x_1, x_2), (y_1, y_2) \in \sim_\varphi$. Dann ist

$$(x_1, x_2) * (y_1, y_2) = (x_1 * y_1, x_2 * y_2). \quad (4.6)$$

Weil φ ein Morphismus ist, gilt aber auch

$$\varphi(x_1 * y_1) = \varphi(x_1) \# \varphi(y_1) = \varphi(x_2) \# \varphi(y_2) = \varphi(x_2 * y_2), \quad (4.7)$$

wobei wir beim mittleren Gleichheitszeichen die Äquivalenzrelation ausgenutzt haben. Damit ist dann aber auch $(x_1 * y_1) \sim_\varphi (x_2 * y_2)$, mit anderen Worten $(x_1, x_2) * (y_1, y_2) \in \sim_\varphi$.

³Wer sich näher mit den Geheimnissen der ganzen Zahlen und der Teilungsreste befassen möchte, sei auf [SS16] verwiesen.

3. Der Schluss des vorigen Beweisteils zeigt dass das Ergebnis der Verknüpfung in M/φ unabhängig von den Repräsentanten ist. Damit ist es eine Abbildung und somit tatsächlich eine Verknüpfung. □

Nun hätten wir gerne, dass sowas analog auch für Halbgruppen, Monoide und Gruppen gilt. Die Kommutativität hätten wir dann auch gerne auf $(\sim_\varphi, *)$ übertragen. Dazu schauen wir uns den Beweis nochmal genauer an. Ziel ist es, die jeweilige algebraische Struktur in M/φ wieder zu finden bzw. dort nachzuweisen.

Für Halbgruppen: Dass sich die Assoziativität bei einem Halbgruppenhomomorphismus überträgt, sieht man, indem man das Spiel aus Gleichungen (4.6) und (4.7) mit drei Paaren durchspielt. Damit gilt sie aber auch in M/φ .

Für Monoide: Ist $x_1 = e$ das Neutralelement in $(M, *)$, so sind $x_1, x_2 \in \ker \varphi$, und Gleichung (4.7) wird zu

$$\varphi(y_1) = \varphi(y_2) = \varphi(x_2 * y_2).$$

Damit bleibt aber die Äquivalenzklasse unter der Verknüpfung gleich. Es ist also $[e]_{\sim_\varphi}$ wie gewünscht ein Neutralelement in M/φ . (Von rechts geht es nämlich genau so.)

Für Gruppen: Hier nehmen wir an, dass $x_1 * y_1 = y_1 * x_1 = e$. Damit wird Gleichung (4.7) zu

$$e' = \varphi(e) = \varphi(x_2 * y_2).$$

Damit ist wie gewünscht $x_2 * y_2 \in [e]_{\sim_\varphi}$. Von rechts geht es genauso, und wir sind fertig.

Zum Kommutativgesetz: Dass sich das auf M/φ vererbt, ist nun offensichtlich.

Damit haben wir Beispiel 201 erfolgreich verallgemeinert. Man nennt M/φ auch, je nachdem *Quotientenmagma*, *Quotientenhalbgruppe*, *Quotientenmonoid* bzw. *Quotientengruppe*, wobei man auch davon spricht, dass man M nach φ faktorisiert hat, und deshalb auch statt „Quotienten-“ auch „Faktor-“ sagt.

4.3.2 Zwei Verknüpfungen in der selben Menge

Nun wissen wir von den Zahlen, dass es durchaus möglich ist, auf ein- und derselben Menge verschiedene (innere) Verknüpfungen zu definieren, etwa eine Addition und eine Multiplikation. Auch bei den Teilmengen einer Menge, d. h. auf der Potenzmenge $\mathcal{P}(A)$ einer beliebigen Menge A konnten wir das. Dort war allerdings der spannendere Fall der, dass wir zwei Verknüpfungen, nämlich Vereinigung und Schnitt, haben, die zueinander in gewisser Weise dual sind. Deshalb wollen wir uns die Verallgemeinerungen der Situation bei den Zahlen und bei den Mengen getrennt anschauen. Zunächst zur Verallgemeinerung der Situation bei den Zahlen:

4.3.2.1 Wei bei den Zahlen: Addition und Multiplikation

Hier hatten wir ja immer eine Addition und eine Multiplikation. Außer bei den natürlichen Zahlen waren die Anforderungen an die algebraische Struktur bei der Addition höher als bei der Multiplikation. Bei \mathbb{Z} , \mathbb{Q} , \mathbb{R} und \mathbb{C} , ja sogar bei Vektoren von solchen Zahlen oder auch quadratischen Matrizen, etwa $\mathbb{R}^{n \times n}$ hatten wir immer eine abelsche Gruppe. Bei der Multiplikation mussten wir dann die Anforderungen runterschrauben. Was wir immer hatten, war die Gültigkeit des *Distributivgesetzes* (4.1), das beide Verknüpfungen miteinander verbindet. *Dabei müssen wir immer daran denken, dass es auch im Fall fehlender Kommutativität bei der Multiplikation von links und von rechts gelten muss*, genau wie bei den Matrizen.

Definition 203. Seien auf der Menge M zwei innere Verknüpfungen $+$ und \cdot definiert, sodass $(M, +)$ eine abelsche Gruppe ist und das Distributivgesetz (4.1) gilt. Wir nennen das additive Neutralelement 0 das *Nullelement* bzw. die *Null* in M und $(M, +, \cdot)$

1. einen *Lie-Ring*⁴, falls das Magma (M, \cdot) keine weitergehenden Eigenschaften erfüllt,
2. einen (*assoziativen*) *Ring*, falls (M, \cdot) eine Halbgruppe bildet,
3. einen *Ring mit Eins*, falls (M, \cdot) ein Monoid ist (mit der Bezeichnung *Einselement* bzw. 1 für das multiplikative Neutralelement),
4. einen *Körper*, falls $(M, +, \cdot)$ ein Ring mit Eins ist und zusätzlich $(M \setminus \{0\}, \cdot)$ eine abelsche Gruppe.

Fehlt zu einem Körper nur die Kommutativität der Multiplikation spricht man noch von einem *Schiefkörper*. Gibt es in einem Ring keine *Nullteiler*, also keine $a, b \neq 0$ mit $a \cdot b = 0$, so heißt der Ring *nullteilerfrei* bzw. *integer*. Gilt in einem Ring das Kommutativgesetz für die Multiplikation, so bekommt er einfach das Attribut *kommutativ*.

Bemerkung 204. In einem Körper $(\mathbb{K}, +, \cdot)$ gelten grundsätzlich:

- \mathbb{K} hat mindestens zwei Elemente, nämlich die Null und die Eins. Mit $\{0, 1\}$ und der naheliegenden Rechenregel $1 + 1 = 0$ - schließlich ist die Eins der einzige potenzielle Kandidat für ihr eigenes additives Inverses - hat man auch schon einen Körper. Da der englische Begriff für einen Körper *field* ist, nennt man diesen \mathbb{F}_2 .
- Für alle $x \in \mathbb{K}$ gilt

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

Es ist nämlich

$$x^2 = x(x + 0) = x^2 + x \cdot 0 \quad \text{sowie} \quad x^2 = (x + 0)x = x^2 + 0 \cdot x .$$

⁴Sophus Lie, 1842-1899

4 Mathematische Grundbegriffe und algebraische Strukturen

Nun muss man nur noch $-x^2$ addieren. Der Beweis funktioniert auch noch für allgemeine Ringe.

Diese Eigenschaft sorgt auch dafür, dass (\mathbb{K}, \cdot) (man beachte, dass die Null noch drin ist) immer noch ein kommutatives Monoid ist.

- \mathbb{K} ist nullteilerfrei. Das folgt daraus, dass $(\mathbb{K} \setminus \{0\}, \cdot)$ eine Gruppe, also bezüglich der Multiplikation abgeschlossen ist.
- Für alle $x, y \in \mathbb{K}$ ist

$$x(-y) = (-x)y = -(xy), \quad (4.8)$$

wobei wir einer alten Gewohnheit folgend den Multiplikationspunkt weggelassen haben. Der Beweis ist eine nette Anwendung des Distributivgesetzes zusammen mit der obigen Feststellung, dass Multiplikation mit der Null die Null ergibt.

Die Sache mit dem kleinsten Körper \mathbb{F}_2 und seiner etwas gewöhnungsbedürftigen Rechenregel motiviert den folgenden Begriff.

Definition 205. Gibt es zu einem Körper eine natürliche Zahl $n \geq 2$ mit

$$\sum_{k=1}^n 1 = 0,$$

so nennt man das kleinste solche n die *Charakteristik* des Körpers. Andernfalls sagt man, der Körper habe die *Charakteristik Null*.

Damit hat der \mathbb{F}_2 die Charakteristik 2.

Beispiel 206 (Restklassenringe und -körper). Gehen wir zurück zu unserem Beispiel 201 mit den Restklassen \mathbb{Z}_m , so finden wir, dass unser \mathbb{F}_2 dort bereits unter dem Namen \mathbb{Z}_2 vorkam, wobei 0 für die geraden, 1 für die ungeraden Zahlen stand. Ganz allgemein sind die \mathbb{Z}_m mit $m = 2, 3, 4, \dots$ schöne Beispiele für kommutative Ringe mit Eins. Ist allerdings m keine Primzahl, so gibt es Nullteiler. Z. B. ist mit der Notation aus Beispiel 201

$$[2]_6 \cdot [3]_6 = [2 \cdot 3]_6 = [6]_6 = [0]_6.$$

Sowas passiert immer, wenn ich $m = d \cdot c$ mit $c, d \in \mathbb{N} \setminus \{1\}$ schreiben kann. Ist p eine Primzahl, kann man sogar zeigen, dass für jedes $d \in \mathbb{Z}_p \setminus \{0\}$ gilt

$$d^{p-1} = 1 \quad \Rightarrow \quad d \cdot d^{p-2} = d^{p-2} \cdot d = 1,$$

sodass d ein multiplikatives Inverses hat. Für eine Primzahl p ist also $\mathbb{Z}_p =: \mathbb{F}_p$ ein Körper. Wie man leicht sieht, hat er die Charakteristik p .

Beispiel 207 (Polynome). In Beispiel 191 hatten wir uns die Polynome $\mathbb{R}[x]$ angeschaut, wobei wir die reellen auch durch die rationalen, komplexen oder gar ganzen Zahlen ersetzen konnten.

- Da sie das Distributivgesetz von den Zahlen direkt erben, handelt es sich um kommutative Ringe mit Eins. Genau genommen, könnten wir statt \mathbb{R} jeden kommutativen Ring mit Eins R nehmen und darüber den Polynomring $R[x]$ bilden. Insbesondere geht das mit jedem Körper \mathbb{K} .
- Dabei fällt auf, dass unsere bei der Trigonalisierung von Matrizen betriebene Schlampligkeit nicht immer gut geht. Wir hatten nicht zwischen dem Polynom $p \in \mathbb{C}[x]$ und der *Polynomfunktion*

$$\begin{aligned} f_p : \mathbb{C} &\rightarrow \mathbb{C}, \\ z &\mapsto p(z) \end{aligned}$$

unterschieden. Bei Körpern mit Charakteristik Null geht das gut. Bei endlichen Körpern gar nicht. So finden wir für den kleinsten Körper

$$|\mathbb{F}_2^{\mathbb{F}_2}| = 4,$$

weil es von einer zweielementigen Menge nur vier verschiedene Abbildungen in eine zweielementige Menge geben kann. Aber Polynome gibt es natürlich unendlich viele, d. h.

$$|\mathbb{F}_2[x]| = \infty.$$

- Interessant ist, dass auch über \mathbb{F}_2 nicht alle Polynome eine Nullstelle haben, etwa die, deren zugehörige Polynomfunktion alles auf die Eins abbildet. Nun könnte man wieder wie beim Weg von den reellen zu den komplexen Zahlen neue künstliche Zahlen als Nullstellen dieser Polynome dazubauen, bis jedes Polynom vom Grad ≥ 1 , das man über dem jeweils erweiterten Körper formulieren kann wieder in Linearfaktoren zerfällt. Das ist aber ein unendlicher Prozess, dessen Theorie Sie sich im Rahmen einer allgemeinen Algebravorlesung zu Gemüte führen können.
- Die Verwandtschaft zu den ganzen Zahlen, ebenfalls ein kommutativer Ring mit Eins, lässt sich dahingehend nutzen, dass man den Algorithmus zu Division mit Rest (den wir aus der Grundschule kennen) übertragen kann. Man nimmt einfach die $1, x, x^2, x^3, \dots$ als Ersatz für die Einer-, Zehner-, Hunderter-, Tausender-, ... stelle. Das ganze nennt sich dann *Polynomdivision*.

Beispiel 208 (Quadratische Matrizen). Betrachten wir die quadratischen Matrizen $(\mathbb{R}^{n \times n}, +, \cdot)$, mit der üblichen Addition und Matrixmultiplikation, so bilden diese einen (nichtkommutativen) Ring mit Eins, nämlich der Einheitsmatrix. Wenn wir uns in Erinnerung rufen, wie wir in Abschnitt 4.1 des ELA-Skripts auf die Rechenregeln gekommen waren, so fällt auf, dass wir von \mathbb{R} nur die algebraischen Eigenschaften verwendet hatten, es also auch mit jedem anderen Körper geht. Was wir dabei gar nicht benutzt hatten war die Existenz von Kehrwerten, also multiplikativen Inversen, in \mathbb{R} . Wir können also über jedem Ring R die Matrixmultiplikation definieren. Um bei den quadratischen Matrizen einen Ring mit Eins zu bekommen, reicht es, dass R selber ein solcher ist, etwa \mathbb{Z} oder $\mathbb{Q}[x]$.

Definition 209 (Körper- und Ringhomomorphismen). Sind $(M, +, \cdot)$ und $(M', +, \cdot)$ (Lie-)Ringe oder Körper, und ist

$$\begin{aligned} \varphi : M &\rightarrow M', \\ x &\mapsto \varphi(x) \end{aligned}$$

sowohl bezüglich der Addition als auch bezüglich der Multiplikation ein (Homo)Morphismus, so heißt φ ein *Lie-Ring- Ring- bzw. Körperhomomorphismus*. Die Namensvorsätze *Endo-* (für $M' = M$), *Iso-* (für bijektive) und *Auto-* (für beides zusammen) verteilen wir wie gewohnt.

Beispiel 210. Wie man leicht nachrechnet bilden die Matrizen

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{mit } a, b \in \mathbb{R}$$

einen Körper. Aufgrund unserer Betrachtungen in Abschnitt 7.1 des ELA-Skripts sowie in den vorangehenden Kapiteln dieses Skripts ahnen wir bereits, dass dieser isomorph zu \mathbb{C} ist. Schreiben wir

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

so wissen wir bereits wie die Abbildungsvorschrift dazu lautet.

Beispiel 211. Betrachten wir die Abbildung

$$\begin{aligned} \phi : \mathbb{K}[x] &\rightarrow \mathbb{K}^{\mathbb{K}}, \\ p &\mapsto f_p, \end{aligned}$$

die jedem Polynom über dem Körper \mathbb{K} die zugehörige Polynomfunktion zuordnet, so ist diese ein Ringhomomorphismus, wenn man $\mathbb{K}^{\mathbb{K}}$ mit den elementweisen Verknüpfungen versieht. Da er für Charakteristik $\neq 0$ nicht injektiv ist, kann er also z. B. für $\mathbb{K} = \mathbb{F}_2$ kein Isomorphismus sein.

4.3.2.2 Wie bei den Mengen: zwei duale Verknüpfungen

Neben den Mengen mit \cap und \cup gibt es noch weitere Beispiele mit solchen dualen Verknüpfungen, etwa ggT und kgV bei den ganzen Zahlen oder min und max bzw. inf und sup bei den reellen Zahlen. (Vergleicht man zwei Zahlen als Verknüpfung sind min und $\in f$ sowie max und sup gleichbedeutend.)

Für solche Gebilde vergeben wir einen eigenen Namen:

Definition 212 (Verbände). Seien auf einer Menge L (wie englisch „Lattice“) zwei Verknüpfungen \wedge und \vee definiert.⁵

⁵Im Englischen heißen sie „meet“ (\wedge) und „join“ (\vee).

4.3 Was es mit den Verknüpfungen auf sich hat

- Sind beide Verknüpfungen assoziativ und kommutativ und gelten die *Absorptionsgesetze*

$$a \wedge (a \vee b) = a \quad \text{sowie} \quad a \vee (a \wedge b) = a \quad \forall a, b \in L,$$

so nennt man (L, \wedge, \vee) einen *Verband*.

- Gelten darüber hinaus die *Distributivgesetze*

$$\begin{aligned} a \wedge (b \vee c) &= (a \wedge b) \vee (a \wedge c), \\ a \vee (b \wedge c) &= (a \vee b) \wedge (a \vee c) \quad \forall a, b, c \in L, \end{aligned}$$

so heißt der Verband *distributiv*.

- Gibt es ein Element $n \in L$ mit

$$n \wedge a = n \quad \text{und} \quad n \vee a = a \quad \forall a \in L,$$

so heißt es *kleinstes Element* bzw. *Nullelement* in (L, \wedge, \vee) .

- Gibt es ein Element $e \in L$ mit

$$e \wedge a = a \quad \text{und} \quad e \vee a = e \quad \forall a \in L,$$

so heißt es *größtes Element* bzw. *Einselement* in (L, \wedge, \vee) .

- Ein Verband (L, \wedge, \vee) mit größtem und kleinstem Element heißt *komplementär*, wenn zu jedem $a \in L$ mindestens ein $b \in L$ existiert mit

$$a \wedge b = n \quad \text{und} \quad a \vee b = e.$$

Ein solches (nicht immer eindeutiges) b heißt *zu a komplementäres Element* (und umgekehrt).

Beispiel 213.

- Ist M eine Menge, so bildet $(\mathcal{P}(M), \cap, \cup)$ einen distributiven Verband mit Nullelement / kleinstem Element \emptyset und Einselement / größtem Element M . Das komplementäre Element \bar{A} zu einer Menge $A \subseteq M$ ist gerade ihr Komplement. Damit ist der Verband komplementär.
- Wie man im Grunde schon aus der Schule weiß, ist $(\mathbb{N}, \text{ggT}, \text{kgV})$ ein distributiver Verband. Das kleinste Element bzw. Nullelement ist die Eins. Ein größtes Element, also ein Einselement im Sinne des Verbands, gibt es nicht.
- (\mathbb{R}, \min, \max) bzw. (\mathbb{R}, \inf, \sup) ist ein distributiver Verband, der weder ein kleinstes noch ein größtes Element besitzt. Ersetzen wir \mathbb{R} durch ein abgeschlossenes Intervall, bekommen wir dessen Unter- und Obergrenze als kleinstes bzw. größtes Element. Ersetzen wir \mathbb{R} durch \mathbb{N} , bekommen wir die Eins als Nullelement, aber kein Einselement.

- Wie man leicht nachrechnet, ist der Schnitt zweier Untergruppen U, V einer Gruppe G wieder eine Gruppe (hatten wir schon bei den Schnitten von Vektorräumen nachgerechnet). Setzen wir

$$U \wedge V := U \cap V$$

und wählen für $U \vee V$ den Schnitt aller Untergruppen, die $U \cup V$ enthalten, so bekommen wir den sogenannten *Untergruppenverband* der Gruppe G .

Weitere Beispiele, auch solche aus der Geometrie, sowie erste Sätze für Verbände finden Sie etwa in [SS16].

4.4 Wenn äußere Verknüpfungen ins Spiel kommen

Um aus dem \mathbb{R}^n einen Vektorraum zu machen, brauchten wir neben der Addition noch eine Multiplikation mit reellen Zahlen, also eine Abbildung

$$\begin{aligned} \cdot : \mathbb{R} \times \mathbb{R}^n &\rightarrow \mathbb{R}^n, \\ (\lambda, (x_1, x_2, \dots, x_n)^T) &\mapsto (\lambda x_1, \lambda x_2, \dots, \lambda x_n)^T. \end{aligned}$$

Man nennt sowas eine *äußere Verknüpfung* auf \mathbb{R}^n und spricht davon, dass *der Körper \mathbb{R} auf der additiven abelschen Gruppe \mathbb{R}^n operiert*. Damit wir zufrieden waren und das ganze Vektorraum nennen konnten, mussten wir verlangen, dass die folgenden Rechengesetze erfüllt sind:

1. Für beliebige Skalare $r, s \in \mathbb{R}$ und Vektoren $\mathbf{v} \in \mathbb{R}^n$ gilt

$$r \cdot (s \cdot \mathbf{v}) = (r \cdot s) \cdot \mathbf{v}. \quad (\text{Assoziativgesetz der skalaren Multiplikation})$$

2. Für alle Skalare $r \in \mathbb{R}$ und Vektoren $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ gilt

$$r \cdot (\mathbf{v} + \mathbf{w}) = r \cdot \mathbf{v} + r \cdot \mathbf{w}. \quad (1. \text{Distributivgesetz})$$

3. Für alle Skalare $r, s \in \mathbb{R}$ und Vektoren $\mathbf{v} \in \mathbb{R}^n$ gilt

$$(r + s) \cdot \mathbf{v} = r \cdot \mathbf{v} + s \cdot \mathbf{v}. \quad (2. \text{Distributivgesetz})$$

4. Für alle Vektoren $\mathbf{v} \in \mathbb{R}^n$ gilt

$$1 \cdot \mathbf{v} = \mathbf{v}. \quad (\text{neutrales Element der skalaren Multiplikation})$$

Damit haben wir aber schon alles, was wir brauchen, um die wichtigsten algebraischen Strukturen mit äußeren Verknüpfungen zu beschreiben:

Definition 214. Sei $(M, +)$ eine abelsche Gruppe. Ersetze in den obigen Rechengesetzen jeweils \mathbb{R}^n durch M .

4.4 Wenn äußere Verknüpfungen ins Spiel kommen

1. Operiert der kommutative Ring mit Eins $(R, +, \cdot)$ so auf M , dass die obigen Rechengesetze mit R statt \mathbb{R} erfüllt sind, nennt man das ganze einen *unitären Modul* über R .
2. Operiert der Körper $(\mathbb{K}, +, \cdot)$ so auf M , dass die obigen Rechengesetze mit \mathbb{K} statt \mathbb{R} erfüllt sind, nennt man das ganze einen *Vektorraum* über \mathbb{K} .
3. Ist $(M, +, \cdot)$ sogar ein (Lie-)Ring und gilt zusätzlich zu den Anforderungen für einen Vektorraum noch die Rechenregel

$$\lambda(\mathbf{x} \cdot \mathbf{y}) = (\lambda\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (\lambda\mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in M, \lambda \in \mathbb{K}, \quad (4.9)$$

so sprechen wir von einer (Lie-)Algebra über \mathbb{K} .

Die Begriffe „integer“, „mit Eins“ etc. werden vom Ring auf die Algebra übertragen.

Beispiel 215.

- \mathbb{Z}^n wird mit den üblichen Operationen ein unitärer Modul über \mathbb{Z} .
- Ist B eine Menge, und lassen wir den Körper \mathbb{F}_2 in natürlicher (man sagt auch „kanonischer“) Weise auf dem Ring $(\mathcal{P}(B), \Delta, \cap)$ operieren, so erhalten wir eine kommutative Algebra mit Eins. Deren Unteralgebren $\mathcal{A} \subseteq \mathcal{P}(B)$ mit der Eigenschaft, dass jede abzählbare Vereinigung von Elementen von \mathcal{A} wieder in \mathcal{A} liegt, nennt man *σ -Algebren*.

Genauer lässt sich zeigen, dass jedes $\mathcal{A} \subseteq \mathcal{P}(B)$ mit

$$B \in \mathcal{A}, \quad (\sigma 1)$$

$$A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}, \quad (\sigma 2)$$

$$A_1, A_2, A_3, \dots \in \mathcal{A} \Rightarrow \bigcup_{k \in \mathbb{N}} A_k \in \mathcal{A} \quad (\sigma 3)$$

ein Unterring von ist und sich auch mit der Operation von \mathbb{F}_2 auf $(\mathcal{P}(B), \Delta, \cap)$ verträgt, also tatsächlich eine Unteralgebra ist, ohne das wir das im Vorhinein hätten fordern müssen.

- Die quadratischen Matrizen $\mathbb{K}^{n \times n}$ über einem Körper \mathbb{K} bilden über diesem bezüglich der üblichen Operationen eine (nichtkommutative) Algebra mit Eins.
- Jeder Körper lässt sich als Vektorraum über jedem seiner Unterkörper auffassen. So entspricht \mathbb{C} ja dem \mathbb{R}^2 . Wir können aber auch \mathbb{R} als Vektorraum über \mathbb{Q} betrachten.
- Wie erhofft ist nach obiger Definition jeder Unterraum des \mathbb{R}^n bzw. \mathbb{C}^n wieder ein Vektorraum.

4 Mathematische Grundbegriffe und algebraische Strukturen

- Ist $[a, b] \subseteq \mathbb{R}$ ein Intervall, so ist $\mathbb{R}^{[a,b]}$ mit den elementweisen Operationen, also

$$\begin{aligned}(f + g)(x) &= f(x) + g(x), \\(f \cdot g)(x) &= f(x) \cdot g(x), \\(\lambda f)(x) &= \lambda f(x) \qquad \forall x \in [a, b]\end{aligned}$$

über \mathbb{R} eine kommutative nullteilerfreie Algebra mit Eins. Die Menge $C[a, b]$ der stetigen Funktionen auf $[a, b]$ ist dann eine Unteralgebra hiervon. Man kann für diese Konstruktion auch allgemeinere kompakte Hausdorffmengen⁶ nehmen, was wir hier aber unterlassen, weil Sie soviel Topologie noch nicht gelernt haben. Allgemein spricht man bei Algebren von Funktionen mit den elementweisen Operationen von *Funktionalgebren*.

Bemerkung 216.

- Da Moduln wie etwa \mathbb{Z}^n eher in der Zahlentheorie interessant sind als in der linearen Algebra und analytischen Geometrie, etwa bei der Betrachtung ganzzahliger Gleichungssysteme, verfolgen wir sie hier nicht weiter.
- Da die quadratischen Matrizen über \mathbb{R} bzw. \mathbb{C} gerade den Endomorphismen von \mathbb{R}^n bzw. \mathbb{C}^n entsprechen, lassen wir die Algebren im Spiel, weil davon auszugehen ist, dass auch die Endomorphismen eines allgemeineren Vektorraums wieder eine Algebra bilden.

Definition 217. Damit ein Gruppen- bzw. Ringhomomorphismus $\varphi : M \rightarrow M'$ den Namen *lineare Abbildung* bzw. *Algebrenhomomorphismus* tragen darf, muss er zusätzlich die äußere Verknüpfung respektieren, also

$$\varphi(\lambda x) = \lambda \varphi(x) \quad \forall \lambda \in \mathbb{K}, x \in M.$$

Insbesondere müssen beide, M und M' Vektorräume bzw. Algebren über dem selben Körper \mathbb{K} sein.

Die Forderung wirkt zunächst etwas streng, hätte man doch vielleicht auf der rechten Seite statt des Skalars so etwas wie $\psi(\lambda)$ mit einem Körperhomomorphismus ψ verlangen können. Aber weil es sich um einen Gruppenhomomorphismus handelt, entspricht für eine natürliche Zahl n der Ausdruck $n \cdot x$ der n -fachen Summation von x , weshalb gelten muss

$$\varphi(n \cdot x) = \varphi\left(\sum_{i=1}^n x\right) = \sum_{i=1}^n \varphi(x) = n \cdot \varphi(x).$$

Bei einem Körper von Zahlen - auch bei den Restklassenkörpern \mathbb{F}_p und ihren Körpererweiterungen - ist das genau die Definition eingeschränkt auf natürliche Zahlen.

Da die quadratischen Matrizen über einem Körper eine Algebra bilden, und wir wissen, dass $\mathbb{R}^{n \times n}$ gerade die Endomorphismen von \mathbb{R}^n darstellt, werden wir mutig:

⁶Felix Hausdorff, 1868-1942

Satz 218. Die Endomorphismen $\text{hom}(V, V)$ eines Vektorraums V über einem Körper \mathbb{K} bilden mit der elementweisen Addition und der Hintereinanderausführung als Multiplikation sowie der elementweisen Multiplikation von Skalaren eine Algebra über \mathbb{K} .

Beweis. Wir basteln uns den Beweis u. a. aus Versatzstücken, die wir schon gezeigt hatten:

- Entsprechend dem zweiten Teil von Beispiel 192 ist $\text{hom}(V, V)$ mit der elementweisen Addition eine abelsche Gruppe.
- Nach Punkt 4 von Satz 197 ist $\text{hom}(V, V)$ bzgl. der Verkettung ein Monoid (als Untermonoid von V^V).
- Es gilt auch das Distributivgesetz, da für jedes $v \in V$

$$\begin{aligned} (\varphi \circ (\psi + \tau))(v) &= \varphi(\psi(v) + \tau(v)) \\ &= \varphi(\psi(v)) + \varphi(\tau(v)) \\ &= (\varphi \circ \psi)(v) + (\varphi \circ \tau)(v), \\ ((\psi + \tau) \circ \varphi)(v) &= (\psi + \tau)(\varphi(v)) \\ &= \psi(\varphi(v)) + \tau(\varphi(v)) \\ &= (\psi \circ \varphi)(v) + (\tau \circ \varphi)(v). \end{aligned}$$

Damit hätten wir schon mal einen Ring mit Eins.

- Die anderen vier Rechengesetze, die es braucht, um eine Algebra zu sein, rechnet man in der selben Weise nach.

□

Auch die Ergebnisse über Kern und Bild solcher Homomorphismen haben wir schon fast fertig:

Satz 219.

- Bild und Kern einer linearen Abbildung sind wieder Vektorräume.
- Bild und Kern eines Algebrenhomomorphismus sind wieder Algebren.

Beweis.

- Für lineare Abbildungen können wir den Beweis wie bei Satz 199 direkt aus dem zu Lemma 62 im ELA-Skript übernehmen.
- Rechnen müssen wir eigentlich nur bei den Produkten in der Algebra, etwa A . Seien $x, y \in \ker \varphi$. Dann ist

$$\varphi(x \cdot y) = \varphi(x) \cdot \varphi(y) = \mathbf{0} \cdot \mathbf{0}.$$

Nun hatten wir bereits festgestellt, dass auch in einem Ring die Multiplikation mit Null alles zu Null macht. Damit ist auch $x \cdot y \in \ker \varphi$. Für Produkte λx mit einem Skalar λ geht es analog.

Beim Bild können wir noch zusätzlich auf Satz 199 zurückgreifen.

□

Nun können wir uns daran machen, die Begriffe und Ergebnisse aus der ELA sukzessive zu übertragen, natürlich nur soweit das geht. Wir beginnen mit *Linearkombinationen* und *linearen Hüllen*. Allerdings ersetzen wir diesmal Mengen von Vektoren durch Systeme von Vektoren. Wir verwenden dafür Indexmengen I , die im Grunde aussehen dürfen wie sie wollen. Wir benutzen ja nur ihre Elemente als Indizes.

Definition 220. Sei V ein Vektorraum über dem Körper \mathbb{K} , I eine Indexmenge und $E = [\mathbf{v}_i \mid i \in I]$ ein System von Vektoren.

- Jede *endliche* Summe $\sum_{k=1}^K \lambda_{i_k} \mathbf{v}_{i_k}$ (alle $\lambda_{i_k} \in \mathbb{K}$) heißt *Linearkombination* über E .
- Die Menge aller solcher Linearkombinationen heißt die *lineare Hülle* $\langle E \rangle$ von E .
- Ist $U \subseteq V$ ein Unterraum und $\langle E \rangle = U$, so heißt E ein *Erzeugendensystem* von U .
- E heißt *linear unabhängig*, wenn sich der Nullvektor über E nur als triviale Linearkombination darstellen lässt.
- Ist $V = \langle E \rangle$ und E l. u., so heißt E eine *Basis* von V .
- Ist E ein EZS von V und verliert diese Eigenschaft, wenn ich irgend ein \mathbf{v}_i weglasse, so ist es ein *minimales* EZS.
- Ist E l. u. und verliert diese Eigenschaft, wenn ich irgend einen Vektor dazu packe, so ist es ein *maximales* l. u. System.

Damit überträgt sich Satz 37 aus der ELA direkt:

Satz 221. Sei V ein VR und B ein System von Vektoren in V . Dann sind folgende Aussagen äquivalent:

- (i) B ist eine Basis von V .
- (ii) B ist ein minimales Erzeugendensystem von V .
- (iii) B ist ein maximales l. u. System in V .
- (iv) Jeder Vektor $\mathbf{v} \in V$ lässt sich eindeutig als LK über B schreiben.

Beweis. Wir hatten zwar in der ELA mit Mengen statt Systemen gearbeitet, aber nirgends wirklich spezielle Eigenschaften des \mathbb{R}^n ausgenutzt, sondern nur die algebraischen Eigenschaften von Vektorräumen. □

Definition 222. Besitzt der Vektorraum V ein endliches Erzeugendensystem, so heißt er *endlichdimensional*.

Bemerkung 223.

- Wir können nun alle Resultate aus den Kapiteln 1-6 der ELA, für die wir ein endliches Erzeugendensystem vorausgesetzt hatten auch für allgemeine endlichdimensionale Vektorräume in Anspruch nehmen. Schließlich haben wir dort keine speziellen Eigenschaften der reellen Zahlen benutzt, sondern nur ihre algebraischen als Körper. In Kapitel 5 muss man natürlich die Determinante als Abbildung $\mathbb{K}^{n \times n} \rightarrow \mathbb{K}$ schreiben. Sonst klappt es nicht.
- Die anderen Resultate gelten sogar für alle Vektorräume.
- Um die Existenz einer Basis für beliebige Vektorräume zu zeigen, muss man in der Mengentheorie das etwas unanschauliche Auswahlaxiom ansetzen. Dann geht es mit der sogenannten transfiniten Induktion. Das würde aber den Rahmen unseres Kurses sprengen.
- Das mit der Darstellung der linearen Abbildungen durch Matrizen klappt nur im Endlichdimensionalen richtig.
- Das Eigenwertproblem lässt sich für beliebige Vektorräume betrachten. Allerdings muss man bei anderen Körpern als \mathbb{R} meist mehr als ein Element dazu bauen, um einen algebraisch abgeschlossenen Körper zu bekommen, im allgemeinen sogar unendlich viele.

Genaugenommen haben wir das Eigenwertproblem auch in der ELA schon auf einem unendlichdimensionalen Vektorraum betrachtet:

Die Menge

$$V := \{f \in C^\infty([a, b]) : f(a) = f(b) = 0, f^{(k)}(a) = f^{(k)}(b) = 0 \text{ für gerades } k\}$$

aus Beispiel 109 der ELA ist tatsächlich ein Unterraum des Vektorraums $C^\infty[a, b]$ der beliebig oft stetig differenzierbaren Funktionen auf dem Intervall $[a, b]$. Dieser ist wiederum ein Unterraum von $\mathbb{R}^{[a, b]}$. Die Abbildung

$$-\frac{d^2}{dx^2} : V \rightarrow V, \quad f \mapsto -f'',$$

die eine Funktion auf das Negative ihrer zweiten Ableitung abbildet ist nach den Ableitungsregeln linear. Der Operator $-\frac{d^2}{dx^2}$ ist somit ein Endomorphismus von V .

Die Physiker unter Ihnen haben vielleicht schon bemerkt, dass es sich hier um das *Problem der schwingenden Saite* handelt, das einfachste *Schwingungsproblem* mit einer räumlichen Komponente. Die Beschreibung von Schwingungen von Membranen ist im Zweidimensionalen die einfachste und lässt sich bei einem rechteckigen Gebiet ähnlich lösen. Für komplexere Formen, Platten mit echter Dicke und dreidimensionale Bauteile,

wie man sie im Ingenieurwesen untersucht, sieht es etwas komplizierter aus. Es bleibt aber immer ein Eigenwertproblem.

Für die Approximation durch ein endlichdimensionales Problem verwendet man dann oft nicht mehr *Finite Differenzen*, wie wir sie in der ELA benutzt haben, sondern sogenannte *Finite Elemente*.

Wir betrachten noch ein paar Beispiele:

Beispiel 224.

- Die Polynome $\mathbb{K}[x]$ bilden eine Algebra über \mathbb{K} . Betrachtet man diese als Vektorraum, so findet man leicht die Monombasis

$$1, x, x^2, x^3, \dots$$

Man kann aber auch, wenn man einen unendlich großen Körper und darin eine Folge $x_0, x_1, x_2, x_3, \dots$ paarweise verschiedener Punkte vorgegeben hat, die *Newton-Basis* bilden:

$$1, (x - x_0), (x - x_0)(x - x_1), (x - x_0)(x - x_1)(x - x_2), \dots$$

- Beschränken wir den Grad der Polynome, etwa

$$\Pi_n := \{p \in \mathbb{K}[x] \mid \deg p \leq n\},$$

was auf einen $n + 1$ -dimensionalen Vektorraum führt, dann können wir bei den angeführten Basen von $\mathbb{K}[x]$ einfach die ersten $n + 1$ nehmen, die ja den passenden Grad haben, oder die *Lagrange-Basis* zu x_0, \dots, x_n mit

$$L_j(x) = \frac{\prod_{\substack{s=0 \\ s \neq j}}^n (x - x_s)}{\prod_{\substack{s=0 \\ s \neq j}}^n (x_j - x_s)} = \prod_{\substack{s=0 \\ s \neq j}}^n \frac{x - x_s}{x_j - x_s}.$$

Für die zugehörigen Polynomfunktionen gilt dann

$$L_j(x_k) = \delta_{j,k}.$$

Hat \mathbb{K} die Charakteristik Null, so ist diese Basis im Raum der zugehörigen Polynomfunktionen quasi die Inkarnation der Standardbasis.

Nun kommen wir zurück zu den linearen Abbildungen. Auch hier können wir die Faktorisierung nach dem Kern durchführen. In Satz 202 haben wir bereits gezeigt, dass sich beim Faktorisieren eines Morphismus φ nach der Relation $\sim_\varphi = \{(x, y) \in M \times M \mid \varphi(x) = \varphi(y)\}$

4.4 Wenn äußere Verknüpfungen ins Spiel kommen

die Verknüpfung auf die Äquivalenzklassen „vererbt“. Im Anschluss daran hatten wir verifiziert, dass sich auch die algebraischen Eigenschaften vererben. Nun hätten wir noch gerne, dass das bei linearen Abbildungen auch für die skalare Multiplikation von außen gilt. Wir betrachten also mit angepasster Notation $(\mathbf{u}, \mathbf{v}) \in \sim_\varphi \subseteq V \times V$ für unseren Vektorraum V . Wegen der Linearität, die wir nun für φ voraussetzen, ist dann für jedes $\lambda \in \mathbb{K}$

$$\varphi(\lambda \mathbf{u}) = \lambda \varphi(\mathbf{u}) = \lambda \varphi(\mathbf{v}) = \varphi(\lambda \mathbf{v}),$$

also auch $(\lambda \mathbf{u}, \lambda \mathbf{v}) \in \sim_\varphi$. Damit haben wir die Operationen komplett an die Menge der Äquivalenzklassen vererbt und bezeichnen diesen neuen Vektorraum als V/φ .

Bemerkenswert ist das folgende:

Satz 225. *Ist bezüglich obiger Äquivalenzrelation $\mathbf{u} \sim_\varphi \mathbf{v}$, so gilt für deren Differenz*

$$\mathbf{u} - \mathbf{v} \in \ker \varphi .$$

Beweis. Das ist so offensichtlich, dass wir es schon bei den Gruppen gemacht haben könnten. \square

Um unsere nächste Konstruktion zu rechtfertigen brauchen wir den allgemeinen Basisergänzungssatz:

Satz 226 (Allgemeiner Basisergänzungssatz). *Ist V ein beliebiger Vektorraum und A ein linear unabhängiges System (oder eine linear unabhängige Menge) in V . Dann gibt es eine Basis B von V , die A als Teilsystem (bzw. Teilmenge) enthält.*

Beweis. Braucht wie der Existenzsatz für die Basis ebenfalls transfiniten Induktion, weshalb wir ihn weglassen. \square

Wir erinnern uns, dass wir bei Satz 66 der ELA zwar ein endliches Erzeugendensystem vorausgesetzt hatten, im Beweis aber nur verwendet hatten, dass sich jedes $\mathbf{v} \in V$ als endliche LK darüber schreiben lässt. Damit gilt auch für allgemeine Vektorräume:

Satz 227. *Sind V, W Vektorräume, und ist B eine Basis von V , dann ist eine lineare Abbildung $\varphi : V \rightarrow W$ durch die Bilder der Basisvektoren aus B eindeutig festgelegt. Insbesondere existiert eine solche Abbildung.*

Damit funktioniert die folgende Konstruktion:

- Wir wählen uns aus unserem Vektorraum V einen Unterraum U und dort eine Basis A .
- Diese ergänzen wir nach dem allgemeinen Basisergänzungssatz 226 zu einer Basis B von V .
- Nun definieren wir uns nach Satz 227 eine lineare Abbildung φ_U , die auf A verschwindet, auf den restlichen Basisvektoren nicht.

4 Mathematische Grundbegriffe und algebraische Strukturen

- Damit ist $\ker \varphi_U = U$.
- Nun faktorisieren wir unseren großen Vektorraum V nach φ_u und schreiben

$$V/U := V/\varphi_U.$$

Dieser enthält dann die Äquivalenzklassen

$$[\mathbf{w}]_{\sim_\varphi} = \mathbf{w} + U := \{\mathbf{w} + \mathbf{u} \mid \mathbf{u} \in U\}.$$

Definition 228 (Faktorraum/Quotientenraum). Der so gewonnene Vektorraum V/U heißt der *Quotientenraum* bzw. *Faktorraum* von V nach U .

Beispiel 229 (Lebesgue-Räume). Ein wichtiges Beispiel für diese Faktorisiererei wird Ihnen in der Analysis 2 begegnen. Dort führt man den Lebesgueschen Integralbegriff ein, der dazu führt, dass jede Menge Funktionen integrierbar sind, von denen man das sonst nie gedacht hätte. Ist z. B.

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q}, \\ 0 & \text{sonst,} \end{cases}$$

so ergibt sich

$$\int_0^1 \chi_{\mathbb{Q}}(x) dx = 0.$$

Da die Nullfunktion aber das gleiche Integral hat, lässt sich damit keine Norm stricken, weil es mit der Definitheit nicht klappt. Nun ist aber durch $\int_0^1 f(x) dx$ auf dem Vektorraum der über $[0, 1]$ Lebesgue-integrierbaren Funktionen eine lineare Abbildung in die reellen Zahlen, also ein lineares Funktional definiert. Faktorisiert man den Raum der über $[0, 1]$ Lebesgue-integrierbaren Funktionen nach diesem Funktional, so bekommt man einen Raum, in dem es tatsächlich nur ein Element gibt, für das $\int_0^1 [f]_{\sim}(x) dx = 0$. Damit kann man dort tatsächlich eine Norm definieren (als Verallgemeinerung der Betragssummennorm).

4.5 Allgemeine Innenprodukträume

In Kapitel 2 hatten wir auf Unterräumen von \mathbb{R}^n bzw. \mathbb{C}^n das *Innenprodukt* und die *Norm* eingeführt. Ganz offensichtlich funktionieren die Definitionen für allgemeine Vektorräume über \mathbb{R} bzw. \mathbb{C} genauso. Damit haben wir bereits die allgemeinen *Innenprodukträume* definiert. Auch die Begriffe *orthogonal* und *Orthonormalsystem* sind damit klar. Nun müssen wir lediglich ausfiltern, welche Ergebnisse von der Existenz einer endlichen Basis Gebrauch gemacht hatten und welche nicht. Da stoßen wir zunächst auf die Sache mit den *Orthonormalbasen*. Diese sollten im Endlichdimensionalen Basen sein. Im Unendlichdimensionalen verlangt man das aus praktischen Gründen nicht mehr:

Definition 230 (Allgemeine Orthonormalbasis). Sei $V, (\cdot, \cdot)$ ein Innenproduktraum. eine Teilmenge $S \subseteq V$ heißt *Orthonormalbasis* von V , wenn

- für je zwei $e, q \in S$ gilt

$$(e, q) = \begin{cases} 1 & \text{falls } e = q, \\ 0 & \text{sonst,} \end{cases}$$

- und zu jedem $v \in V$ gibt es eine Folge $\lambda_1, \lambda_2, \dots$ von Skalaren sowie eine Folge q_1, q_2, \dots in S mit

$$v = \sum_{k=1}^{\infty} \lambda_k q_k.$$

Am Symbol ∞ sehen Sie bereits, dass dieses in die Analysis führt, weshalb wir die unendlichen Orthonormalbasen nicht weiter verfolgen. Wir wollen lediglich festhalten, dass in einigen Fällen eine abzählbare ONB existiert, die wir mittels *Gram-Schmidt* „berechnen“ können.

Beispiel 231 (Reelle Polynomfunktionen). Im Reellen gehört zu jedem Polynom $p \in \mathbb{R}[x]$ genau eine Polynomfunktion, die wir einfach mit dem selben Buchstaben bezeichnen. Betrachten wir die Polynomfunktionen über dem Intervall $[-1, 1]$, so ist dort durch

$$(p, q) = \int_{-1}^1 f(x) \cdot g(x) dx \quad (4.10)$$

ein Innenprodukt definiert. Beginnen wir nun mit der Monombasis $1, x, x^2, \dots$ und orthogonalisieren diese nach Gram-Schmidt, so erhalten wir die *Legendre-Polynome*. Ersetzen wir das Innenprodukt durch

$$(p, q) = \int_{-1}^1 f(x) \cdot g(x) \cdot \frac{1}{1-x^2} dx,$$

so erhalten wir auf die gleiche Weise die *Tschebyschow-Polynome* erster Art. Es gibt noch weitere Möglichkeiten, solche *Orthogonalpolynome* zu konstruieren. Meist versieht man sie allerdings so mit einem Vorfaktor, dass sie in Eins den Funktionswert Eins haben.

Die Faktorräume vom Ende des vorigen Abschnitts kommen auch hier ins Spiel. Wenn der Raum „zu groß“ ist, um darauf ein hübsches Innenprodukt zu definieren, meist, weil es keine Definitheit gibt, kann man wieder geeignet faktorisieren:

Beispiel 232 (Der Raum L^2). Das Innenprodukt (4.10) lässt sich auch für allgemeinere Funktionen hinschreiben, deren Quadrat über $[0, 1]$ Lebesgue-integrierbar ist. Diese bilden den Vektorraum $\mathcal{L}^2([0, 1])$. Das Problem ist, dass wir wie in Beispiel 229 keine Definitheit bekommen. Wir faktorisieren diesen wieder nach dem linearen Funktional $\int_0^1 f(x) dx$ und

bezeichnen den Faktorraum als $L^2([0, 1])$. Auf können wir ein Innenprodukt definieren durch

$$([f], [g]) = \int_0^1 f(x) \cdot g(x) dx . \quad (4.11)$$

Dieses ist unabhängig von den gewählten Repräsentanten in der jeweiligen Äquivalenzklasse. Damit die konstante Funktion 1 auch Norm Eins bekommt, fügt man bei anderen Integrationsgebieten üblicherweise noch einen Normierungsfaktor ein.

Wie man an den Beispielen sieht, kommt man bei allgemeinen Innenprodukträumen ganz schnell in den Bereich der Analysis, weshalb wir die Betrachtung an dieser Stelle kurz halten. Wir halten lediglich die folgenden Merkgeln fest:

Bemerkung 233 (Merkgeln).

- Alle Begriffe und Resultate aus Kapitel 2, die eine endliche Basis oder ONB verwenden, gelten auch in anderen endlichdimensionalen Innenprodukträumen.
- Haben wir eine solche endliche ONB in einem n -dimensionalen Innenproduktraum vorgegeben, lässt sich wie in Abschnitt 2.7 alles wieder auf das Standardskalarprodukt im \mathbb{R}^n bzw. \mathbb{C}^n zurückspielen.
- Die Begriffe und Resultate, bei denen wir keinen Gebrauch von der endlichen Dimension gemacht haben, gelten für alle Innenprodukträume.
- Weil diese aber mit analytischen Konzepten behandelt werden müssen, sprengen sie den Rahmen dieses Kurses.

5 Analytische Geometrie: Wir binden die Vektoren an Punkten fest

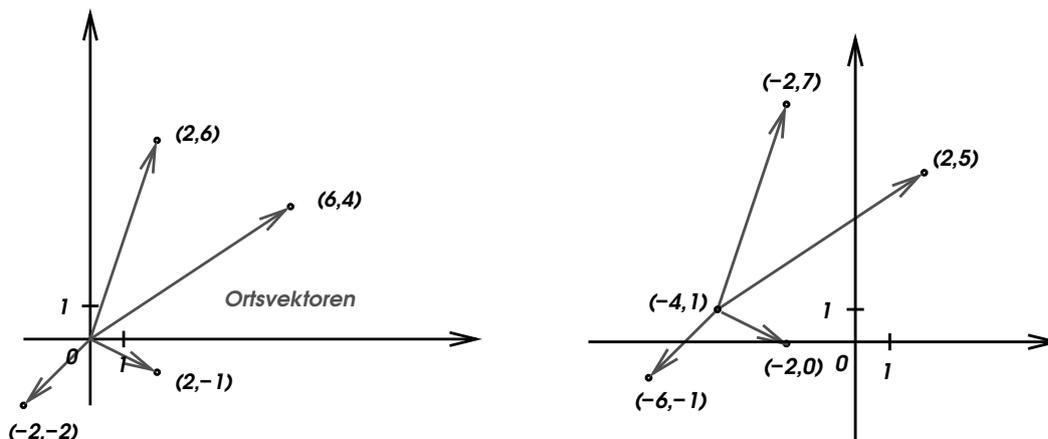
Wir können die analytische Geometrie in diesem Kurs nur kurz streifen. Es sollen aber die Dinge vorkommen, die Ihnen am ehesten im weiteren Verlauf des Studiums begegnen werden bzw., die Sie im Lehramt benötigen. Dabei sei für Lehramtskandidaten auch auf den Kurs zur Elementargeometrie verwiesen, in dem auch Teile der analytischen Geometrie mitbehandelt werden.

Für alle, die es genauer wissen wollen, sei auf das Lehrbuch [Fis92] verwiesen, das selber einen Umfang von mehr als einem einsemestrigen Kurs hat.

5.1 Affine Räume: Wir bauen uns eine geeignete Mathematische Struktur

Wir hatten zu Beginn der ELA festgestellt, dass wir mit dem \mathbb{R}^n sowohl Punkte im n -dimensionalen Raum als auch Vektoren beschreiben können. Da man mit Punkten schlecht rechnen kann, hatten wir uns zunächst mit der Rolle des \mathbb{R}^n als Vektorraum befasst. Damit kamen wir schon recht weit und konnten das Ganze auch auf andere Räume wie den \mathbb{C}^n bzw. ganz allgemeine Vektorräume übertragen. Wenn wir die Lineare Algebra für die Geometrie nutzbar machen wollen, müssen wir wieder Punkte haben, an denen wir dann die Vektoren festzurren können. Als Motivation benutzen wir ein Beispiel aus dem \mathbb{R}^n :

Beispiel 234. Wir schauen uns das erste Bild aus der ELA nochmals an, zeichnen aber zu den Punkten die Ortsvektoren ein:



5 Analytische Geometrie: Wir binden die Vektoren an Punkten fest

Das Ergebnis sehen Sie im linken Bild. Im rechten sehen Sie, was passiert, wenn man die gleichen Vektoren statt im Ursprung $(0, 0)$ im Punkt $(-4, 1)$ ansetzt. Bei den Punkten an der Spitze der Vektoren wird von der ersten Komponente Vier abgezogen, zur zweiten Eins addiert. Allgemein bekommt man als Verbindungsvektor von (a, b) und (p, q) einfach $(p - a, q - b)^T$.

Das Prinzip des *Verbindungsvektors* wollen wir nun zum Ausgangspunkt unserer Definition des *Affinen Raums* machen und die Sache mit der Differenz der Koordinaten zwischen Ausgangs- und Zielpunkt des Vektors ausnutzen, um alle Vektorräume, insbesondere \mathbb{K}^n für einen beliebigen Körper \mathbb{K} zum affinen Raum über sich selbst zu machen.

Definition 235 (Affine Räume). Gegeben seien ein Vektorraum V und eine nichtleere Menge \mathbb{A} , deren Elemente wir als *Punkte* und üblicherweise mit Großbuchstaben bezeichnen, sowie eine Abbildung

$$f : \mathbb{A} \times \mathbb{A} \rightarrow V, \\ (P, Q) \mapsto \overrightarrow{PQ}.$$

Wir nennen \mathbb{A} einen *affinen Raum (über dem Vektorraum V)*, wenn die folgenden beiden Bedingungen erfüllt sind:

1. Zu jedem Punkt $P \in \mathbb{A}$ und jedem Vektor $\mathbf{v} \in V$ gibt es genau einen Punkt $Q \in \mathbb{A}$ mit $\overrightarrow{PQ} = \mathbf{v}$.
2. Für alle $P, Q, R \in \mathbb{A}$ ist $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$.

Hat V die Dimension n , so bezeichnen wir \mathbb{A} als *n -dimensionalen Affinen Raum*. Im Fall $n = 1$ sprechen wir von einer *Geraden* bei $n = 2$ von einer *Ebene*.

Wir fordern also genau das, was uns im \mathbb{R}^n , wenn wir seine Elemente als Punkte auffassen, selbstverständlich ist. Damit haben wir auch schon die erste naheliegende Klasse von Beispielen:

Satz 236. Mit der naheliegenden Setzung $\overrightarrow{uv} = \mathbf{v} - \mathbf{u}$ wird jeder Vektorraum zum Affinen Raum über sich selbst.

Beweis. Wir müssen nur die beiden Forderungen, genauer *Axiome* auf ihre Gültigkeit überprüfen:

1. Wir suchen also zu \mathbf{v} als Punkt und \mathbf{w} als Vektor einen Punkt \mathbf{u} mit $\mathbf{w} = \mathbf{v} - \mathbf{u}$. Da V aber bezüglich der Addition eine abelsche Gruppe bildet, lässt sich diese Gleichung eindeutig auflösen, und wir bekommen $\mathbf{u} = \mathbf{v} - \mathbf{w}$.
2. Wählen wir $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, so finden wir

$$\begin{aligned} \overrightarrow{uv} + \overrightarrow{vw} &= (\mathbf{v} - \mathbf{u}) + (\mathbf{w} - \mathbf{v}) \\ &= \mathbf{w} - \mathbf{u} \\ &= \overrightarrow{uw}. \end{aligned}$$

□

Wir können aber auch erste einfache Resultate gewinnen:

Lemma 237. In einem affinen Raum gelten

- $\overrightarrow{PQ} = \mathbf{0} \Leftrightarrow P = Q.$
- $\overrightarrow{PQ} = -\overrightarrow{QP}$
- Sowohl aus $\overrightarrow{PQ} = \overrightarrow{PR}$ als auch aus $\overrightarrow{QS} = \overrightarrow{RS}$ folgt $Q = R.$

Beweis. Übung.

□

Wir können natürlich auch *affine Unterräume* definieren. Dabei möchten wir eine Teilmenge $L \subseteq \mathbb{A}$ eines affinen Raums genau dann als *affinen Unterraum* von \mathbb{A} bezeichnen, wenn sie mit der von \mathbb{A} ererbten Struktur ein affiner Raum ist. Damit lautet die Definition:

Definition 238 (Affiner Unterraum). Sei \mathbb{A} ein affiner Raum über dem Vektorraum V und $L \subseteq \mathbb{A}$.

- Ist

$$U_L := \{\overrightarrow{PQ} \mid P, Q \in L\}$$

ein Untervektorraum von V , so heißt L ein *affiner Unterraum* von \mathbb{A} und U_L der zugehörige *Richtungsraum*.

- Dessen *Dimension* ist dann die Dimension des Unterraums.
- Wir bezeichnen $\text{codim } L := \text{codim } U_L := \dim(V/U_L)$ als die *Kodimension* sowohl des affinen Unterraums L als auch seines Richtungsraums U_L
- Ist $\text{codim } L = 1$, so nennt man L eine *Hyperebene* in \mathbb{A} .

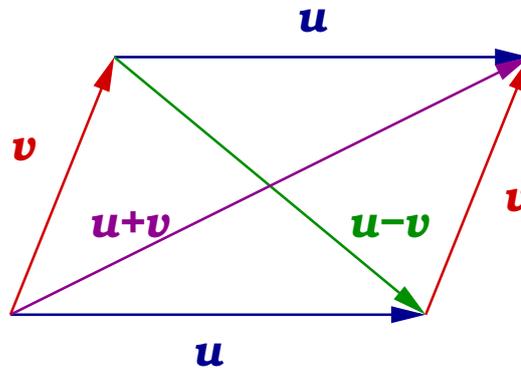
Wie sich leicht nachvollziehen lässt, ist ein solcher Unterraum tatsächlich ein affiner Raum über dem Vektorraum U_L . Als Unterräume bekommen die Begriffe *Ebene* und *Gerade* erst ihre volle Bedeutung. So können wir also Geraden in einer Ebene finden und Ebenen in einem höherdimensionalen Affinen Raum. Ein nulldimensionaler Unterraum hat dann genau einen Punkt, weshalb wir den Begriff *Punkt* auch allgemein für nulldimensionale (Unter)räume verwenden. Der Name *Hyperebene* kommt daher, dass es im Anschauungsraum tatsächlich Ebenen sind.

Beispiel 239.

- Die Lösungsmenge $L(\mathbf{A}, \mathbf{b})$ eines linearen Gleichungssystems mit Systemmatrix $\mathbb{K}^{m \times n}$ ist ein affiner Unterraum von \mathbb{K}^n .
- Die Äquivalenzklassen $[\mathbf{w}] \in V/U$, also die Elemente des Faktorraums, sind affine Unterräume von V (siehe Übung).

5 Analytische Geometrie: Wir binden die Vektoren an Punkten fest

- Der Anschauungsraum lässt sich als affiner Raum auffassen:
 - Die Relation zwischen gerichteten Strecken „lässt sich durch Translation überführen in“ ist eine Äquivalenzrelation. Die Äquivalenzklassen sind nun unsere *Vektoren im Raum*.
 - Die Addition ist so definiert, wie wir es schon in vielen Abbildungen verwendet haben, etwa der Parallelogrammgleichung



- Zur Multiplikation werden die Vektoren einfach um den entsprechenden Faktor gestreckt (bzw. gestaucht). Bei negativen Faktoren dreht sich natürlich die Richtung um.
- Damit haben wir einen Vektorraum
- Ordnen wir nun jedem Punktepaar (P, Q) im Anschauungsraum die Äquivalenzklasse der Verbindungsstrecke von P nach Q als Verbindungsvektor zu, so haben wir den fertigen affinen Raum.

In diesem sind dann die ein- und zweidimensionalen affinen Unterräume genau das, was man sich landläufig unter Geraden und Ebenen vorstellt. Auch der Begriff des Richtungsraums ist hier selbsterklärend.

- Im Raum \mathbb{F}_2^n sind die affinen Unterräume sehr klein. Der Raum selber hat ja nur 2^n Elemente. Jede Gerade hat genau zwei Punkte, jede Ebene genau drei. Treibt man das Spiel mit dem \mathbb{F}_2^n weiter, so kommt man zu Objekten wie etwa Fano-Ebenen, die in der diskreten Mathematik und der Informatik bei der Betrachtung von Nachbarschaftsstrukturen eine Rolle spielen.

Wir können nun auch ein Äquivalent zum Begriff der linearen Hülle finden. Dabei machen wir uns das folgende Ergebnis zunutze:

Satz 240. *Ist \mathbb{A} ein affiner Raum über dem Vektorraum V und $\mathcal{M} \subseteq \mathcal{P}(\mathbb{A})$ eine Menge affiner Unterräume von \mathbb{A} . Dann ist der Schnitt*

$$M := \bigcap_{L \in \mathcal{M}} L$$

5.1 Affine Räume: Wir bauen uns eine geeignete Mathematische Struktur

entweder leer oder selber ein affiner Unterraum von \mathbb{A} , dessen Richtungsraum dann durch

$$U_M := \bigcap_{L \in \mathcal{M}} U_L$$

gegeben ist.

Beweis. Wir betrachten den Fall, dass M nicht leer ist. Dann können wir dort ein $P \in M$ wählen und die folgende Menge bilden:

$$\{\overrightarrow{PX} \mid X \in M\} = \bigcap_{L \in \mathcal{M}} \{\overrightarrow{PX} \mid X \in L\} = \bigcap_{L \in \mathcal{M}} U_L = U_M.$$

Da wir aber wissen, dass der Schnitt von Untervektorräumen eines Vektorraums V wieder ein solcher ist, sind wir fertig. \square

Damit ist das folgende wohldefiniert:

Definition 241 (Affine Hülle). Ist $C \subseteq \mathbb{A} \neq \emptyset$, so heißt der Durchschnitt aller affinen Unterräume von \mathbb{A} , die C enthalten die *affine Hülle* von C .

Enthält C nur endlich viele Punkte, also $C = \{P_1, \dots, P_k\}$, so nennen wir die affine Hülle von C den *Verbindungsraum* $P_1 \vee \dots \vee P_k$ der Punkte P_1, \dots, P_k .

Bemerkung 242. Der Richtungsraum des Verbindungsraums der Punkten P_0, \dots, P_k (man beachte, dass wir bei Null zu zählen beginnen) lässt sich darstellen durch die lineare Hülle $\langle \overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_k} \rangle$, was die Begriffe der affinen und der linearen Hülle miteinander verbindet.

Das motiviert auch den Begriff der *affinen Abhängigkeit* bzw. *Unabhängigkeit*:

Definition 243. In einem affinen Raum \mathbb{A} heißen die Punkte P_0, \dots, P_k *affin unabhängig* oder *in allgemeiner Lage*, wenn $\dim(P_0 \vee \dots \vee P_k) = k$. Ansonsten heißen sie *affin abhängig*.

Die Verbindung zur linearen (Un)abhängigkeit lässt sich leicht zeigen:

Satz 244. In einem Affinen Raum \mathbb{A} über dem Vektorraum V sind äquivalent:

1. Die Punkte $P_1, \dots, P_k \in \mathbb{A}$ sind affin unabhängig.
2. Die Vektoren $\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_k} \in V$ sind linear unabhängig.

Beweis. Sind die Punkte $P_1, \dots, P_k \in \mathbb{A}$ affin unabhängig, so gilt für den Verbindungsraum $\dim(P_1 \vee \dots \vee P_k) = \dim\langle \overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_k} \rangle = k$. Das geht aber nur, wenn die Vektoren $\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_k} \in V$ linear unabhängig sind.

Für die umgekehrte Implikation müssen wir nur das Ganze rückwärts spielen. \square

Nun hätten wir gerne auch für unsere - diesmal eher geometrische als algebraische - Struktur eine Charakterisierung der strukturerhaltenden Abbildungen. Wir versuchen es folgendermaßen:

Definition 245 (Affine Abbildungen). Seien \mathbb{A} und \mathbb{B} affine Räume über den Vektorräumen $V_{\mathbb{A}}$ bzw. $V_{\mathbb{B}}$. Eine Abbildung $f : \mathbb{A} \rightarrow \mathbb{B}$ heißt *affin*, wenn es eine zugehörige lineare Abbildung $\varphi_f : V_{\mathbb{A}} \rightarrow V_{\mathbb{B}}$ gibt mit

$$\overrightarrow{f(P)f(Q)} = \varphi_f(\overrightarrow{PQ}). \quad (5.1)$$

Ist $\mathbb{B} = \mathbb{A}$, also f eine *affine Selbstabbildung*, und f bijektiv, so spricht man von einer *Affinität*.

Die lineare Abbildung φ_f ist tatsächlich eindeutig. Es kann nämlich nach Satz 66 der ELA höchstens eine lineare Abbildung geben, die Gleichung (5.1) erfüllt.

Hier noch ein paar allgemeine Resultate:

Satz 246.

1. Eine Abbildung $f : \mathbb{A} \rightarrow \mathbb{B}$ ist genau dann *affin*, wenn es ein $P_0 \in \mathbb{A}$ gibt, sodass die Abbildung

$$\begin{aligned} \varphi_f : V_{\mathbb{A}} &\rightarrow V_{\mathbb{B}}, \\ \overrightarrow{P_0Q} &\mapsto \overrightarrow{f(P_0)f(Q)} \end{aligned} \quad (5.2)$$

linear ist.

2. Seien mit den Bezeichnungen aus der Definition $P_0 \in \mathbb{A}, Q_0 \in \mathbb{B}$ und $\psi : V_{\mathbb{A}} \rightarrow V_{\mathbb{B}}$ linear. Dann gibt es genau eine affine Abbildung $f : \mathbb{A} \rightarrow \mathbb{B}$ mit $f(P_0) = Q_0$ und $\varphi_f = \psi$.
3. Eine affine Abbildung f ist genau dann bijektiv, wenn φ_f bijektiv ist. Dann ist auch die Umkehrabbildung f^{-1} affin, und es gilt

$$\varphi_{f^{-1}} = (\varphi_f)^{-1}.$$

4. Ist $g : \mathbb{B} \rightarrow \mathbb{D}$ eine weitere affine Abbildung, dann ist auch $g \circ f$ affin.

Beweis.

1. Dass für eine affine Abbildung f die Abbildung φ_f aus Gleichung (5.2) linear sein muss, folgt direkt aus der Definition.

Sei diese nun umgekehrt als linear vorausgesetzt, und wir versuchen zu zeigen, dass f affin ist.

Offenbar gilt für jedes $P \in \mathbb{A}$

$$\overrightarrow{PQ} = \overrightarrow{PP_0} + \overrightarrow{P_0Q} \quad \text{und} \quad \overrightarrow{f(P)f(Q)} = \overrightarrow{f(P)f(P_0)} + \overrightarrow{f(P_0)f(Q)}.$$

Wenn man beachtet, dass sich beim Tausch von Anfangs- und Endpunkt das Vorzeichen des Vektors dreht, reicht das, um sofort nachzurechnen, dass

$$\overrightarrow{PQ} \mapsto \overrightarrow{f(P)f(Q)}.$$

2. Funktioniert analog.
3. Übung.
4. Folgt direkt aus der Definition durch Einsetzen.

□

Die beiden ersten Punkte motivieren zusammen mit unserem Vorgehen bei der affinen Unabhängigkeit den Begriff des *Ursprungs* in einem affinen Raum. Wir bezeichnen dafür einen Punkt mit O als Symbol für den Ursprung. Dann ist jeder Punkt $P \in \mathbb{A}$ durch seinen Ortsvektor \overrightarrow{OP} festgelegt. Wählt man etwa in \mathbb{K}^n den Punkt $(0, \dots, 0)$ als Ursprung, so hat jeder Punkt die selben Komponenten wie sein Ortsvektor. Man kann sie also in gewisser Weise gleichsetzen.

Bemerkung 247. Bezeichnen wir den Ursprung in \mathbb{B} mit O' , so bekommen wir mit obigem Satz

$$\overrightarrow{O'f(P)} = \overrightarrow{O'f(O)} + \overrightarrow{f(O)f(P)} =: \overrightarrow{f(O)f(P)} + \mathbf{a} = \varphi_f(\overrightarrow{OP}) + \mathbf{a}$$

mit dem festen Vektor $\mathbf{a} \in V_{\mathbb{B}}$, der nur von der Abbildung f abhängt, nicht aber vom abgebildeten Punkt P . Daher kann man sich für die meisten Betrachtungen auf den Fall $\mathbb{A} = V_{\mathbb{A}} =: V, \mathbb{B} = V_{\mathbb{B}} =: W$ zurückziehen und Abbildung schreiben als

$$\begin{aligned} f : V &\rightarrow W, \\ \mathbf{v} &\mapsto \varphi(\mathbf{v}) + \mathbf{a} \end{aligned}$$

mit $\varphi \in \text{hom}(V, W)$, $\mathbf{a} \in W$. Der Vektor \mathbf{a} heißt dann der *Translationsvektor* zur affinen Abbildung f .

Mit dieser Bemerkung im Rücken lassen sich dann viele Dinge, die man gerne zeigen würde, wie etwa, dass affine Abbildungen affine Unterräume auf affine Unterräume abbilden und affine Hüllen auf affine Hüllen direkt auf unsere Resultate für lineare Abbildungen zurückführen. Da bei linearen Abbildungen die Urbilder linear unabhängiger Vektoren wieder linear unabhängig sind, sind dann mit Satz 244 die Urbilder affin unabhängiger Punkte affin unabhängig. Und so weiter.

5.2 Geometrie im \mathbb{R}^n : Koordinatenbehaftete Geometrie

Im Anschauungsraum ist es üblich, Koordinaten bezüglich eines kartesischen Koordinatensystems zu verwenden. Diese liegen dann für die Ebene im \mathbb{R}^2 , für den Anschauungsraum im \mathbb{R}^3 . Um dieses zu tun, müssen wir uns erst mal überlegen, was das denn genau ist.

5 Analytische Geometrie: Wir binden die Vektoren an Punkten fest

- Sowohl im Anschauungsraum bzw. der euklidischen Ebene als auch im \mathbb{R}^n (versehen mit dem Standardskalarprodukt) haben wir einen Begriff von Orthogonalität, von Winkeln und von Längen.
- Damit können wir uns im Vektorraum der Vektoren im Raum bzw. der Ebene jeweils eine ONB wählen, bezüglich der die Vektoren darstellen.
- Haben wir einen Ursprung gewählt, so bilden die Vektoren der ONB verstanden als Ortsvektoren von Punkten ein kartesisches Koordinatensystem. Wir müssen lediglich auf die Sortierung der Vektoren des Koordinatensystems achten, schließlich sollten diese aus geometrischen Gründen ein *rechtshändiges* System bilden.
- Damit haben wir die Rechtfertigung, die Geometrie in der euklidischen Ebene im \mathbb{R}^2 und im Anschauungsraum im \mathbb{R}^3 abzuhandeln. Schließlich gilt das, wie wir in Abschnitt 2.7 gesehen haben für allgemeine Innenprodukträume auch.
- Der Vollständigkeit halber sei erwähnt, dass das Innenprodukt im Anschauungsraum bzw. der Ebene definiert ist durch

$$\mathbf{x} \cdot \mathbf{y} := \begin{cases} \|\mathbf{x}\| \|\mathbf{y}\| \cos(\sphericalangle(\mathbf{x}, \mathbf{y})) & \text{falls } \mathbf{x}, \mathbf{y} \neq \mathbf{0}, \\ 0 & \text{sonst,} \end{cases}$$

wobei man, wenn keine Verwechslungsgefahr besteht, statt der Norm auch Betragssstriche schreiben kann. Diese Formel hatten wir im Grunde genommen schon umgekehrt als Definition des Cosinus in allgemeinen Innenprodukträumen verwendet.

5.2.1 Die Ebene: \mathbb{R}^2

Wir betrachten zunächst die affinen Unterräume, in diesem Fall nur die Geraden, dann einfache geometrische Figuren, nämlich Parallelogramme und Dreiecke und zuletzt sogenannte Quadriken, wo die Sache anfängt, rund zu werden.

5.2.1.1 Geraden in der Ebene

Die einzigen spannenden affinen Unterräume im \mathbb{R}^2 sind die Geraden, die ja bereits die Hyperebenen sind. Mit Dimension zwei hätten wir ja schon wieder den gesamten Raum, in diesem Fall also die *Koordinatenebene*.

- Eine Gerade durch den Punkt P (mit Ortsvektor \mathbf{p}) mit Richtungsraum $\langle \mathbf{v} \rangle$ lässt sich dadurch charakterisieren, dass man für Ihre Punkte X den allgemeinen Ortsvektor hinschreibt:

$$\mathbf{x} := \overrightarrow{OX} = \overrightarrow{OP} + \lambda \mathbf{v} = \mathbf{p} + \lambda \mathbf{v},$$

wobei λ ganz \mathbb{R} durchläuft.

- Das ist aber genau die Gestalt, die wir von der Lösungsmenge eines inhomogenen LGS $(A|\mathbf{b})$ mit $A \in \mathbb{R}^{(n-1) \times n}$, $\mathbf{b} \in \mathbb{R}^{n-1}$ und $\text{rk}A = n - 1$ kennen. In unserem Fall ist dann A eine einzelne Zeile der Länge Zwei und $\mathbf{b} = b$ ein Skalar.
- Unser Richtungsvektor \mathbf{v} ist nun gerade eine Basis des Lösungsraums des homogenen Systems.
- Schreiben wir also $A = \mathbf{a}^T$, so gilt gerade $\mathbf{a} \cdot \mathbf{v} = 0$. Der Vektor \mathbf{a} steht also genau senkrecht auf dem Lösungsraum/Richtungsvektor.
- Normieren wir diesen auf Länge Eins zu \mathbf{n} und sorgen dafür, dass er zum Ortsvektor \mathbf{p} unseres *Aufpunkts* P einen Winkel hat, der kleiner als ein rechter ist, d. h., dass gilt

$$\cos(\sphericalangle(\mathbf{n}, \mathbf{p})) > 0$$

- notfalls müssen wir einfach das Vorzeichen drehen -, so ist die Gerade einfach die Lösungsmenge von

$$\mathbf{n} \cdot \mathbf{x} = d \quad (\text{Hessesche Normalform})$$

mit $0 < d = \mathbf{n} \cdot \mathbf{p}$.

Die *Hessesche Normalform (HNF)*¹ ist so allgemein, dass wir auch in höheren Dimensionen die Hyperebenen durch eine solche beschreiben können. Die Frage ist jeweils, wie man sich einen Vektor besorgt, der senkrecht auf dem Lösungsraum steht. Im \mathbb{R}^2 ist es einfach:

1. Starte mit dem Richtungsvektor \mathbf{v} und bilde den zu diesem orthogonalen Vektor $\mathbf{v}^\perp = (v_2, -v_1)^T$.

2. Normiere \mathbf{v}^\perp zu

$$\tilde{\mathbf{n}} = \frac{1}{\|\mathbf{v}^\perp\|} \mathbf{v}^\perp.$$

3. Berechne das Skalarprodukt mit dem Ortsvektor \mathbf{p} des Aufpunkts.

- Ist dieses Null, geht die Gerade durch den Ursprung, und wir haben die rechte Seite $d = 0$.
- Ist dieses positiv, so setzen wir $\mathbf{n} = \tilde{\mathbf{n}}$ und $d = \mathbf{n} \cdot \mathbf{p}$.
- Ansonsten setzen wir $\mathbf{n} = -\tilde{\mathbf{n}}$ und $d = \mathbf{n} \cdot \mathbf{p} = -\tilde{\mathbf{n}} \cdot \mathbf{p}$.

Damit haben wir dann die Hessesche Normalform beisammen. Der umgekehrte Weg geht natürlich durch Lösen des LGS.

Hat man bereits die Form

$$\mathbf{a} \cdot \mathbf{x} = b,$$

also ein LGS mit einer Gleichung und zwei Unbekannten, so muss man lediglich durch $\|\mathbf{a}\|$ teilen und gegebenenfalls die Vorzeichen drehen, um die Hessesche Normalform zu bekommen.

¹Otto Hesse, 1811-1874

Definition 248. Der Vektor \mathbf{n} in der Hesseschen Normalform heißt der *Normalenvektor* der Gerade.

Beispiel 249. Wir betrachten die Gerade mit

$$\mathbf{x}(\lambda) = \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} -3 \\ 4 \end{pmatrix}.$$

Damit ist unser auf dem Richtungsvektor senkrecht stehender Vektor $\mathbf{v}^\perp = (4 \ 3)^T$. Wir berechnen

$$\|\mathbf{v}^\perp\|^2 = 3^2 + 4^2 = 25 \quad \Rightarrow \|\mathbf{v}^\perp\| = 5.$$

Damit haben wir $\tilde{\mathbf{n}} = (4/5 \ 3/5)^T$. Als nächstes müssen wir das Skalarprodukt mit dem Ortsvektor des Aufpunkts bilden, also

$$\tilde{\mathbf{n}} \cdot \mathbf{x}(0) = \begin{pmatrix} 4/5 \\ 3/5 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 4/5 + 6/5 = 2 > 0.$$

Damit ist $\mathbf{n} = \tilde{\mathbf{n}}$, und die Hessesche Normalform der Geraden lautet

$$\frac{4}{5}x_1 + \frac{3}{5}x_2 = 2.$$

Beispiel 250. Haben wir die Gerade bereits gegeben durch die Gleichung

$$5x_1 + 7x_2 = -4,$$

so müssen wir sie mit -1 durchmultiplizieren und durch

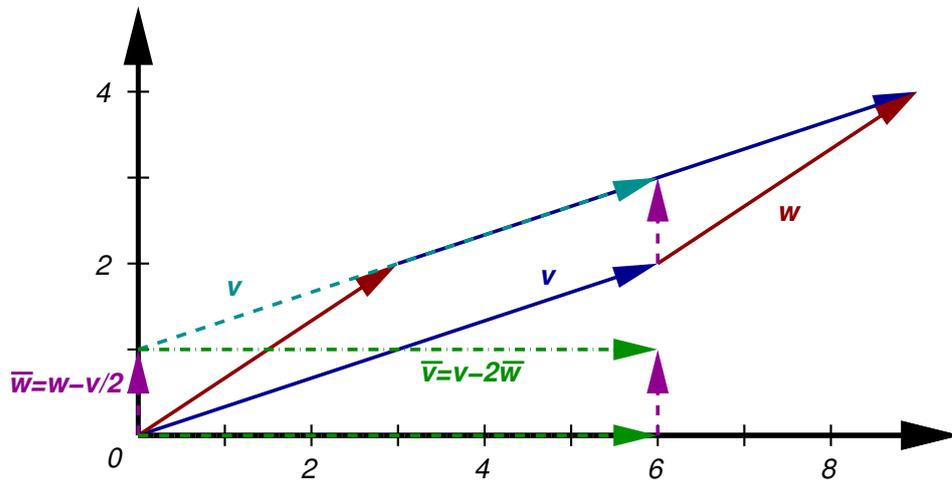
$$\left\| \begin{pmatrix} 5 \\ 7 \end{pmatrix} \right\| = \sqrt{74}$$

dividieren. Damit lautet die HNF

$$\frac{-5}{\sqrt{74}}x_1 + \frac{-7}{\sqrt{74}}x_2 = \frac{4}{\sqrt{74}}.$$

5.2.1.2 Flächen von Parallelogrammen und Dreiecken

Wir können in kartesischen Koordinaten auch bequem Flächen von Figuren ausrechnen, z. B. von Parallelogrammen. Dabei erinnern wir uns, wie wir in der Elementaren Linearen Algebra die Determinante motiviert hatten:



Dabei hatten wir das von \mathbf{v} und \mathbf{w} aufgespannte Parallelogramm durch Scherungen flächenerhaltend in ein Rechteck überführt, dessen Inhalt wir leicht berechnen konnten. Schreiben wir also die aufspannenden Vektoren in eine Matrix $(\mathbf{v}|\mathbf{w})$, so erhalten wir den Flächeninhalt durch $\det(\mathbf{v}|\mathbf{w})$, ohne explizit die Höhe des Parallelogramms konstruieren zu müssen. Im abgebildeten Beispiel also

$$\det(\mathbf{v}|\mathbf{w}) = \begin{vmatrix} 6 & 3 \\ 2 & 2 \end{vmatrix} = 12 - 6 = 6.$$

Wollen wir die Orientierung außer acht lassen, so nehmen wir den Betrag und finden allgemein für das von \mathbf{v} und \mathbf{w} aufgespannte Parallelogramm die Fläche

$$A_{\mathbf{v},\mathbf{w}} = |v_1 w_2 - v_2 w_1|. \quad (5.3)$$

Die Fläche des von \mathbf{v} und \mathbf{w} aufgespannten Dreiecks ist dann genau die Hälfte, was wir plötzlich ohne umständliche Konstruktion der Höhe berechnen können.

5.2.1.3 Wir werden vorsichtig nichtlinear: Quadriken in der Ebene

Nun wollen wir schauen, was wir aus der linearen Algebra für nichtlineare Probleme, im einfachsten Fall quadratische, gewinnen können. Wir starten mit der Kreisgleichung, wobei wir unseren Ortsvektoren jetzt die Koordinatendarstellung $\mathbf{x} = (x, y)^T$ geben, wie sie in der Physik oder den Ingenieurwissenschaften üblich ist:

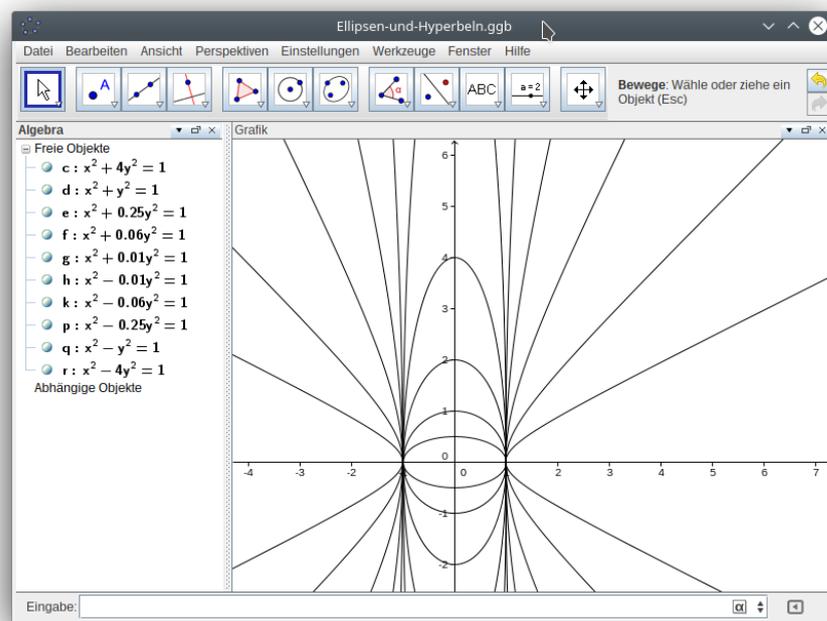
$$x^2 + y^2 = r^2 \quad \Leftrightarrow \quad \frac{1}{r^2}x^2 + \frac{1}{r^2}y^2 = 1$$

mit einem Radius $r > 0$. Dabei eignet sich die zweite Formulierung eher für Verallgemeinerungen wie etwa das Ändern des Radius nur in einer Koordinatenrichtung bzw. in beide Koordinatenrichtungen getrennt:

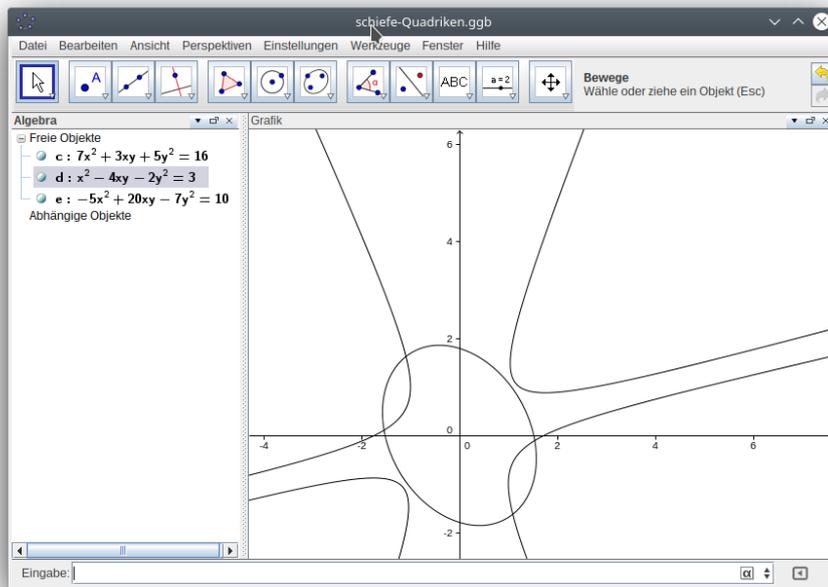
$$\frac{1}{a^2}x^2 + \frac{1}{b^2}y^2 = 1. \quad (5.4)$$

5 Analytische Geometrie: Wir binden die Vektoren an Punkten fest

Im folgenden Bild sind einige Fälle dargestellt, inklusive solcher, bei denen wir mit $1/b^2$ nicht nur nahe an Null gehen, sondern mit negativem Vorzeichen, quasi einem imaginären Radius, weitermachen:



Man sieht leicht, dass man dabei von Ellipsen zu Hyperbeln übergeht. Was allen diesen Varianten gemein ist, sind die Hauptachsen, nämlich die Koordinatenachsen. Kommen gemischte Terme dazu, also Vielfache von xy , so ist nicht mehr genau zu sehen, welches die Hauptachsen und dementsprechend die Achsenabschnitte etc. sind:



Im Bild kann man nicht einmal auf den ersten Blick erkennen, welche Kurven zu welcher Eingabe gehören.

Wir beginnen mit der allgemeinen Form

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + b_1x + b_2y + c = 0. \quad (5.5)$$

Diese können wir mit einer symmetrischen Matrix und einem Vektor

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \quad \text{und} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

darstellen als

$$\mathbf{x}^T A \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = 0. \quad (5.6)$$

Nun wissen wir aus dem reellen Spektralsatz 92, dass A orthogonal diagonalisierbar ist. Also besitzt die Quadrik bezüglich einer anderen ONB als der, bezüglich der wir die Koordinaten genommen hatten, eine Darstellung ohne die gemischten Terme. Die neue ONB beschreibt damit die Hauptachsen, und wir sprechen bei der zugehörigen Transformation auf die neuen Koordinaten von einer *Hauptachsentransformation*. Die linearen Terme müssen dabei nicht automatisch verschwinden. So beschreibt etwa

$$\frac{x^2}{\alpha^2} - 2y = 0$$

eine Parabel. Aber auch eine Verschiebung des Mittelpunkts der Quadrik nach außerhalb des Ursprungs wird darin sichtbar.

Wir gehen hier nicht in alle Details, sondern betrachten ein Beispiel:

Beispiel 251. Wir haben eine Quadrik gegeben durch

$$2x^2 - y^2 + 4xy - 2x + y - 6 = 0$$

und schreiben diese entsprechend Gleichung (5.6) um. Dabei erhalten wir

$$A = \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix} \quad \text{und} \quad \mathbf{b} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Das charakteristische Polynom der Matrix lautet

$$\chi_A(\lambda) = (2 - \lambda)(-1 - \lambda) - 4 = \lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2).$$

Damit haben wir die Eigenwerte $\lambda_1 = 3$, $\lambda_2 = -2$. Praktischerweise müssen damit die Eigenvektoren zu den beiden Eigenwerten bereits orthogonal sein. Wir müssen sie nur noch normieren und auf die richtige Orientierung achten:

$$\mathbf{q}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{q}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

Damit bilden wir die orthogonale Matrix $S = (\mathbf{q}_1 | \mathbf{q}_2)$ und erhalten

$$S^T A S =: B = \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$$

Nun transformieren wir den Koordinatenvektor \mathbf{x} zu $\mathbf{y} = (\xi, \eta) = S^T \mathbf{x}$ und bilden $\mathbf{d} := S^T \mathbf{b}$. Dann liest sich unsere Quadrik als

$$\mathbf{y}^T B \mathbf{y} + \mathbf{d}^T \mathbf{y} - 6 = 0$$

bzw.

$$3\xi^2 - 2\eta^2 - \frac{3}{\sqrt{5}}\xi + \frac{4}{\sqrt{5}}\eta - 6 = 0.$$

Die verbliebenen linearen Terme weisen darauf hin, dass der Mittelpunkt der Quadrik noch außerhalb des Ursprungs liegt. Durch quadratische Ergänzung kommen wir auf die noch durchzuführende Translation

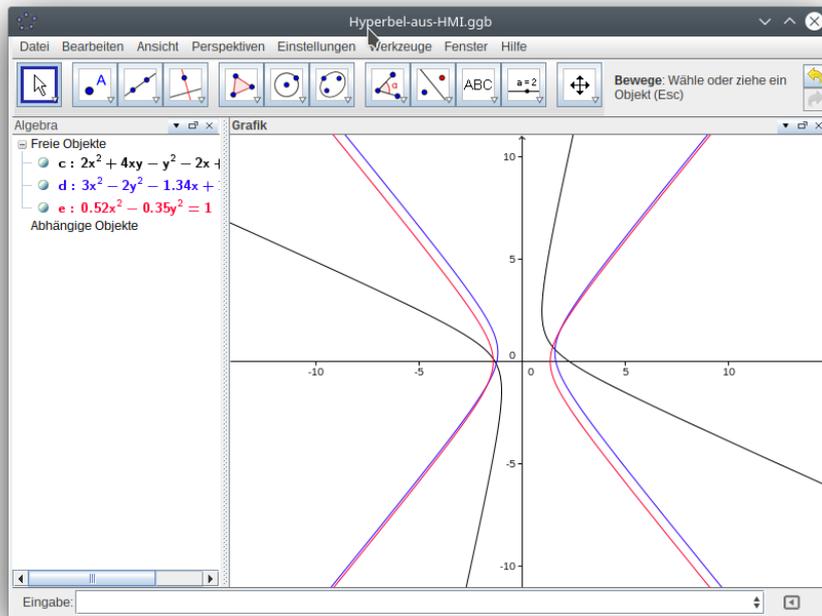
$$\mathbf{z} := \begin{pmatrix} u \\ v \end{pmatrix} := \begin{pmatrix} \xi - \frac{1}{2\sqrt{5}} \\ \eta - \frac{1}{\sqrt{5}} \end{pmatrix} = \mathbf{y} - \begin{pmatrix} \frac{1}{2\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix} =: \mathbf{y} - \mathbf{t}$$

Bezüglich dieser Koordinaten bekommen wir tatsächlich die Normalform entsprechend dem Anfang dieses Abschnitts:

$$\frac{u^2}{23/12} - \frac{v^2}{23/8} = 1.$$

Dies ist demnach eine Hyperbel mit Mittelpunkt $\left(\frac{1}{2\sqrt{5}}, \frac{1}{\sqrt{5}}\right)$ und den Hauptachsenrichtungen \mathbf{q}_1 und \mathbf{q}_2 . Auch die Exzentrizität und die Koordinaten der Scheitel lassen sich mit den berechneten Daten bestimmen.

Im Bild sehen Sie die Ursprungsvariante schwarz, die Zwischenform mit den Achsenparallelen Hauptachsen blau und die Normalform mit dem Mittelpunkt im Ursprung rot: Tippfehler)



5.2.2 Der Raum: \mathbb{R}^3

Hier drehen wir die Reihenfolge etwas und befassen uns zuerst mit Raum- und Flächeninhalten, ehe wir uns mit den Geraden und Ebenen als den echten nichttrivialen affinen Unterräumen befassen.

5.2.2.1 Raum- und Flächeninhalte

Das dreidimensionale Analogon zum Parallelogramm ist ein *Parallelepiped* bzw. *Spat*, wobei sich der Begriff auch für höhere Dimensionen erweitern lässt. Ein Spat wird von drei Vektoren aufgespannt, etwa \mathbf{a} , \mathbf{b} , \mathbf{c} . Der Rauminhalt, das sogenannte *Spatprodukt* lässt sich in kartesischen Koordinaten wieder über die Determinante berechnen, also

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \det(\mathbf{a}|\mathbf{b}|\mathbf{c}).$$

5 Analytische Geometrie: Wir binden die Vektoren an Punkten fest

Wie man das ohne Koordinaten getan hätte, sieht man hier noch nicht. Nun bringen wir eine kleine Anwendung des Laplaceschen Entwicklungssatzes ins Spiel:

Sei $M = (\mathbf{a} \mid \mathbf{b} \mid \mathbf{c}) \in \mathbb{R}^{3 \times 3}$. Dann ist nach dem Laplaceschen Entwicklungssatz

$$\det(M) = \mathbf{a} \cdot \begin{pmatrix} \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} \\ - \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} \\ \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \end{pmatrix} = \mathbf{a} \cdot \begin{pmatrix} b_2 c_3 - b_3 c_2 \\ b_3 c_1 - b_1 c_3 \\ b_1 c_2 - b_2 c_1 \end{pmatrix} =: \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}). \quad (5.7)$$

Den Vektor $\mathbf{b} \times \mathbf{c}$ nennt man das *Vektorprodukt* bzw. *Kreuzprodukt* von \mathbf{b} und \mathbf{c} . Offenbar steht dieses senkrecht auf \mathbf{b} und \mathbf{c} , und es gilt

$$\det(\mathbf{b}, \mathbf{c}, \mathbf{b} \times \mathbf{c}) \geq 0,$$

für linear unabhängige Vektoren sogar „ > 0 “, sodass dann $\mathbf{b}, \mathbf{c}, \mathbf{b} \times \mathbf{c}$ ein rechtshändiges System bilden.

Zu klären bleibt nur noch die Frage nach der Norm dieses Vektors, also seiner Länge. Projizieren wir \mathbf{a} auf die Richtung von $\mathbf{b} \times \mathbf{c}$, so finden wir

$$\mathbf{p} := \frac{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}{\|\mathbf{b} \times \mathbf{c}\|^2} (\mathbf{b} \times \mathbf{c}).$$

Dies Projektion steht also ebenfalls senkrecht auf der durch \mathbf{b}, \mathbf{c} definierten Ebene. Für seine Länge ergibt sich entsprechend der Konstruktion unseres Vektorprodukts

$$\begin{aligned} \|\mathbf{p}\|^2 &= \left\| \frac{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}{\|\mathbf{b} \times \mathbf{c}\|^2} (\mathbf{b} \times \mathbf{c}) \right\|^2 \\ &= \frac{\det(M)^2}{\|\mathbf{b} \times \mathbf{c}\|^4} \|\mathbf{b} \times \mathbf{c}\|^2 \\ &= \left(\frac{\det(M)}{\|\mathbf{b} \times \mathbf{c}\|} \right)^2 \end{aligned}$$

Damit ist

$$\|\mathbf{b} \times \mathbf{c}\| = \frac{[\mathbf{a}, \mathbf{b}, \mathbf{c}]}{\|\mathbf{p}\|} = \frac{\text{Volumen des Spats}}{\text{Höhe des Spats über der Seite } \mathbf{b}, \mathbf{c}}.$$

Damit ist $\|\mathbf{b} \times \mathbf{c}\|$ gerade die Fläche des von \mathbf{b} und \mathbf{c} aufgespannten Parallelogramms. Das ist auch der Schlüssel dazu, wie man das Spatprodukt koordinatenfrei bestimmen kann:

- Die Fläche des von \mathbf{b} und \mathbf{c} aufgespannten Parallelogramms lässt sich über bekannte Formeln bestimmen.
- Dann braucht man nur noch von einem Punkt außerhalb der von \mathbf{b} und \mathbf{c} aufgespannten Ebene das Lot auf diese fällen und dem so entstandenen Vektor die richtige Länge verpassen. Das Resultat ist dann $\mathbf{b} \times \mathbf{c}$.

- Danach kommt noch ein Skalarprodukt mit \mathbf{a} , und das Spatprodukt ist berechnet.

Wir haben nun insgesamt drei verschiedene Produkte von Vektoren im \mathbb{R}^3 : Das (Standard)skalarprodukt, das Vektorprodukt und das Spatprodukt.

Für das Vektorprodukt gelten folgende Formeln:

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= -\mathbf{b} \times \mathbf{a}, && \text{(Antikommutativität)} \\ \lambda(\mathbf{a} \times \mathbf{b}) &= (\lambda\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\lambda\mathbf{b}), \\ \mathbf{a} \times (\mathbf{b} + \mathbf{c}) &= \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}, \\ (\mathbf{a} + \mathbf{b}) \times \mathbf{c} &= \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}, \\ \|\mathbf{a} \times \mathbf{b}\|^2 &= \|\mathbf{a}\|^2\|\mathbf{b}\|^2 - (\mathbf{a} \cdot \mathbf{b})^2, \\ \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) &= (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}, && \text{(Grassmannscher Entwicklungssatz)} \\ (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) &= (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{b} \cdot \mathbf{c})(\mathbf{a} \cdot \mathbf{d}) && \text{(Lagrange-Identität)} \end{aligned}$$

mit $\lambda \in \mathbb{R}$, $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \mathbb{R}^3$. Das Schöne ist, dass man alle diese durch schlichtes Einsetzen mit Koordinaten nachweisen kann. Hätte man das Ganze ohne kartesische Koordinaten probiert, wäre es sehr kompliziert geworden.

5.2.2.2 Geraden und Ebenen im Raum

Während im \mathbb{R}^2 die Geraden bereits die Hyperebenen sind, fällt diese Rolle im \mathbb{R}^3 den Ebenen zu. Die Hessesche Normalform, die wir für die Geraden in der Ebene konstruiert hatten, benutzt mit dem Normalenvektor die Tatsache, dass der Orthogonalraum zum Richtungsraum eindimensional ist. Damit ist es genau genommen ein Fall für Hyperebenen, im \mathbb{R}^3 also für Ebenen. Wir starten also mit der Geraden

$$\mathbf{x} := \overrightarrow{OX} = \overrightarrow{OP} + \lambda\mathbf{v} = \mathbf{p} + \lambda\mathbf{v} + \mu\mathbf{w},$$

wobei diesmal zwei Parameter, nämlich λ und μ ganz \mathbb{R} durchlaufen. Der Richtungsraum ist also $\langle \mathbf{v}, \mathbf{w} \rangle$, wobei wir \mathbf{v}, \mathbf{w} als linear unabhängig voraussetzen müssen, damit dieser auch wirklich zweidimensional ist. Da wir zwischenzeitlich das Vektorprodukt kennen, können wir den Ablauf zur Konstruktion der Hesseschen Normalform leicht aufstellen:

- Berechne $\mathbf{v} \times \mathbf{w}$. Dieser Vektor steht automatisch senkrecht auf dem Richtungsraum.
- Normiere diesen zu

$$\tilde{\mathbf{n}} = \frac{1}{\|\mathbf{v} \times \mathbf{w}\|} \mathbf{v} \times \mathbf{w}.$$

Ab hier geht es weiter wie im \mathbb{R}^2 :

- Berechne das Skalarprodukt mit dem Ortsvektor \mathbf{p} des Aufpunkts.
 - Ist dieses Null, geht die Ebene durch den Ursprung, und wir haben die rechte Seite $d = 0$.

5 Analytische Geometrie: Wir binden die Vektoren an Punkten fest

- Ist dieses positiv, so setzen wir $\mathbf{n} = \tilde{\mathbf{n}}$ und $d = \mathbf{n} \cdot \mathbf{p}$.
- Ansonsten setzen wir $\mathbf{n} = -\tilde{\mathbf{n}}$ und $d = \mathbf{n} \cdot \mathbf{p} = -\tilde{\mathbf{n}} \cdot \mathbf{p}$.

Möchten wir eine Gerade

$$\mathbf{x} := \overrightarrow{OX} = \overrightarrow{OP} + \lambda \mathbf{v} = \mathbf{p} + \lambda \mathbf{v} ,$$

durch eine analoge Form darstellen, so benötigen wir eine Basis, vorzugsweise eine Orthonormalbasis $\mathbf{n}_1, \mathbf{n}_2$ des Orthogonalraums zum Richtungsraum $\langle \mathbf{v} \rangle$. Damit bekommen wir dann zwei Gleichungen

$$\begin{aligned} \mathbf{n}_1 \cdot \mathbf{x} &= d_1 , \\ \mathbf{n}_2 \cdot \mathbf{x} &= d_2 , \end{aligned}$$

die wieder ein unterbestimmtes LGS mit unserer Geraden als Lösungsraum bilden.

Man kann im \mathbb{R}^3 aber auch lustige Dinge tun, wie etwa den Abstand zweier *windschiefer* Geraden berechnen, also von Geraden mit verschiedenem Richtungsraum, die sich auch nicht schneiden:

- Seien zwei Geraden

$$\mathbf{x} = \mathbf{p} + \lambda \mathbf{v}$$

und

$$\mathbf{y} = \mathbf{q} + \mu \mathbf{w}$$

gegeben.

- Bezeichne den Verbindungsvektor $\mathbf{q} - \mathbf{p}$ der Aufpunkte der beiden Geraden mit \mathbf{u} .
- Nun spannen die Vektoren $\mathbf{u}, \mathbf{v}, \mathbf{w}$ ein Parallelepipid mit dem Volumen $[\mathbf{u}, \mathbf{v}, \mathbf{w}]$ auf.
- Die Grundfläche dieses Spats ist $\|\mathbf{v} \times \mathbf{w}\|$, die Höhe dementsprechend

$$d := \frac{[\mathbf{u}, \mathbf{v}, \mathbf{w}]}{\|\mathbf{v} \times \mathbf{w}\|} .$$

- Damit wissen wir zum einen, dass sich die Geraden für kein (λ, μ) näher als auf die Entfernung d kommen können. Zum anderen wissen wir, dass sie sich tatsächlich so nahe kommen. Die Höhe d des Spats ist also die Entfernung der beiden Geraden.

Man könnte noch mehr lustige Dinge tun, was aber den zeitlichen Rahmen unseres Kurses sprengen würde.

5.2.3 Ausblick auf höhere Dimensionen

Bei der Berechnung der Hesseschen Normalform aus einem Aufpunkt und einer Basis des Richtungsraums unterscheiden sich die Vorgehensweisen im Zwei- und Dreidimensionalen nicht wirklich. Genau genommen kann die Berechnung des Normalenvektors im \mathbb{R}^2 auf den selben Trick mit dem Laplaceschen Entwicklungssatz zurückgeführt werden wie im \mathbb{R}^3 . Dieser klappt auch im \mathbb{R}^n :

- Schreibe die Basis des Richtungsraums der Hyperebene als zweite bis n -te Spalte einer Matrix $M \in \mathbb{R}^{n \times n}$.
- Dann steht mit den Bezeichnungen, die wir beim Laplaceschen Entwicklungssatz verwendet hatten der Vektor

$$\left(\det(M_{11}), -\det(M_{21}), \det(M_{31}), \dots, (-1)^{n+1} \det(M_{n1}) \right)^T$$

senkrecht auf dem Richtungsraum. Dabei bezeichnet $M_{k\ell}$ die Matrix, die aus M durch Streichen der k -ten Zeile und der ℓ -ten Spalte entsteht.

Damit lässt sich dann der Rest des Vorgehens direkt übertragen. Lediglich die Berechnung der vielen Determinanten wird bei großem n sehr aufwendig. Geht es uns nur um den Normalenvektor einer Hyperebene, ist es schon bei $n = 5$ einfacher, den Orthogonalraum mit einem 4×5 -LGS zu berechnen, als fünf 4×4 -Determinanten mit Gauß.

Dass die Determinante ein n -dimensionales Volumen beschreibt, wussten wir auch schon.

Auch die Sache mit den Quadriken lässt sich auf beliebig hohe Dimensionen erweitern, wobei sich dieses zunehmend unserem visuellen Vorstellungsvermögen entzieht, weil wir nun mal nur in einer dreidimensionalen Welt leben.

5.3 Und wie das Ganze koordinatenfrei aussieht

Dieser Abschnitt lässt sich ganz kurz halten: Da wir mit den kartesischen Koordinaten auf die Darstellung bezüglich einer Orthonormalbasis zurückgreifen, gelten alle Erkenntnisse des vorigen Abschnitts auch ohne Koordinaten. Lediglich auf die Formeln, in denen explizit Koordinaten vorkommen, müssen wir verzichten. So müssen wir etwa beim Vektorprodukt $\mathbf{a} \times \mathbf{b}$ damit auskommen, dass wir wissen, dass es auf \mathbf{a} und \mathbf{b} senkrecht steht, als Länge die Fläche des von \mathbf{a} und \mathbf{b} aufgespannten Parallelogramms hat, und dass \mathbf{a} , \mathbf{b} , $\mathbf{a} \times \mathbf{b}$ ein rechtshändiges System bilden.

Literaturverzeichnis

- [Axl15] AXLER, Sheldon: *Linear algebra done right*. 3. Auflage. Springer, 2015
- [Fis92] FISCHER, Gerd: *Analytische Geometrie*. 6. Auflage. Vieweg, 1992
- [Fis19] FISCHER, Gerd: *Lernbuch Lineare Algebra und Analytische Geometrie*. 4. Auflage. Springer Spektrum, 2019
- [SS16] SCHEID, Harald ; SCHWARZ, Wolfgang: *Elemente der Arithmetik und Algebra*. 6. Auflage. Springer Spektrum, 2016